Expert Systems with Applications 39 (2012) 10292–10302

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

A unified framework for market segmentation and its applications

Ying Liu [a,*], Melody Kiang [a,1], Michael Brusco [b,2]

[a] Department of Information Systems, College of Business Administration, California State University Long Beach, 1250 Bellflower Blvd, Long Beach, CA 90840, United States
[b] Department of Marketing, College of Business, 359 RBB, Florida State University, Tallahassee, FL 32306-1110, United States

ARTICLE INFO

Keywords:
Market segmentation
Computational complexity
Multi-criteria market segmentation
Multi-objective optimization
Pareto optimal solution

ABSTRACT

Market segmentation is a core marketing concept that is conceptually simple to define and understand, but inherently a multi-criteria problem that is hard to measure and computationally difficult in many aspects. This paper reviews the development of market segmentation techniques and identifies the computational issues of the applications of market segmentation. A multidimensional unified framework for market segmentation is proposed based on the relationship among segmentation variables, data measures, and the multi-objective optimization techniques implemented. We conduct an empirical comparison of two prominent methods: a concomitant finite mixture model and a multi-objective evolutionary algorithm. The result shows that the proposed framework helps to understand different segmentation models and solutions and to guide the development of new market segmentation solution techniques.

Published by Elsevier Ltd.

1. Introduction

Understanding and differentiating customers by their needs and responses to marketing mix plays a vital role in business management. Wind (1978) suggested that management should employ the concept of segmentation in most studies and analyze data at the segment level. The market segmentation was first defined by Smith (1956) to provide a conceptual view of an inherently heterogeneous market. Rather than offering the same marketing mix to vastly heterogeneous customers, firms should divide customers into segments and tailor the marketing mix for targeted segments to improve customer satisfaction and achieve maximum efficiency. The development of market segmentation theory has been partially contingent on the availability of marketing data, the advances in analytical techniques, and the progress of segmentation methodology (Wedel & Kamakura, 2000).

Market segmentation has been constantly under intensive investigation by researchers. There is an abundance of segmentation methods available including *K*-means clustering, hierarchical clustering, automatic interaction detection (AID), classification and regression trees (CART), conjoint analysis, clusterwise regression, hierarchical Bayes (Allenby & Rossi, 1998), finite mixture model (FMM), neural network (Vellido, Lisboa, & Meehan, 1999), simulated annealing (Brusco, Cradit, & Stahl, 2002), self-organizing map (SOM) (Lee, Suh, Kim, & Lee, 2004) and multi-objective evolutionary algorithm (MOEA) (Liu, Ram, Lusch, & Brusco, 2010). Previous research has suggested different ways to categorize market segmentation approaches and they can be classified according to four aspects: (1) application domains (Desarbo, Grewal, & Scott, 2008; Wind, 1978), such as new product development, customer retention, price discrimination, etc., (2) application objectives pertaining to behavior-oriented segmentation, the objective is to explain the differences in customer choice behavior. In the so called normative segmentation, the objective is to maximize the overall effectiveness of market resources allocated among a number of segments (Mahajan & Jain, 1978), (3) application data (Kim, Jung, Suh, & Hwang, 2005; Wind, 1978), such as benefit sought, product usage, price sensitivity, demographic data, etc., and (4) application techniques (Vriens, Wedel, & Wilms, 1996; Wedel & Kistemaker, 1989), such as conjoint analysis, clustering, etc. The four aspects are application-driven rather than theory-driven. There is relatively little work on the development of a comprehensive framework that reveals the essential characteristics of market segmentation solution techniques.

Wind (1978) and Green (1977) categorized market segmentation methods into a-priori and post-hoc approaches. In a-priori customer segmentation, the customer memberships such as young (age between 25 and 35) and middle-income ($50,000–$80,000) are determined based on intuition rather than data analyses. While in this research we consider the more common post-hoc approaches where the number and type of segments are data driven.

Wedel and Kamakura (2000) categorized market segmentation into predictive and descriptive methods. Predictive methods analyze the relationships between a set of independent variables and one or more dependent variables while descriptive methods do not distinguish between dependent and independent variables.

* Corresponding author. Tel.: +1 562 544 4271; fax: +1 562 985 4080.
E-mail addresses: yliu2@csulb.edu (Y. Liu), mkiang@csulb.edu (M. Kiang), mbrusco@fsu.edu (M. Brusco).
[1] Tel.: +1 562 985 8944; fax: +1 562 985 4080.
[2] Tel.: +1 850 644 6512; fax: +1 850 655 4098.

0957-4174/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.eswa.2012.02.161

Predictive methods allow decision makers to predict customer behaviors and the profitability of a marketing mix based on customer demographics or past behaviors. Descriptive methods are used to profile customer segments. Though helpful in showing the managerial intentions of market segmentation, this conceptually simple categorization gives few clues of the computational complexity of the market segmentation problem and the diversity and abundance of market segmentation techniques.

The lack of a comprehensive classification of market segmentation methods represents a gap between the theoretical research and the real-world applications of market segmentation. We need a framework that covers key aspects of a market segmentation problem: the managerial intentions, the properties of different segmentation data, and the characteristics of the solution techniques. The lack of such a general framework is primarily attributed to the complexity and the multi-criteria nature of market segmentation. The objectives of this research include: (1) review and identify a number of computational issues of market segmentation techniques; (2) propose a unified framework to help understand and categorize market segmentation methods; and (3) conduct an empirical comparison of two prominent methods using the framework. The rest of the paper is organized as follows: Section 2 reviews the development of market segmentation techniques that leads to the discussion of computational complexity of market segmentation problems in Section 3. A unified framework and its application to market segmentation are discussed in Section 4. We describe the empirical comparison in Section 5. The conclusions and future research directions are discussed in Section 6.

2. The development of market segmentation techniques

Clustering could be defined as a technique that groups entities similar in measured characteristics (Jain & Dubes, 1988). In this study we use this definition to refer to a number of traditional clustering methods such as hierarchical clustering, *K*-means clustering, self organization maps and neural networks that only optimize one within-segment homogeneity objective. This definition of clustering is more restrictive than the cluster analysis definition used by Punj and Stewart (1983) in their discussion of clustering methods in market segmentation. Specifically, methods such as clusterwise regression and automatic interaction detection (AID) are not clustering methods because these methods do not optimize within-segment homogeneity. A fundamental task of market segmentation is to group customers based on similarities in their needs and preferences, and clustering is a common tool for such purpose. Each segment is a group of homogeneous customers that marketers can identify, target, and communicate. In early market segmentation research, clustering was considered almost synonymous with market segmentation (Wind, 1978). As the spectrum of market segmentation expanded from customer profiling to behavioral and attitudinal studies concerning customer interaction with marketing mix, the market segmentation techniques evolved to simultaneously considering multiple sets of variables in both descriptive and predictive models. This section reviews the evolution of market segmentation techniques and discusses issues of those techniques.

2.1. Clustering techniques in descriptive market segmentation

Traditional cluster analysis is suitable for descriptive segmentation when there is only one set of variables (one segment base) need to be clustered. A segmentation base describes a certain aspect of the customer. Different segmentation bases describe different features of the customer or marketing mix. For example, the two variable sets, promotion response data and customer demographics data, are two completely different segment bases but often times are used together in marketing research studies to understand their interactions and correlations. A descriptive market segmentation model that utilizes more than one segmentation base is called a joint descriptive market segmentation (Morwitz & Schmittlein, 1992). In joint descriptive market segmentation, each segmentation base has its homogeneity objective to be optimized not only individually but also simultaneously with other segmentation bases. However the traditional clustering techniques are not designed for multi-objective optimization.

2.2. Clustering techniques in predictive market segmentation

Traditional clustering techniques only focus on segment homogeneity and generally ignore the relationships between segmentation descriptive variables and exogenous response variable(s). In predictive market segmentation, the decision makers seek to optimize both within-segment homogeneity and segment level predictability. For example, the outcome of a segmentation study will allow marketing manager to predict segment profit (response) based on customer demographics or behavioral information. The actionable segments derived from predictive market segmentation can assist marketing managers to formulate effective marketing campaign. Vriens et al. (1996) provided a review and comparison of different predictive segmentation methods. Methods such as clusterwise regression focus more on the prediction of customer responses than on the clustering of customers because they do not contain an explicit optimization objective to maximize the within segment homogeneity of the predictor variables. Clustering can be used to augment predictive segmentation methods but cannot be used alone to solve predictive segmentation problems. As a result, in both descriptive and predictive market segmentations, clustering is only a partial solution to the general model of market segmentation.

2.3. Clustering and segmentation definitions of market segmentation

The operational definitions of market segmentation come in many different forms. Market segmentation was framed as a clustering problem in early marketing research when the clustering technique was the dominant approach (Arabie & Hubert, 1994). When the research focus shifted from descriptive to response variables, market segmentation was reframed as a segmentation problem solved by solution techniques such as automatic interaction detection (AID), chi-squared automatic interaction detector (CHAID), classification and regression trees (CART), and clusterwise regression. Unlike the clustering definition that aims to maximize within segment homogeneity, these predictive methods seek to optimize an aggregated predictive function of the customer segments. Their problem definitions can be viewed as instances of a more general segmentation problem defined by Kleinberg, Papadimitriou, and Raghavan (1998). The segmentation problem is defined in a microeconomic framework that aims to maximize the total utility based on different ways to segment customers. There are two fundamental differences between a segmentation problem and a clustering problem. First of all, the segment homogeneity is not a part of the objective of the utility functions in segmentation problems. Secondly, there could be multiple utility functions in a segmentation problem reflecting the managerial intentions to treat each segment differently to maximize the total utility. In order to achieve both segment homogeneity and maximum total utility as required by market segmentation, the operational definition of market segmentation needs to fulfill dual properties of clustering and segmentation.

2.4. The multi-criteria nature of market segmentation

Marketing researchers realized that market segmentation is a multi-criteria problem because customers in a segment should not only have similar profiles (identifiability) but also respond similarly to a marketing mix (responsiveness) (Smith, 1956). Identifiability and responsiveness are two fundamental criteria for market segmentation. Throughout the evolution of market segmentation concepts, more criteria have been added. DeSarbo and Grisaffe (1998) stated that market segmentation should also satisfies accessibility, feasibility, membership identification, and profitability criteria. Wedel and Kamakura (2000) suggested to include substantiality, stability, and actionability criteria when evaluating the quality of a segmentation solution. Recently, DeSarbo and DeSarbo (2007) added projectability to the list of criteria. Cluster analysis can only address the identifiability criterion (Brusco, Cradit, & Tashchian, 2003) while the other criteria such as responsiveness, profitability, and actionability have to be handled by other solution techniques.

Market segmentation did not have an operational multi-objective definition until DeSarbo and Grisaffe (1998) proposed a combinatorial optimization framework. For a multiple objective problem, usually there is not a single "optimal" solution available because there are tradeoffs among multiple objectives. In this study we use the term "Pareto optimal" to define the acceptable optimal solutions. A solution is Pareto optimal if there is no other solution that has better values than it in all objectives. The collection of Pareto optimal solutions is called a Pareto optimal set.

For example, identifiability and responsiveness are two somehow antagonistic measures, and one specific segment assignment of customers may not be optimized for both criteria. An illustrative solution set that optimize both criteria is depicted in Fig. 1 where the two axes represent the two optimization objectives. In Fig. 1, solutions A, B, and C are Pareto optimal because there does not exist a solution whose two measures are both better (assuming smaller values are better) than those of the Pareto optimal solutions. On the contrary, solution D is not Pareto optimal because solution A is better in both measures than D. The Pareto optimal solution set forms a surface called Pareto front.

It is observed that different segmentation bases have different levels of effectiveness towards different segmentation criteria (Frank, Massey, & Wind, 1972; McDonald & Dunbar, 2004). For example, geodemographic data are stable, easy to collect hence are effective for identifiability and stability criteria, but lack actionability and responsiveness power (McCann, 1974). On the other hand, customer usage and brand loyalty data are effective for actionability criterion but possess medium support for accessibility and stability (Frank et al., 1972). In both predictive and joint descriptive segmentations, multiple segmentation bases are often used to meet multiple segmentation criteria. Therefore the selection and measurements of segmentation bases play an important role in addressing the multi-criteria nature of market segmentation.

Moreover, market segmentation is subject to various administrative and resource constraints which tend to further complicate

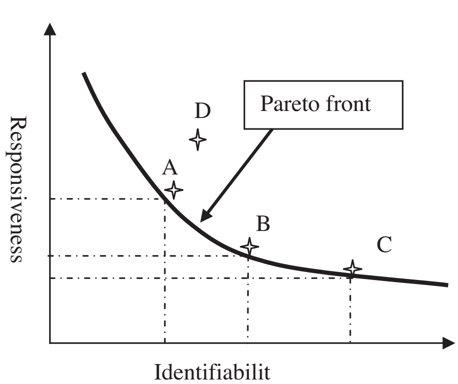

Fig. 1. Pareto optimal solutions.

the problem. Although marketing researchers are well aware of the multi-criteria nature of market segmentation problem, and many multi-objective optimization approaches (Brusco et al., 2002, 2003; Green & Krieger, 1991; Krieger & Green, 1996; Mo, Kiang, Zou, & Li, 2010) have been proposed, due to the limitations of their solution techniques, the multi-criteria market segmentation problem was either solved as a sequence of single-objective optimization problems or was transformed into a single-objective optimization problem.

3. The computational issues of market segmentation

The computational challenges of the market segmentation are often underestimated and have brought many issues in market segmentation methods. These issues include data measurement, the choice of computation models, the algorithm complexity, and multi-objective optimization methods. These issues explain many commonalities and differences among market segmentation algorithms.

3.1. The similarity measures and the clustering process

To understand a market segmentation method, one needs to identify and understand the similarity measure and the clustering process underlying the method because the clustering procedure is part of the process of developing a segmentation plan. Clustering by itself is vaguely defined and the clustering process is hard and fuzzy (Jain, Murthy, & Flynn, 1999). The vagueness lies in the measurement of so-called "homogeneity" or "similarity". The difficulties in determining the appropriate similarity measure in market segmentation have been discussed extensively in previous studies (Frank & Green, 1968; Green, Frank, & Robinson, 1967; Punj & Stewart, 1983). Skinne (1978) has identified three aspects of similarity: elevation, scatter, and shape. Generally speaking, elevation represents the means of all attributes for a single subject. Scatter describes deviation, while shape depicts the direction (up/down) of the data. There is no single distance or similarity measure that can take into account of all aspects of similarity. For example, Euclidian distance measures the elevation and scatter of the data while a correlation coefficient only measures the scatter aspect of data.

The degree of complexity of a clustering process has been well described by the impossibility theory of clustering by Kleinberg (2002). The impossibility theory stated that it is intuitive to think of three desired properties of a clustering process: scale-invariance, richness, and consistency. Scale-invariance property means that changing the unit of distance measure should not change the clustering result. Richness requires that a clustering process be able to generate all possible partitions of clustering entities. Lastly, consistency is satisfied when the clustering result remains the same when we increase the between cluster distances and decrease the within cluster distances. The impossibility theory of clustering states that there does not exist a clustering process that can satisfy all three properties simultaneously. To avoid the limitation of the existing clustering process, Handl and Knowles (2007) developed a multi-objective clustering algorithm that uses multiple criteria to measure the quality of the cluster cohesion. Their results showed that complementary measures often provide better solutions than a single measure.

3.2. Discriminative vs. generative methods

The three aspects of similarity measure and the impossibility theory of clustering process suggest that each similarity measure and clustering process pertain to certain data characteristics, and there is no generally accepted criterion for determining the quality of a clustering method. Zhong and Ghosh (2003) proposed to classify the clustering methods into discriminative and generative methods to expose the underlying data assumptions. In a discriminative method, also called distance/similarity-based method, the similarity function is defined between pairs of objects. Common discriminative measures include Euclidean distance, Pearson linear correlation, single linkage and average linkage. In a generative method the similarity is defined indirectly through the assumption of the data distribution. A generative method assumes that the overall distribution of the data is a mixture of probability distributions, each describes a different cluster (Fraley & Raftery, 1998). Common generative data assumptions are multinomial distribution for categorical data or multivariate Gaussian distribution for interval or ratio data. Each assumption has its intrinsic similarity aspects. For example, Gaussian distribution measures elevation (mean) and scatter (variance).

3.3. Computational complexity and multi-objective optimization

Aloise, Deshpande, Hansen, and Popat (2009) showed that Euclidean distance clustering problems are NP-hard, even for the two-cluster case. Kleinberg et al. (1998) proved that many optimization problems become NP-complete if they are defined in a segmentation form. Krieger and Green (1996) defined market segmentation problem as a 0–1 programming problem whose computational complexity is NP-hard. Consequently, many market segmentation problems cannot be solved in polynomial time. Existing methods either transform the market segmentation problem into a simplified form or apply heuristics to solve the problem.

Moreover, the multi-criteria nature of market segmentation raises many challenges that cannot be appropriately addressed using traditional market segmentation methods such as *K*-means and clusterwise regression that were designed for single-objective optimization. Multiple criteria can be mapped to multiple data measures and multiple optimization objectives. For example, the identifiability criterion is often operationalized as a clustering problem that uses different clustering objectives and processes. The responsiveness criterion has been defined to optimize the expected mean square prediction error, log likelihood, sum of squared errors, or weighted sum of squared errors for different data models and solution techniques (Krieger & Green, 1996). Most segmentation studies attempt to achieve both identifiability and responsiveness criteria because they represent the essential requirements of a market segmentation problem. A number of heuristics have been developed to address the multi-criteria requirement of market segmentation (Brusco et al., 2002, 2003; Green & Krieger, 1991; Krieger & Green, 1996; Liu et al., 2010), and different mechanisms (multi-stage, transformation, and multi-objective) have been proposed to tackle the thorny issues of multi-objective optimization.

4. A unified framework for market segmentation

The computational challenges and the multi-criteria nature of market segmentation suggest a multi-dimensional framework of market segmentation as depicted in Fig. 2. The three dimensions include relationship among segmentation bases, data measurement, and the solution techniques for multi-objective optimization. In practice, the relationship among segmentation bases comes up early during the planning and design stage of market segmentation. The data measurement becomes the focus during the data collection and modeling stage. Solution techniques are selected or developed by marketing researchers during the implementation stage.

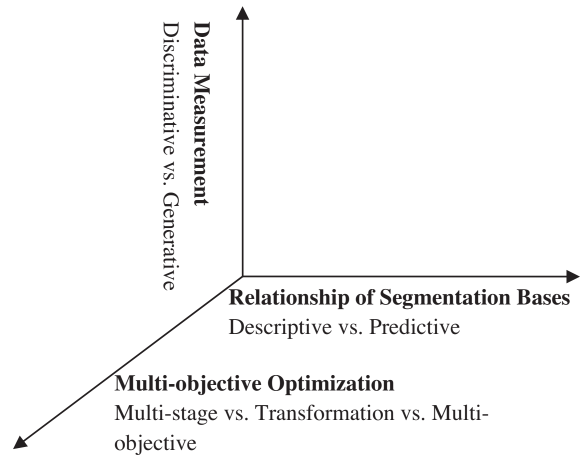

Fig. 2. A unified market segment framework.

4.1. Relationship among segmentation bases

Based on the relationship among segmentation bases, segmentation methods can be classified into descriptive or predictive approaches depending on whether the method distinguishes between independent and dependent variables. Descriptive methods include the clustering-based methods such as the variations of *K*-means (Chiu, Chen, Kuo, & Ku, 2009), hierarchical clustering techniques (Arabie & Hubert, 1994; Jain & Dubes, 1988; Punj & Stewart, 1983), *p*-median clustering (Klastorin, 1985), case-based reasoning (CBR) (Chen, Wang, & Feng, 2010) and self-organizing map (SOM) (Lee et al., 2004) methods. If the assumption of the density distribution can be made for a data set, probabilistic models are often used for clustering tasks with the additional advantages of statistical inference capability (Fraley & Raftery, 1998). Artificial Neural Networks (ANN), evolutionary approaches, and other search-based approaches recently become popular because the heuristic searching algorithms allow businesses to process large data set in a timely manner. Joint segmentation (Ramaswamy, Chatterjee, & Cohen, 1996) tries to segment customers on more than one bases or dimensions.

A variety of response modeling techniques such as clusterwise regression model (Spath, 1979), clusterwise logistic model, and mixture regression model have been widely used as predictive methods for market segmentation. Predictive methods usually result in a better predictive model for individual segment than for the population as a whole, and the within-segment homogeneity of predictors is relatively low. These methods perform well on responsiveness at the expense of less-than-satisfied segment identifiability.

4.2. Data measurement

The second dimension is based on the assumption and measurement of the segmentation data. Segment identifiability and the corresponding clustering process is an intrinsic part of a market segmentation problem. Similar to the classification of clustering methods, market segmentation methods can be classified into discriminative (or distance/similarity-based) and generative approaches according to their computational assumption of the data distribution.

Different data assumptions and data measures lead to different solutions for the segment identifiability criterion. For example, Euclidean distance is a common similarity measure for a pair of data points and it has been adopted by many discriminative methods. On the other hand, multivariate Gaussian model is based on the distribution assumptions of the data set; it is often used by discriminative methods. Each approach has its advantages and disadvantages. Discriminative-based segmentation methods are usually efficient and intuitive because they directly optimize the within-segment customer similarity. However, the use of the results has no or limited statistical inference. Contrastingly, when the data distribution is known, the results derived from generative model are usually more interpretable and allow for statistical inferences. Generative method such as finite mixture model can be computationally expensive when there are considerable number of variables and/or segments to be modeled (Wedel & Kamakura, 2000).

4.3. Multi-objective optimization

Tackling the multi-criteria nature of segmentation problems defines the third dimension of the unified framework. Existing market segmentation methods can be classified into multi-stage, transformation, and multi-objective optimization approaches. Many traditional descriptive clustering methods and predictive clusterwise methods have been modified to overcome their intrinsic shortcomings in solving multi-criteria market segmentation problems. Vriens et al. (1996) provided an extensive conceptual review and empirical comparison of the results from modifications of clusterwise regression methods and latent class models. These methods use either multi-stage or transformation approaches to make tradeoffs between the identifiability and responsiveness criteria of market segmentation.

The multi-stage approach allows us to deal with one criterion at a time in sequence. The modified *K*-means method proposed by Krieger and Green (1996) is a typical multi-stage segmentation method that first employs a *K*-means clustering algorithm to optimize segment identifiability, then uses a heuristic algorithm to improve the variance of response variable of the segments with a threshold on the increasing within-segment heterogeneity. Since the approach is based on optimizing a local/single objective at each stage, and there is no explicit information sharing between stages, the problem of multi-criteria segmentation is addressed in a sub-optimal way. Moreover, the order the objectives applied often matters and the results may vary significantly.

The transformation approach transforms and combines multiple objectives into a single one. As a result, many existing single objective optimization methods can be applied. DeSarbo and Grisaffe (1998) defined a total utility function to integrate the different criteria and proposed a single-objective combinatorial optimization approach. Brusco et al. (2002, 2003) used the transformation approach to solve two-criteria and three-criteria segmentation problems. Most finite mixture models also fall into the transformation category where multiple criteria are integrated into a single maximum likelihood function. Although the simplicity of the transformation approach has helped to gain its popularity, the approach has many drawbacks yet to be resolved. It is usually difficult to define an appropriate total utility or weighted sum function. Additionally, the transformation procedure may put unnecessary limitations on the search space. A weighted sum transformation of a multi-objective optimization problem may only identify a subset of all possible Pareto-optimal solutions (Ehrgott & Gandibleux, 2000).

Compared with multi-stage and transformation approaches, the multi-objective optimization approach has two distinctive advantages. First of all, there is no need to change the multi-criteria problem definition. Secondly, it generates a set of Pareto optimal solutions that gives a holistic view of the possible solution spaces. Nonetheless, the multi-objective approach has not been widely used in market segmentation due to the complexity of both the market segmentation problem and the multi-objective optimization solution techniques.

4.4. Applications of the unified framework to market segmentation

Since the three dimensions of the proposed framework are orthogonal, it is easy to categorize a market segmentation method and understand the managerial intentions (profiling or predicting), the data assumptions, and the solution techniques of that method using the framework. Table 1 summarizes each component of the proposed framework. For example, clusterwise regression segmentation method (Spath, 1979) is a predictive, generative, and transformation method. Modified *K*-means (Krieger & Green, 1996) is a descriptive, discriminative, and multi-stage method. Finite mixture models (FMM) are mostly used for predictive market segmentation, but there are cases where FMM models are applied to descriptive market segmentation problems (Andrews, Brusco, Currim, & Davis, 2010).

Nonetheless, there are many hybrid market segmentation methods resulted from the multi-criteria nature of market segmentation. Different data assumptions can be assumed simultaneously in a transformation method or a multi-objective method. For example, DeSarbo and Grisaffe (1998) proposed a general multi-criteria programming paradigm that combines the discriminative data assumption (the *K*-means type variance) and the generative data assumption (the clusterwise-linear-regression model) in a weighted-sum total utility function. The predictive simulated annealing heuristic (SAH) segmentation method (Brusco et al., 2003) is a transformation method that uses both discriminative and generative data assumptions. Secondly, it is also possible that different types of the same dimension are used in different optimization stages as suggested by Krieger and Green (1996). Consequently, the possible types of market segmentation methods are not a simple combination of the one method type from each dimension. Table 2 categorizes the possible orthogonal combination of market segmentation methods based on the proposed framework.

The classification helps marketing researchers to understand the characteristics of existing methods. The framework explicates the underlying assumptions of a specific method and suggests alternative solution techniques. For example, a generative method is often used for predictive market segmentation, nevertheless it has also been applied to descriptive market segmentation problems for customer profiling (Wedel & Kamakura, 2000). Vis-à-vis, a discriminative method has also been applied to predictive segmentation problems (Brusco et al., 2002, 2003; Krieger & Green, 1996). Therefore it might be helpful for a decision maker to explore both methods and compare their solution results before settling on one approach.

The proposed framework helps the decision makers to visualize the relationship among the key components of the segmentation problem and suggest alternatives during the segmentation process. Decision makers define the segmentation model (descriptive/predictive) to reflect the business requirements. Krieger and Green (1996) discussed cases where the same segmentation problem can be formulated as either a descriptive or a predictive model that could lead to different segmentation strategies. In the data collection stage, the knowledge and assumptions of the segmentation data have direct impact on the selection of discriminative/generative segmentation method. The proposed framework can help to emphasize the importance of the data assumptions and draw the decision maker's attention to the possible consequences of the different data measurements on the segmentation results. Additionally, the framework can assist in selecting suitable solution techniques for the problem. When the weights of each criterion can be easily assessed or the total utility function is known, the transformation solution technique maybe a viable solution due to the availability of many highly efficient single objective optimization methods. On the

Table 1
Summary of the framework components.

	Definition	Features	Examples
Relationship of segmentation bases			
Descriptive model	Does not distinguish between independent and dependent variables. May use more than one segmentation bases.	Often used for profiling customers from their segmentation bases Optimizes segmentation homogeneity of each segmentation base	*K*-means; self-organizing map (SOM); joint descriptive segmentation methods (Brusco et al., 2002; Krieger & Green, 1996); finite mixture model (FMM) (Dillon & Kumar, 1994)
Predictive model	Has independent and dependent variables. Use different regression models (linear or logistic)	Use to predict marketing responses from predictor variables. Optimize the regression model criteria	Clusterwise regression (Spath, 1979); predictive finite mixture model (FMM) (Wedel & Desarbo, 1995)
Data measurement			
Discriminative measurement	Measures the similarity/ distances between pairs of objects	Is easy to define and implement efficient algorithms but the results does not offer statistical inferences	*K*-means; SOM; finite mixture model (FMM) (Dillon & Kumar, 1994)
Generative measurement	Assumes one or more distribution models for a data set	Allows statistical inference. Provide good results if assumptions match data set. Is not efficient and scalable for large data sets and/ or with a large number of variables	Clusterwise regression (Spath, 1979); finite mixture model (FMM) (Wedel & Desarbo, 1995)
Multiobjective optimization			
Multistage method	Optimizes one object at a time	Allows the use of many single objective optimization methods. Does not share information between stages. Generates only one solution	Modified *K*-means (Krieger & Green, 1996)
Transformation method	Transform multiple objectives into a single objective problem	Allows the use of many single objective optimization methods. Needs to combine multiple objectives into one. May lose the global optimal solution due to the transformation. Generates only one optimal solution with subjectively assigned objective weights	Combinatory optimization approaches (DeSarbo & Grisaffe, 1998); most finite mixture model methods (Wedel & Kamakura, 2000); simulated annealing heuristic (SAH) (Brusco et al., 2002)
Multiobjective method	Optimizes multiple objectives simultaneously and generates the Pareto-optimal solutions	Simultaneously optimizes multiple objectives without loss of information. May find all potential Pareto optimal solutions. Needs to develop multiobjective optimization algorithms. Requires multiobjective analysis	Multiobjective evolutionary algorithm (Liu et al., 2010)

other hand, addressing one criterion at each stage using the multistage solution technique can sometimes generate good segmentation results when there is no significant interaction among the criteria processed at different stages. The emerging multiobjective optimization solution techniques can serve as effective alternative methods to multistage and transformation techniques when the two methods fail to perform.

The three dimensions in the framework are independent of each other and can be assembled in all possible combinations. The framework facilitates the identification of the research gap in existing methods and suggests possible solution techniques. For example, in Table 2, the two categories that do not have a solution technique are the descriptive, generative, multi-objective optimization category and the predictive, discriminative, multistage category.

Market segmentation, from a computational point of view, is a combination of clustering and segmentation problems that are both NP hard, most existing segmentation methods are either multi-stage or transformation methods thus the design of an efficient multi-objective algorithm specially tailored to market segmentation problems will benefit the field. Recently, Liu et al. (2010) developed a multi-objective market segmentation method using a multi-objective evolutionary algorithm called MMSEA. Unlike the transformation or the multi-stage approaches, the MMSEA algorithm directly optimizes both identifiability and responsiveness criteria and generates a set of Pareto optimal solutions. It is a hybrid of the generative method and the discriminative method in its data measurement. It simultaneously optimizes a regression prediction model and the segment homogeneity. The model may assume normal or logistic distribution of data while the segment homogeneity is measured by the Euclidean distances among customers.

The modified MMSEA algorithm is a descriptive, discriminative, and multi-objective optimization method. In the following case study, we present the modified MMSEA method and compare its performance with that of a popular transformation method in market segmentation.

5. A case study

In this section we compare two state-of-the-art market segmentation methods, the concomitant FMM and the modified MMSEA methods, and apply them to a real-world market segmentation problem. Concomitant FMM is considered one of the most popular predictive transformation methods that simultaneously optimize segment homogeneity (identification) and predictive (responsiveness) performance (Dayton & MacReady, 1988). The MMSEA method, recently developed by Liu et al. (2010) is a multi-objective optimization method for market segmentation that can generate a set of Pareto optimal solutions. It used both discriminative and generative data measurements for predictive market segmentation. For comparison purpose, we modify the MMSEA algorithm to use only discriminative data measurement in a joint descriptive segmentation model. The results show that though the Concomitant FMM and the modified MMSEA are applied to the same market segmentation data with identical high-level market segmentation criteria, there are significant differences in their data assumptions, solution techniques, segmentation results, and managerial implications.

Table 2
Categories of the segmentation methods.

Relationships of segmentation bases	Data measurements	Multiobjective optimization	Examples
Descriptive	Discriminative	Multistage	Step-wise conjoint analysis (Green & Krieger, 1991); Modified cluster-based segmentation (Krieger & Green, 1996)
Descriptive	Discriminative	Transformation	Descriptive simulated annealing heuristic (Brusco et al., 2002)
Descriptive	Discriminative	Multiobjective optimization	Modified MMSEA (our proposed method)
Descriptive	Generative	Multistage	Clustering with Latent Class (Dillon & Kumar, 1994)
Descriptive	Generative	Transformation	Stochastic unfolding mixture model (Wedel & Desarbo, 1996)
Descriptive	Generative	Multiobjective optimization	Not available
Predictive	Discriminative	Multistage	Not available
Predictive	Discriminative	Transformation	FMM for generalized linear model (Wedel & Desarbo, 1995)
Predictive	Discriminative	Multiobjective optimization	Multiobjective evolutionary algorithm (MMSEA) (Liu et al., 2010)
Predictive	Generative	Multistage	Two-stage clusterwise method
Predictive	Generative	Transformation	Multi-criterion clusterwise regression (Brusco et al., 2003)
Predictive	Generative	Multiobjective optimization	Multiobjective evolutionary algorithm (MMSEA) (Liu et al., 2010)

5.1. The market segmentation model

The goal of this study, which is a typical of market segmentation, is to identify segments among customers based on customer profit and household-level sociodemographic data. Segments formed on the basis of sociodemographic data are easy to identify and communicate with. The response variable is the firm profit of a customer in a 6-month period. The data set is a random sample of 1500 customers from about 95,000 members of a large retailer's premium club. Table 3 gives the descriptive statistics of the raw segmentation variables.

The selection of profit as the response variable is often considered more cost-effective than other variables and is a good choice for normative segmentation. To simplify the problem discussion, we consider only the essential identifiability and responsiveness criteria in this study. A simple yet effective identifiability measure is the within segment omega squared (WSOS) of customer descriptive variables. WSOS is the ratio of the within-segment sum of squares to the total sum of squares. A smaller WSOS value represents better segment homogeneity. The responsiveness criterion can be formed differently based on different managerial concerns. The two segmentation methods evaluated here differ with respect to the response criterion. If a decision maker is interested in the prediction of responses from the explanatory variables, they often use linear or logistic regression models (Desarbo & Ramaswamy, 1994), and the total residual sum of squares (TRSS) can be used as the optimization objective in the regression model. TRSS is the sum of all segment level residual sum of squares. A small TRSS value represents a better predictive result. The concomitant FMM optimize the segment-level regression model thus indirectly optimizes the TRSS objective. Contrastingly, if the decision makers are interested in optimizing response homogeneity for each segment, they can optimize the WSOS value of the response variables. Krieger and Green (1996) use this objective in a bank case where the managers intended to segment the market for cross selling and setting profit objectives. Similarly, the modified MMSEA method tries to minimize the WSOS of customer profit for the responsiveness criterion. In the segmentation process, the response variable (profit) was transformed using a natural logarithmic function, $\ln(\cdot)$, to reduce its skewness. The calculation of TRSS uses ln(Profit) as the dependent variable and the raw sociodemographic data as the independent variables. In the segmentation process, we normalize the explanatory variables to eliminate the possible biases due to differences in data scales as suggested by Milligan and Cooper (1988). WSOS(Y) uses the same ln(Profit) as the response variable.

Table 3
Descriptive statistics of the segment variables.

Attributes	Min	Max	Mean	Std. dev.
Profit	0.84	117.22	14.49	11.98
Adult	1	4	2.17	.71
Age	18	86	46.80	12.67
Children	0	1	.57	.50
Gender	0	1	.72	.45
Income	1	9	6.28	2.05
Marital status	0	1	.79	.411
Working woman	0	1	.48	.50

Therefore the segmentation goal of this study is to assign customers to identifiable (smaller WSOS of the explanatory variables for both methods) segments that will either improve the predictability of the explanatory variables (smaller TRSS in the concomitant FMM) or the homogeneity among the response variables (smaller WSOS of the customer response variable in the modified MMSEA). In the following, we formally define the segmentation model used in both the concomitant FMM and the modified MMSEA methods. Let x_{ij} = the value of attribute j for customer i; $i = 1,\ldots,I$, I is the number of customers; $j = 1,\ldots,J$, J is the number of attributes in a segmentation base (the predictors); $\bar{x}_{j_k}$ = the mean of attribute j in segment k; $k = 1,\ldots,K$; K is the number of segments; $I(k)$ = the set of customers in segment k; $\bar{x}_j$ = the mean of attribute j for all customers; y_i = the value of the response variable for customer i; $\bar{y}_k$ = the mean of the response variable for customers in segment k; $\bar{y}$ = the mean of the response variable for all customers; f_i = the predicted value of customer i in its segment-level linear regression model.

The three optimization objectives are defined as follows:

$$\mathrm{WSOS}(X) = \frac{\sum_{k=1}^{K}\sum_{j=1}^{J}\sum_{i\in I(k)}(x_{ij}-\bar{x}_{j_k})^2}{\sum_{j=1}^{J}\sum_{i=1}^{I}(x_{ij}-\bar{x}_j)^2}$$
(used as the first objective for both methods),

$$\mathrm{WSOS}(Y) = \frac{\sum_{k=1}^{K}\sum_{i\in I(k)}(y_i-\bar{y}_k)^2}{\sum_{i=1}^{I}(y_i-\bar{y})^2}$$
(used as the second objective for the modified MMSEA),

$$\mathrm{TRSS} = \sum_{k=1}^{K}\sum_{i\in I(k)}(y_i-f_i)^2 \text{(used as the second objective for the concomitant FMM).}$$

5.2. A predictive, generative, and transformation method

FMM is a generative approach that assumes the data are a mixture of subgroups with different density functions (Dillon & Kumar, 1994). It implements a transformation mechanism to combine the liner regression modeling and customer partitioning in one process. The Concomitant FMM has been applied in many market segmentation studies (Wedel & Desarbo, 2002). Grün and Leisch (2008) developed an open source concomitant FMM software module called FlexMix, implemented using the open source R statistical software package.

The FlexMix module utilizes a mixture of regression models, which consists of a finite number of parametrically distributed segments. A weight is assigned to each segment to represent the a-priori probability for an observation to belong to that segment. To optimize both the performance of the segment-level regression models and the segment homogeneity, the weight is assumed to be dependent on the predictors. The distribution model of the above segmentation problem is formulated as $f(y|x,\phi) = \sum_{k=1}^{K}\pi_{k|x}f_k(y|x,\beta_k)$, where y denotes the dependent variable (the profit) and x denotes the vector of predictors (the sociodemographic variables). ϕ denotes all parameters of the mixture density $f(\cdot)$. The density function of segment k is $f(\cdot)$, where β_k denotes its regression parameters. The sub-model $\pi_{k|x}$ denotes the segment weight. This sub-model represents segment homogeneity requirement of the segmentation model. It holds $\forall x$ that $\sum_{k=1}^{K}\pi_{k|x} = 1$ and $\pi_{k|x} > 0$, $\forall k$. In this mixture model, the vector of the socio-demographic variables (x) plays the dual roles of predictors and concomitant variables. The sub-model $\pi_{k|x}$ often uses a logistic function $\pi_{k|x} = \frac{\exp\left(\sum_{j=1}^{j} y_{kj}x_{ij}\right)}{\sum_{k=1}^{k}\exp\left(\sum_{j=1}^{j} y_{kj}x_{ij}\right)}$, where the parameter γ_{kj} denotes the impact of predictor j on the prior probability of segment k. The optimization goal is to maximize the likelihood function $L(\phi|y) = \sum_{i=1}^{I}\ln\left(\sum_{k=1}^{K}\pi_{k|x}f_k(y|x,\beta_k)\right)$. The identifiability and the predictive objectives are transformed using generative data models and combined into this likelihood function. The FlexMix implements a variant of the Expectation-Maximization (EM) algorithm as the problem solver. The EM algorithm maximizes this likelihood function and indirectly minimizes WSOS(X) and TRSS.

5.3. A descriptive, discriminative and multiobjective method

The modified MMSEA uses descriptive segmentation and discriminative data measures for the same segmentation problem. Instead of minimizing the total residual sum of squares (TRSS) of the segment regression models, it uses the within segment omega squared (WSOS) of the response variable, ln(Profit), to represent the response objective. The two optimization objectives are to minimize WSOS(X) and WSOS(Y). It is a joint descriptive segmentation model that simultaneously optimizes segment homogeneity for two discriminative data measures. The defined objectives do not consider the relationship between the two segmentation bases because the two objectives are independently defined. Similar operational definitions have been used in other segmentation studies (Brusco et al., 2002; Krieger & Green, 1996).

Since the underlying multi-objective evolutionary algorithm is a meta-heuristic algorithm, we are able to modify the MMSEA algorithm to fit the defined market segmentation model in this evaluation. The optimization objectives are changed to minimize both WSOS(X) and WSOS(Y), and we replace clusterwise regression with K-means clustering to generate the initial solution set. To generate a more diversified initial solution set, we apply K-means clustering to two dimensional data sets as wells as the whole data set containing both X and Y. Combining the two dimension data sets together generates a new set of initial solutions that considering both data sets simultaneously. Our experiment shows that the added initial solution set helps to reduce the time it takes for the evolutionary algorithm to converge. In the original MMSEA, the regression model calculation is an expensive operation to calculate the fitness of every solution generated in the evolutionary process. We replace the linear regression model with the calculation of WSOS(Y) in the modified MMSEA method that has significantly reduced the processing time of the model. We modify the solution density control algorithm and the fitness function based on the new optimization objectives. In our evaluation, the modified MMSEA algorithm runs more than two times faster than the original MMSEA algorithm.

5.4. Comparison of the two methods

Both the modified MMSEA and the Concomitant FMM algorithms are applied to the same segmentation data set to optimize segment identifiability and responsiveness. The concomitant FMM algorithm generates one solution at a time while the modified MMSEA algorithm generates a set of Pareto optimal solutions for each specified number of segments. In practice it is impossible to find the true Pareto optimal solutions for a NP-hard multi-criteria market segmentation problem. A solution set of the modified MMSEA only represents an approximation to the real Pareto front. We ran the modified MMSEA algorithm to generate a total of 600 solutions which consist of three sets of 200 Pareto optimal solutions for the three different number of segments (3, 4, and 5). The results of both methods are depicted in Figs. 3 and 4.

Fig. 3 compares the solutions using the descriptive segmentation objectives. The horizontal axis is the within segment omega squared (WSOS) of the standardized sociodemographic variables. The vertical axis is the WSOS of the response variable ln(profit). From the top-right to the bottom-left of Fig. 3, the three dotted lines are the three sets of Pareto optimal solutions of the modified MMSEA algorithm and are labeled as new 3-seg, new 4-seg and new 5-seg. Each set contains 200 solutions and a total of 600 solutions were generated by a single run of the modified algorithm. For each desired number of segments, we ran the concomitant FMM model five times and presented the best result in Fig. 3. It shows that the modified MMSEA tends to generates better solutions be-

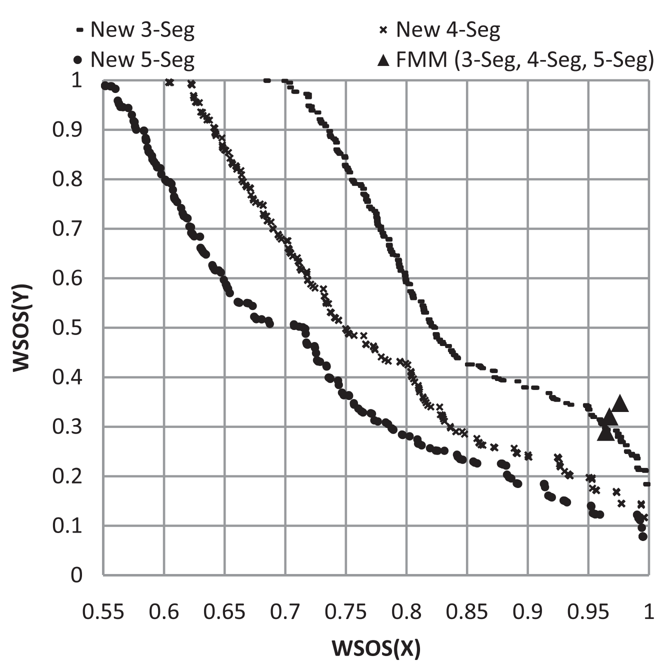

Fig. 3. Comparison of the homogeneity measures between the modified MMSEA and the concomitant FMM methods.

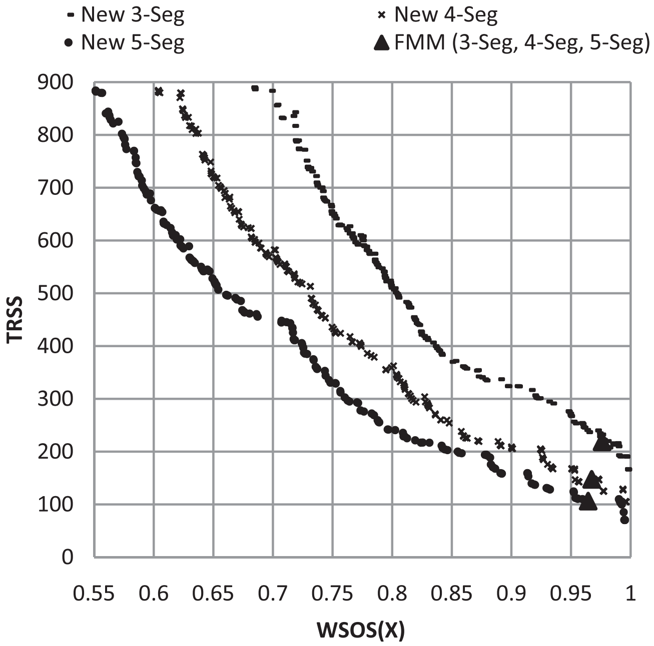

Fig. 4. Comparison of the regression measures between the modified MMSEA and the concomitant FMM methods.

cause it optimizes both objectives directly while the FMM method optimizes different objectives.

In Fig. 4 we compare the two methods using the predictive segmentation objectives. The vertical axis is the total residual sum of squares (TRSS) of all segment-level linear regression models while the horizontal axis is WSOS(X), same as in Fig. 3. Both methods do not optimize the TRSS values directly. The TRSS values are calculated afterwards from segment membership of the solutions from both methods. As a result, not all modified MMSEA solutions are Pareto optimal based on the new set of objectives. It is interesting to notice that the concomitant FMM solutions are very close to the lines formed by the set of corresponding modified MMSEA solutions. For comparison purpose, we use a solution in the middle of the Pareto front of the 3-segment Pareto front of the modified MMSEA as the selected solution and compare it with the 3-segment solution of the concomitant FMM method. The objective values and the segment level linear regression models of the two solutions are summarized in Table 4.

The 3-segment modified MMSEA solution represents a balanced trade off of the two objectives while the concomitant FMM solution is biased towards the responsiveness criterion. The concomitant FMM method is a predictive method that directly optimizes the segment regression models. The method renders much better segment R^2 values than the modified MMSEA solution. Segment 2 of the concomitant FMM solution has a strong relationship between the response and the explanatory variables. Variables such as gender, marital status, adult, income, and age are significant in predicting the customer profit. The modified MMSEA solution does not consider the relationship between the two sets of variables. The low R^2 values show a weak relationship in all segment regression models. However, the variance reduced by the regression model is not as significant as it looks when we consider the variance reduced by both segmentation and regression. For example, in Table 4, the largest R^2 value of the concomitant FMM solution is in segment 2 that explains a variance value of 44.02, which represents only 4.7% of the total sum of squares value of 930.80.

The advantage of the descriptive method is the segment homogeneity of the sociodemographic variables. It has a smaller WSOS(X) value (.840) than the predictive concomitant FMM method (.977). Table 5 shows the segment means and standard deviations of all variables from both solutions. Compare with those in the concomitant FMM method, the adult and marital status variables in the modified MMSEA solution have smaller standard deviations and more distinct segment means. Segment 1 of the descriptive method solution has high profit mean and is distinguished by a smaller number of adults (1.39), a lower income level (4.86), and single marital status. It is a highly identifiable segment. The concomitant FMM solution does not have any highly identifiable sociodemographic variable.

The modified MMSEA method gives a holistic view of the possible solutions while the concomitant FMM method only generates a single solution with unknown tradeoffs between the two objectives. The Pareto optimal solution set allows decision makers to select a solution with desired objective values easily. The tradeoffs are visually illustrated by the shape of the Pareto front. The segment sizes and other statistics of a solution can also be used in solution selection.

From the computational perspective, the concomitant FMM method uses EM algorithm that guarantees convergence but may result in local optimum. Multiple runs with different starting values are suggested when using an EM algorithm. In this study, we ran FlexMix five times for each of the 3, 4, and 5 segment solutions.

Table 4
Linear regression models of the 3-segment solutions.

	The concomitant FMM solution			The modified MMSEA solution		
WSOS(X)	.977			.840		
TRSS	218.36			383.36		
WSOS(Y)	.348			.439		
	Seg 1	Seg 2	Seg 3	Seg 1	Seg 2	Seg 3
Segment size	446	462	592	234	563	703
R-squared	.264	.717	.212	.012	.091	.058
Predicted sum of squares	31.543	44.015	30.311	1.063	14.315	9.415
Residual sum of squares	88.084	17.379	112.898	86.340	143.735	153.289
Total sum of squares	119.627	61.394	143.209	87.403	158.050	162.705
Intercept	.194	**.282**	**1.857**	**2.721**	**1.293**	3.**296**
Adult	**.116**	**.162**	**.175**	−.064	.039	−.014
Age	**.006**	**.016**	**.009**	.000	−.003	**.000**
Children	−.096	.020	.023	.003	−**.142**	−**.147**
Gender	**.385**	**.373**	**.241**	−.102	**.248**	**.174**
Income	**.105**	**.123**	**.070**	.005	.009	−.020
Marital status	−**.202**	−**.183**	−**.168**	n/a	**.282**	−.215
Working woman	−.042	.010	.010	.125	−.040	−.073

Notes: The top panel contains the objective values. The middle panel contains the segment model summaries. The bottom panel contains the within-segment slope coefficient (significant at 0.01 level shown in bold). The marital status coefficient of segment 1 in the new solution is not available because it is a constant.

Table 5
The within-segment variable means and the standard deviations (in parentheses).

	The concomitant FMM solution			The modified MMSEA solution		
	Seg 1	Seg 2	Seg 3	Seg 1	Seg 2	Seg 3
ln(Profit)	1.51 (.52)	2.37 (.36)	3.05 (.49)	2.62 (.61)	1.63 (.53)	2.90 (.48)
Adult	2.26 (.72)	2.25 (.70)	2.03 (.67)	**1.39 (.61)**	**2.23 (.63)**	**2.37 (.62)**
Age	51.44 (12.98)	47.48 (12.02)	42.77 (11.61)	44.81 (13.16)	49.39 (13.66)	45.39 (11.29)
Children	.36 (.50)	.61 (.49)	.53 (.50)	.36 (.48)	.56 (.50)	.63 (.48)
Gender	.77 (.42)	.76 (.43)	.65 (.48)	.77 (42)	.74 (.44)	.69 (.47)
Income	6.61 (1.89)	6.58 (1.93)	5.80 (2.16)	4.86 (2.09)	6.16 (1.88)	6.86 (1.91)
Marital status	.80 (.40)	.83 (.38)	.74 (.44)	**.00 (.00)**	**.89 (.32)**	**.96 (.16)**
Working woman	.49 (.50)	.50 (.50)	.46 (.50)	.56 (.50)	.49 (.50)	.46 (.50)

Notes: The bold numbers highlight the highly identifiable segment attribute values.

The average time to generate a solution is 21 min and 22 s. The modified MMSEA guarantee convergence and facilitates escape from local optima (Coello, Veldhuizen, & Lamont, 2002). It takes 57 min and 46 s to generate 600 solutions in a single run. It averages to around 5.78 seconds for each solution.

6. Discussion and future research directions

In this study, we propose a unified framework that encompasses the computational properties of the market segmentation problem. The framework can assist decision makers to analyze and evaluate market segmentation problems in a structured manner through the three orthogonal perspectives: segmentation variables, data measures, and the solution techniques implemented. The framework can support the decision maker in various ways which include: (1) the framework provides a structured categorization of market segmentation based on the three important key components that can assist the decision makers to better understand the characteristics of each method. (2) The framework helps the decision makers to understand the relationship among the key components of the segmentation problem through the three framework dimensions and suggest alternative approaches to formulate/analyze the problem at each stage of the segmentation process and (3) the framework suggests possible solution techniques for future research direction.

The framework highlights the multi-criteria nature of market segmentation, which is a long standing problem facing marketing researchers. The computational complexity and the lack of multi-objective optimization methods suggest that more efficient multi-objective optimization segmentation approaches should be developed and evaluated. Tabu/scatter search, ant systems, and memetic algorithms are commonly used multi-objective optimization methods (Coello et al., 2002). Their applications to market segmentation are potential future research topics.

A multi-criteria market segmentation problem has a set of Pareto-optimal solutions that requires the decision makers to select one best solution to implement. Though many multi-criteria decision making methods (Chan, Cheng, & Hsien, 2011) can be used to select a solution from the Pareto optimal solution set, the selection is a non-trivial process. Yun, Nakayama, Tanino, and Arakawa (2001) proposed a multi-objective optimization method that combines generalized data envelopment analysis (GDEA) and genetic algorithm to yield fewer desirable solutions. It is a future research topic to develop similar market segmentation methods that can simultaneously optimize multiple objectives and suggest segmentation solutions.

References

Allenby, G. M., & Rossi, P. E. (1998). Marketing models of consumer heterogeneity. *Journal of Econometrics, 89*, 57–78.
Aloise, D., Deshpande, A., Hansen, P., & Popat, P. (2009). NP-hardness of Euclidean sum-of-squares clustering. *Machine Learning, 75*, 245–248.
Andrews, R. L., Brusco, M. J., Currim, I. S., & Davis, B. (2010). An empirical comparison of methods for clustering problems: Are there benefits from having a statistical model? *Review of Marketing Science, 8*, 1–32.
Arabie, P., & Hubert, L. (1994). *Cluster analysis in marketing research*. Malden, MA: Blackwell Business.
Brusco, M. J., Cradit, J. D., & Stahl, S. (2002). A simulated annealing heuristic for a bicriterion partitioning problem in market segmentation. *Journal of Marketing Research, 39*, 99–109.
Brusco, M. J., Cradit, J. D., & Tashchian, A. (2003). Multicriterion clusterwise regression for joint segmentation settings: An application to customer value. *Journal of Marketing Research, 40*, 225–234.
Chan, C.-C. H., Cheng, C.-B., & Hsien, W.-C. (2011). Pricing and promotion strategies of an online shop based on customer segmentation and multiple objective decision making. *Expert Systems with Applications, 38*, 14585–14591.
Chen, Y.-K., Wang, C.-Y., & Feng, Y.-Y. (2010). Application of a 3NN+1 based CBR system to segmentation of the notebook computers market. *Expert Systems with Applications, 37*, 276–281.
Chiu, C.-Y., Chen, Y.-F., Kuo, I.-T., & Ku, H. C. (2009). An intelligent market segmentation system using *k*-means and particle swarm optimization. *Expert Systems with Applications, 36*, 4558–4565.
Coello, C. A. C., Veldhuizen, D. A. V., & Lamont, G. B. (2002). *Evolutionary algorithms for solving multi-objective problems*. New York, NY: Kluwer Academic Publishers.
Dayton, C. M., & MacReady, G. B. (1988). Concomitant variable latent class models. *Journal of American Statistical Association, 83*, 173–178.
Desarbo, W. S., & DeSarbo, C. F. (2007). A generalized normative segmentation methodology employing conjoint analysis. In A. Gustafsson, A. Herrmann, & F. Huber (Eds.). *Conjoint measurement: Methods and applications* (4th ed.) (Vols. 321–345). Berlin: Springer.
Desarbo, W. S., Grewal, R., & Scott, C. J. (2008). A clusterwise bilinear multidimensional scaling methodology for simultaneous segmentation and positioning analyses. *Journal of Marketing Research, 45*, 280–292.
DeSarbo, W. S., & Grisaffe, D. (1998). Combinatorial optimization approaches to constrained market segmentation: An application to industrial market segmentation. *Marketing Letters, 9*, 115–134.
Desarbo, W. S., & Ramaswamy, V. (1994). CRISP: Customer response based iterative segmentation procedures for response modeling in direct marketing. *Journal of Direct Marketing, 8*, 7–20.
Dillon, W. R., & Kumar, A. (1994). Latent structure and other mixture models in marketing: An integrative survey and overview. In R. P. Bagozzi (Ed.), *Advanced methods for marketing research* (pp. 353–388). Malden, MA: Blackwell Business.
Ehrgott, M., & Gandibleux, X. (2000). A survey and annotated bibliography of multiobjective combinatorial optimization. *OR Spektrum, 22*, 425–460.
Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering methods? Answers via model-based cluster analysis. *The Computer Journal, 41*, 578–588.
Frank, R. E., & Green, P. E. (1968). Numerical taxonomy in marketing analysis: A review article. *Journal of Marketing Research, 5*, 83–94.
Frank, R. E., Massey, W. F., & Wind, Y. (1972). *Market segmentation*. Englewood Cliffs, New Jersey: Prentice-Hall Inc..
Green, P. E. (1977). A new approach to market segmentation. *Business Horizons, 20*, 61–73.
Green, P. E., Frank, R. E., & Robinson, P. J. (1967). Cluster analysis in test market selection. *Management Science, 13*, B387–400.
Green, P. E., & Krieger, A. M. (1991). Segmenting markets with conjoint analysis. *Journal of Marketing, 55*, 20–31.
Grün, B., & Leisch, F. (2008). FlexMix Version 2: Finite mixtures with concomitant variables and varying and constant parameters. *Journal of Statistical Software, 28*, 1–35.
Handl, J., & Knowles, J. (2007). An evolutionary approach to multiobjective clustering. *IEEE Transactions on Evolutionary Computation, 11*, 56–76.
Jain, A. K., & Dubes, R. C. (1988). *Algorithms for clustering data*. Englewood Cliffs, NJ: Prentice Hall.
Jain, A. K., Murthy, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys, 31*, 264–323.
Kim, S.-Y., Jung, T.-S., Suh, E.-H., & Hwang, H.-S. (2005). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert Systems with Applications, 31*, 101–107.

Klastorin, T. D. (1985). The p-median problem for cluster analysis: A comparative test using the mixture model approach. *Management Science, 31*, 84–95.
Kleinberg, J. (2002). An impossibility theorem for clustering. In *Proceedings of the 15th conference on neural information processing systems (NIPS). Vancouver, Canada*.
Kleinberg, J., Papadimitriou, C., & Raghavan, P. (1998). Segmentation problems. In *Proceedings of the thirtieth annual ACM symposium on theory of computing* (pp. 473–482). Dallas, Texas: ACM Press.
Krieger, A. M., & Green, P. E. (1996). Modifying cluster-based segments to enhance agreement with an exogenous response variable. *Journal of Marketing Research, 33*, 351–363.
Lee, S. C., Suh, Y. H., Kim, J. K., & Lee, K. J. (2004). A cross-national market segmentation of online game industry using SOM. *Expert Systems with Applications, 27*, 449–570.
Liu, Y., Ram, S., Lusch, R. F., & Brusco, M. (2010). Multicriterion market segmentation: A new model, implementation, and evaluation. *Marketing Science, 29*, 880–894.
Mahajan, V., & Jain, A. K. (1978). An approach to normative segmentation. *Journal of Marketing Research, 15*, 338–345.
McCann, J. M. (1974). Market segment response to the marketing decision variables. *Journal of Marketing Research, 11*, 399–412.
McDonald, M., & Dunbar, I. (2004). *Market segmentation: How to do it, how to profit from it*. Burlington, MA: Elsevier Butterworth–Heinemann.
Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification, 5*, 181–204.
Mo, J., Kiang, M. Y., Zou, P., & Li, Y. (2010). A two-stage clustering approach for multi-region segmentation. *Expert Systems with Applications, 37*, 7120–7131.
Morwitz, V. G., & Schmittlein, D. C. (1992). Using segmentation to improve sales forecasts based on purchase intent: Which intenders actually busy"? *Journal of Marketing Research, 29*, 391–405.
Punj, G., & Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestion for application. *Journal of Marketing Research, 20*, 118–134.
Ramaswamy, V., Chatterjee, R., & Cohen, S. H. (1996). Joint segmentation on distinct interdependent bases with categorical data. *Journal of Marketing Research, 33*, 337–350.
Skinne, H. A. (1978). Differentiating the contribution of elevation, scatter and shape in profile similarity. *Educational and Psychological Measurement, 38*, 297–308.
Smith, W. R. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing, 21*, 3–8.
Spath, H. (1979). Clusterwise linear regression. *Computing, 22*, 367–373.
Vellido, A., Lisboa, P. J. G., & Meehan, K. (1999). Segmentation of the on-line shopping market using neural networks. *Expert Systems with Applications, 17*, 303–314.
Vriens, M., Wedel, M., & Wilms, T. (1996). Metric conjoint segmentation methods: A Monte Carlo comparison. *Journal of Marketing Research, 33*, 73–85.
Wedel, M., & Desarbo, W. S. (1995). A mixture likelihood approach for generalized linear models. *Journal of Classification, 12*, 1–35.
Wedel, M., & Desarbo, W. S. (1996). An exponential-family multidimensional scaling mixture methodology. *Journal of Business and Economic Statistics, 14*, 447–459.
Wedel, M., & Desarbo, W. S. (2002). Market segment derivation and profiling via a finite mixture model framework. *Marketing Letters, 13*, 17–25.
Wedel, M., & Kamakura, W. (2000). *Market segmentation: Conceptual and methodological foundations* (2nd ed.). Norwell, Massachusetts: Kluwer Academic Publishers.
Wedel, M., & Kistemaker, C. (1989). Consumer benefit segmentation using clusterwise linear regression. *International Journal of Research in Marketing, 6*, 241–258.
Wind, Y. (1978). Issues and advances in segmentation research. *Journal of Marketing Research, 15*, 317–337.
Yun, Y. B., Nakayama, H., Tanino, T., & Arakawa, M. (2001). Generation of efficient frontiers in multi-objective optimization problems by generalized data envelopment analysis. *European Journal of Operational Research, 129*, 586–595.
Zhong, S., & Ghosh, J. (2003). A unified framework for model-based clustering. *Journal of Machine Learning Research, 4*, 1001–1037.