Engineering Optimization, 2014
Vol. 46, No. 12, 1729–1747, http://dx.doi.org/10.1080/0305215X.2013.861458

Creating targeted initial populations for genetic product searches in heterogeneous markets

Garrett Foster[a], Callaway Turner[a], Scott Ferguson[a*] and Joseph Donndelinger[b]

[a]*Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, North Carolina, USA;*
[b]*General Motors Research and Development Center, Warren, Michigan, USA*

(*Received 1 February 2013; accepted 21 October 2013*)

Genetic searches often use randomly generated initial populations to maximize diversity and enable a thorough sampling of the design space. While many of these initial configurations perform poorly, the trade-off between population diversity and solution quality is typically acceptable for small-scale problems. Navigating complex design spaces, however, often requires computationally intelligent approaches that improve solution quality. This article draws on research advances in market-based product design and heuristic optimization to strategically construct 'targeted' initial populations. Targeted initial designs are created using respondent-level part-worths estimated from discrete choice models. These designs are then integrated into a traditional genetic search. Two case study problems of differing complexity are presented to illustrate the benefits of this approach. In both problems, targeted populations lead to computational savings and product configurations with improved market share of preferences. Future research efforts to tailor this approach and extend it towards multiple objectives are also discussed.

Keywords: genetic search; market-based product design; product line optimization; mixed-integer optimization

1. Introduction

Product designers face the challenge of creating products in markets where customers have highly heterogeneous preferences. Improving performance in market-related objectives, such as market share of preference, requires a product line, *i.e.* a set of related products that are offered by a single company. This is different from a product family where commonality of features, components and subsystems is often explicitly enforced (Simpson, Maier, and Mistree 2001).

Customers also desire products that maximize their value for money (Prahalad and Mashelkar 2010). This leads to complex design problems that require: (1) advanced techniques to capture and model customer preferences for product features; and (2) optimization techniques capable of searching the expansive mixed-integer design space associated with the resulting combinatorial problem. The nature of many feature-packaging problems supports the use of heuristic optimization approaches. Genetic techniques (Holland 1975), in particular, have been readily demonstrated in recent publications to be effective for a variety of problem types (Jaberipour and Khorram 2011; Rafique *et al.* 2011; Srivastava and Deb 2012; Tang, Chen, and Wei 2012; Villeneuve and Mavris 2012). However, a non-tailored genetic search can have difficulty identifying the most optimal regions of performance or require exorbitant computational costs. Improving the performance of a

*Corresponding author. Email: scott_ferguson@ncsu.edu

genetic search often begins by tailoring aspects of the algorithm, such as changing the population size (Kouchakpour, Zaknich, and Braunl 2007) and trying different approaches for selection (Kwak and Lee 2011), crossover (Lee *et al.* 2010) and mutation.

Research efforts exploring the design of the initial population often champion the use of random draws or points that are equally distributed throughout the design space (Karci 2004; Maaranen, Miettinen, and Makela 2004; Maaranen, Miettinen, and Penttinen 2007; Garcia-Arnau *et al.* 2007; Jaberipour and Khorram 2011; Poles *et al.* 2009; Rafique *et al.* 2011). While many of these initial product configurations may perform poorly when evaluated, the unknown topography of the solution contours requires genetic diversity to enable a proper search. From a design perspective, initial populations created from random draws can lead to product configurations ranging from acceptable to undesirable to absurd. Yet, when heuristic rules (Bruha 2005; Chong, Omar, and Bakar 2008; Kuczapski *et al.* 2010) or knowledge of problem aspects (Smith and Smith 2003; Nakamura *et al.* 2005; Togan and Daloglu 2008; Lin 2012; Sadrzadeh 2012) are used, tailored initial populations can lead to improvements in solution quality and computational cost.

This article explores how estimates of customer preference from statistical market models can be used to create candidates for an initial population. Specifically, initial product lines are created by combining candidate designs that maximize the product attribute part-worths estimated for a choice-based conjoint study. The hypothesis driving this work is that these candidate designs, which perform well for a subset of the market, can be combined to create product lines with better market performance than a randomly generated product line. By offering an improved starting point, it is then expected that the remainder of the genetic search will give improved results, yield computational efficiencies, or both.

2. Background

Research in market-based product design has been aided by the increased availability and capability of computational resources. This section provides a brief background to the history of product line design and how advances in market analysis and optimization have enabled this work.

2.1. *Market-based product design*

Product design decisions driven by utility and preferences have been an active area of research over the past 30 years. Work in the area of decision-based design, for example, initially focused on designer utility for engineering options (Thurston 1991). This work quickly evolved to include consumer utilities (Hazelrigg 1998) and maximizing the net present value of profit (Marston and Mistree 1998). Engineering design methods that leveraged market-demand models began to mature with the use of conjoint analysis (Green and Rao 1971; Green, Wind, and Jain 1972) and the S-Model (Cook and DeVor 1991). Li and Azarm (2000) further built upon Hazzelrigg's framework by using conjoint analysis to fit utility functions at the respondent level and considered the net present values of both market share and profit.

Choice-based conjoint studies and discrete choice analysis (Ben-Akiva and Lerman 1985; Louviere, Hensher, and Swait 2001; Train 2003) provided additional realism by considering product selection from a set of alternatives. These works saw the first application of the logit model (Wassenaar and Chen 2003; Wassenaar *et al.* 2004), experimental methods for profiling the market and mapping to the technical space (Hoyle *et al.* 2008; Kumar *et al.* 2009), exploration of model assumptions and their implication on results (Shiau *et al.* 2007; Donndelinger, Robinson, and Wissmann 2008), integration with existing design-decision tools (Michalek, Feinberg, and Papalambros 2005; Olewnik and Hariharan 2010), exploration of market heterogeneity (Shiau and Michalek 2009), application toward mass customization (Ferguson, Cormier, and Olewnik

2011) and the effect of retail channels (Williams, Kannan, and Azarm 2011). However, these works focused on the design of a single product.

Product line optimization can be traced back to Green and Krieger (1985). This led to the use of conjoint analysis (Moore, Louviere, and Verma 1999; Li and Azarm 2002) and nested-logit models (Kumar, Chen, and Simpson 2009). As research advanced to discrete choice analysis, demand-model formulation progressed from a multinomial logit (Michalek *et al.* 2006) to more advanced forms capable of representing customer heterogeneity. These works have primarily explored the latent-class multinomial logit and hierarchical Bayes mixed logit formulations (Kumar *et al.* 2009; Ferguson, Cormier, and Olewnik 2011; Michalek *et al.* 2011; Sullivan, Ferguson, and Donndelinger 2011; Turner, Ferguson, and Donndelinger 2011).

Research in market-based product design has highlighted the rich heterogeneous preference information captured when product attribute part-worths are estimated using hierarchical Bayes mixed logit models. While many choice model forms can be used in engineering design, hierarchical Bayes has clear advantages (Huber, Arora, and Johnson 1998; Michalek *et al.* 2011; Sullivan, Ferguson, and Donndelinger 2011). However, this richness comes at a price, as the resulting optimization problem can require tens or hundreds of design variables to simultaneously optimize a product line.

2.2. *Algorithms for product line optimization*

Early efforts in product line optimization used greedy heuristics (Green and Krieger 1985) and other heuristic rule-based approaches (Kohli and Krishnamurti 1987; Sudharshan, May, and Shocker 1987; Dobson and Kalish 1993; Nair, Thakur, and Wen 1995; Thakur *et al.* 2000). Analytical methods such as linear programming and analytical target cascading (McBride and Zufryden 1988; Hanson and Martin 1996; Michalek *et al.* 2006) have also successfully been used with multinomial logit models. However, such approaches are commonly not suitable for problems with discrete attribute levels.

To accommodate discrete and mixed-integer problem formulations, research has turned to heuristic optimization techniques such as branch-and-bound and genetic algorithms. While heuristic techniques were initially used to solve single product problems (Balakrishnan, Gupta, and Jacob 1996; Camm *et al.* 2006), they have been extended to product line optimization (Li and Azarm 2002; Steiner and Hruschka 2003; Belloni *et al.* 2008; Wang, Camm, and Curry 2009; Chapman and Alford 2011; Tsafarakis, Marinakis, and Matsatsinis 2011) and are more effective than greedy algorithms and analytical approaches (Belloni *et al.* 2008). Heuristics have also been applied to create intelligent starting points, though in very limited application. Balakrishnan, Gupta, and Jacob (2006), for example, used a dynamic programming heuristic to seed their initial population that improved solution quality but discarded good solutions early in the process.

Overall, there has been limited effort to improve the performance of optimization algorithms for market-based product design problems. Because of this, the information obtained from the demand model remains largely untapped. The next section of this article presents an approach that leverages this information to gain insight into the product configurations preferred by survey respondents. These designs are then used to create a genetic algorithm's initial population with the desired effects of accelerating convergence and increasing solution quality.

3. Technical approach

The approach presented in this article is intended for problems where the market is heterogeneous and part-worths have been estimated at the respondent level. Subsection 3.1 discusses the estimation and use of these part-worths in product line design problems, while Subsection 3.2 describes the targeted initialization approach.

3.1. *Estimation and use of heterogeneous customer preferences*

When the potential customer base for a product is large, demand models can be estimated using statistical approaches to gauge market response (Ben-Akiva and Lerman 1985; Train 2003). Heterogeneous market models capture the variation of preference within a market, providing the rich information needed to guide product line design decisions. In this work, preference information is estimated from the results of a choice-based conjoint survey. Here, respondents answer a battery of questions (choice tasks) where they select the product they are most likely to purchase, if any.

These data are then used to estimate model parameters capable of representing heterogeneous customer preferences. While many different model forms exist, a hierarchical Bayes mixed logit model (Sawtooth Software 2009; Train 2003) was chosen. The output from these models is a set of part-worth mean and standard deviation estimates. For the purposes of this work, only the estimated part-worth means are reported and used in simulation. Future work should study the validity of the posterior distribution (Gilbride and Lenk 2010) and explore the incorporation of this information in market simulations. These mean part-worth estimates are used to represent the change in product value when product composition decisions are made.

For a proposed product configuration, the estimated part-worths are used to calculate a respondent's observed utility. In this work the observed utility is calculated in two parts using a compensatory model where the attribute levels represent discrete product features. The implication of this formulation is that for the independent variables only one attribute level can be active for a given product. As shown in Figure 1, multiplying the matrix of estimated part-worths and the binary representation of the independent attributes (product configuration matrix) yields the first component of the observed utility. A similar mathematical procedure can be used when an integer encoding scheme is chosen.

To find the second part of the observed utility, relationships between the independent and dependent attributes must be defined. For the problems presented in this article, the only dependent attribute considered is product price. Here, a price model is used that multiplies an attribute cost vector, the product configuration matrix and a predefined mark-up variable to determine the price of a product. An example of this calculation is shown in Figure 2. It should also be noted that this linear additive pricing model is a simplifying assumption used in this article. More advanced techniques for determining product price exist, and their application is left for future work.

As product price can be represented on a continuous spectrum, its part-worth value can be calculated using interpolation. These part-worth values are added to the observed utility portion calculated in Figure 1. This yields a respondent's observed utility for each product in the line, which is often the primary component for many of the market-level objectives considered in product line design problems.

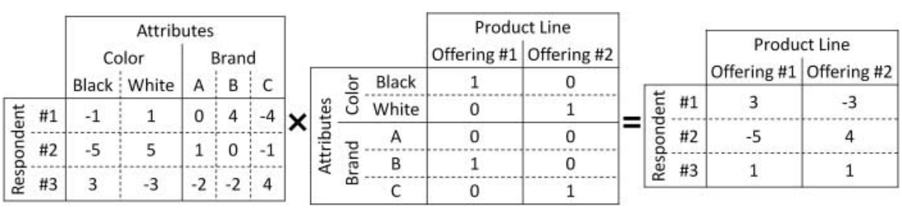

Figure 1. Example calculation of the first component of observed utility.

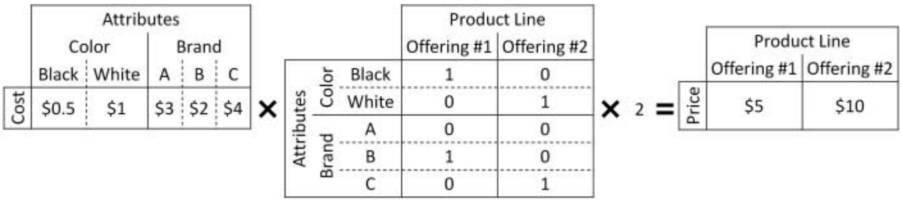

Figure 2. Example calculation of product price.

3.2. *Targeted initialization approach for product lines*

The observed utility value allows products that are most preferred by respondents to be identified and aggregated into initial product lines. It is not guaranteed that an individual's most preferred product will perform well at the market level or capture new pockets of market share in a product line scenario. However, it is reasonable to expect that if at least one respondent finds a product highly desirable, it could be an effective starting point for the genetic search. A flowchart of the general approach is shown in Figure 3, where the steps common to both targeted and random initialization approaches are shown with solid lines. Any steps added or modified by the targeted initialization approach are shown with dashed lines.

3.2.1. *Specify design string format*

The goal of this step is to specify design variable representation and define the size of the product line. Interactions or constraints within the design string are also defined at this stage. An integer encoding scheme was chosen, and a 100% mark-up (*i.e.* double cost) is assumed for each included level.

3.2.2. *Define initial population size*

A challenge of defining population size is balancing computational cost with diversity in the design space. Too few initial designs will cause the search to prematurely converge. Too many initial designs can cause the search to take an excessive amount of time to converge, which is often directly proportional to the computational cost. In this article, five population sizes are investigated. These population sizes are 10, 8, 6, 4 and 2 times the number of design variables

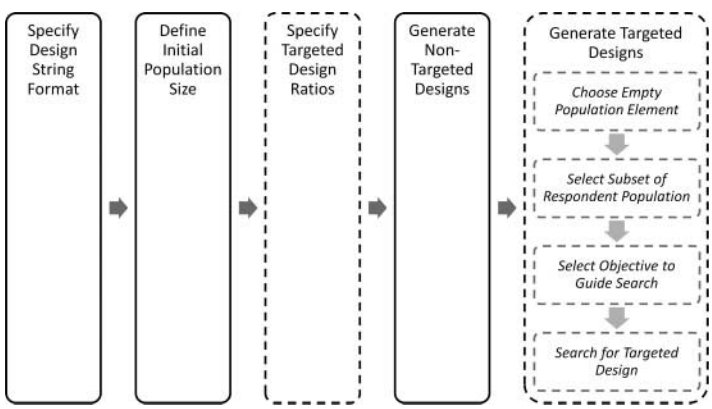

Figure 3. Generalized targeted initialization approach.

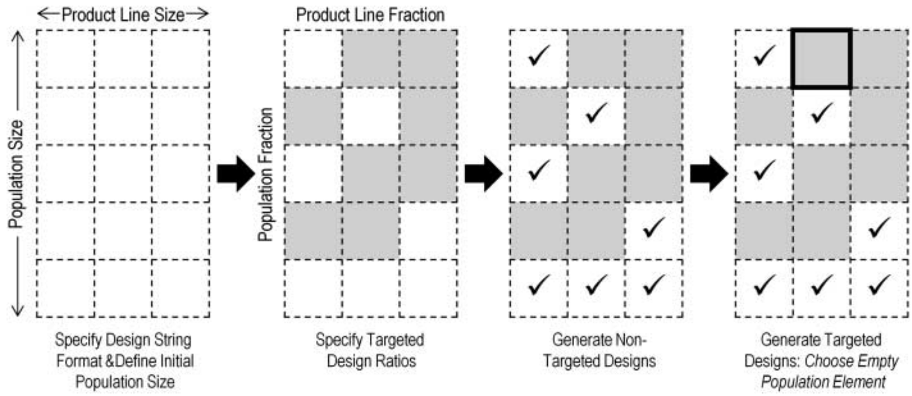

Figure 4. Visualization of generalized targeted initialization approach.

(denoted by 10x, 8x, 6x, 4x and 2x, respectively). The first element of Figure 4 provides a visual representation of the population structure: columns represent a product in the product line, and the number of rows equals the population size.

3.2.3. *Specify targeted design ratios*

A targeted initial population can be constructed using both targeted and non-targeted designs. Two settings are used to specify the ratio between them. The *population fraction* (n_{pf}) defines how many of the product lines in the population contain targeted designs, while the *product line fraction* (n_{lf}) defines how many products within each of these targeted product lines are targeted. Note that setting either fraction to 0 will yield a non-targeted (*i.e.* random) initial population.

The second element of Figure 4 shows a hypothetical population with a 0.8 population fraction and a 0.67 product line fraction. A grey box indicates an element classified for a targeted design, while a white box indicates an element classified for a non-targeted design.

3.2.4. *Generate non-targeted designs*

Non-targeted designs can be generated using an existing initialization strategy such as random variable selection. In a traditional genetic search, this would be the end of the initialization step. The third element of Figure 4 shows each of the non-targeted elements being filled.

3.2.5. *Generate targeted designs*

If targeted designs are desired, the initial population will not yet be fully defined. The four substeps below are used to identify a single product. To complete the definition of the entire population, this process is repeated.

3.2.6. *Choose empty population element*

The first substep involves selecting an empty population element that has been classified for a targeted design. This step is represented by the grey box with a black outline in the fourth element of Figure 4.

3.2.7. *Select focused subset of respondent population*

A targeted design is found by optimizing its performance with respect to a focused subset of the respondent population. This subset can be as small as a single respondent, and is constrained to be less than the entire respondent pool. In this research, each targeted design was found by targeting a randomly selected respondent. To decrease the likelihood of duplicate products entering a design string, a constraint was created that prevented a respondent from being used more than once in the same product line. To prevent unnecessary computational effort, each respondent's targeted design was found once and reused as required.

3.2.8. *Select objective to guide search*

An objective should be defined that will guide the search for the targeted product. This objective should be related to the overlying objective(s) of the problem. The formulation of the problem objective used in this article is discussed in Section 4.

3.2.9. *Search for targeted design*

Prior to searching, problem constraints should be defined. Next, the designer must select a search algorithm to solve for the configuration of the targeted design. A trade-off to be considered is the optimality of the suboptimization versus the computational expense of conducting the search. Understanding designer preferences in the context of this trade-off can help to guide the selection of a suitable algorithm. After conducting the search, the population element is filled. Once all the elements in the initial population have been filled the initialization step is complete.

In this article, two search algorithms are investigated: branch-and-bound and genetic search. The impact of the optimization scheme used to generate the targeted designs is further discussed in Subsection 5.1. Having described the approach for generating targeted initial populations, the next section introduces two case study problems and the testing procedure used to test this approach.

4. Experimental set-up

To test this approach, Sections 4.1 and 4.2 introduce two case study problems that differ greatly in size. Section 4.3 describes the testing procedure used. In this article, it is assumed that the goal of the product line is to maximize market share of preference. The formulation of the market share of preference (S) objective for an entire product line is shown in Equation (1):

$$S = \sum_{i=1}^{n_p} \frac{\sum_{j=1}^{n_r} \frac{e^{V_{ji}}}{\sum_{k=1}^{n_p} e^{V_{jk}} + \sum_{l=1}^{n_c} e^{V_{jl}} + e^{V_j^{none}}}}{n_r} \tag{1}$$

The first summation in Equation (1) is used to combine the market share of preference for all of the products in the product line (n_p). Next, the outermost summation in the numerator combines the probability of purchase for all of the respondents (n_r) for the current product (i) before dividing by the number of respondents to obtain an average market share of preference for the current product. The probability of purchase for the current respondent (j) is calculated by dividing the exponential of the respondent's observed utility (V_{ji}) of the current product by the sum of the exponentials of their utility of all the represented options. These represented options include the observed utility (V_{jk}) of all the products in the product line, the observed utility (V_{jl}) of all the competing products (n_c), and the observed utility of not selecting any of the products (V_j^{none}).

As described in Section 3.2, one of the tasks involves selecting an objective to guide the search. Given the form of Equation (1), maximizing the observable utility of a product (V_{ji}) correlates with the goal of maximizing the product line's market share of preference. In addition, this calculation can be done with reduced computational effort and is used to guide all targeted product searches in this article.

4.1. *MP3 product line design problem*

The first case study problem is based on a choice-based conjoint fielded by researchers at NC State and the University at Buffalo. The choice-based conjoint study for this problem was created using Sawtooth Software's SSI Web (Sawtooth Software 2008). One-hundred and forty students answered 10 choice task questions each. As shown in Table 1, nine product attributes were studied, yielding a total of 23,040 possible build combinations.

4.2. *Vehicle feature packaging problem*

The second case study problem originally motivated this research because of the complexity associated with the product line optimization problem. Results for this case study are based on market data obtained by General Motors, where 2275 respondents each answered 19 choice task questions. Table 2 shows the number of levels associated with each attribute. An enumeration of all possible feature combinations yields over 1,074,954,240 vehicle configurations that must be considered for a single vehicle.

4.3. *Experimental set-up*

The product lines in both case study problems contain five products, as previous work with these problems has suggested that this is the optimal product line size (Turner, Ferguson, and Donndelinger 2011). This means that the design string for the MP3 player case study contained 45 integer variables (nine for each product), while the vehicle case study had a design string with 95 integer variables (19 for each product).

To conduct the optimization, a genetic algorithm was written in Matlab (MathWorks 2011). The size of the population at each generation was constant based on the number of initial design variables. Recall that five population sizes will be investigated ($10x$, $8x$, $6x$, $4x$ and $2x$). The selection operator was random, and the crossover operator was uniform with a mixing ratio of 0.5. The mutation operator was Gaussian with a mutation rate of 5%. Finally, the convergence criterion was met when 20 generations were performed with no improvement in the objective value.

Table 1. Levels per attribute for the MP3 player feature packaging problem.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
2	2	2	2	2	3	5	6	8

Table 2. Levels per attribute for the vehicle feature packaging problem.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}
3	2	5	6	2	3	3	2	4	2	3	2	4	3	3	4	4	3	2

To assess the effectiveness of this approach all operators other than initialization were held constant, meaning that the remainder of the genetic search is executed in a manner identical to that of one initialized with a random population. In addition, multiple runs were performed for both case study problems to account for the stochastic nature of the genetic algorithm. These results were compared with runs initialized with populations full of randomly generated designs.

5. Results

In this section two metrics are used to investigate the results. The first metric is solution quality, which is measured using the value of the objective. The second metric is the number of objective function evaluations. This serves as an unbiased time comparison as it is independent of computer-specific attributes and eliminates computer variability from run to run. However, for this metric, the cost of completing the suboptimizations cannot be ignored. To create a fair comparison, a conservative estimate of initialization cost using targeted designs (F_0) is found by multiplying three main quantities. Those values, shown in Equation (2), are the average cost to find a respondent's optimal product ($\bar{f}$), the number of targeted products needed to seed the entire initial population (n_{tar}), and a conversion factor to convert between the computational cost of evaluating an individual product for a single respondent and the cost of evaluating an entire product line for the whole respondent base (C).

$$F_0 = \bar{f} * n_{\text{tar}} * C \tag{2}$$

The average cost to find a respondent's optimal product ($\bar{f}$) is simply the average of the number of functional evaluations needed to find each respondent's optimal product. The number of targeted products needed to seed the entire initial population (n_{tar}) is calculated under the assumption that no products are repeated in the targeted population. As seen in Equation (3), this quantity has two possible values: (1) the minimum number of unique ideal products needed; or (2) the maximum number of unique ideal products that can possibly be found from the number of respondents.

$$n_{\text{tar}} = \min(n_p * n_{pf} * n_l * n_{lf}, n_r) \tag{3}$$

The first expression in Equation (3) is found by multiplying the number of products in the product line (n_p), the population fraction (n_{pf}), the total number of product lines (n_l) and the product line fraction (n_{lf}). The second possible expression is simply the number of modelled respondents (n_r), since each respondent should have a single optimal product.

To compute the conversion factor (C) between the cost of evaluating an individual product for a single respondent and the cost of evaluating an entire product line for the whole respondent base, consider the number of different calculations carried out to evaluate market share of preference during three scenarios: (1) one product for one respondent; (2) one product for 1000 respondents; and (3) five products for 1000 respondents. Table 3 shows how the number of calculations increase as more products (n_p) and more respondents (n_r) are added.

From these calculations, major increases are seen in the quantity of observed utility calculations and the quantity of individual share calculations. The value of each is equal to the number of

Table 3. Quantity of calculations used to calculate market share of preference.

n_p	n_r	Utility	Share	Market avg.	Line avg.
1	*1*	1	1	–	–
1	*1000*	1000	1000	1	–
5	*1000*	5000	5000	5	1

products in the product line (n_p) times the number of respondents modelled (n_r). Meanwhile, the quantity of average calculations is fairly low throughout. Assuming that their computational expense is relatively low, the average calculations can be neglected, leaving only the quantity of observed utility calculations and the quantity of individual share calculations to consider.

From these quantities, either: (1) one of the calculation types is more expensive and thus dominates; or (2) both calculation types are equivalent. In the first scenario, only the quantity of the more expensive calculation is considered—either utility or share—while the other is neglected. Neglecting the less expensive calculation creates a conservative estimate with a ratio of $1/5000$. The second scenario considers both sets of calculations and results in a ratio of $2/10{,}000$. As seen, both ratios simplify to the same value, and therefore the conversion factor (C) can be generalized as seen in Equation (4):

$$C = \frac{1}{n_p * n_r} \tag{4}$$

5.1. *Computational cost to identify candidate designs*

Many different optimization techniques can be used to solve the candidate designs of the targeted population. This article discusses two different optimization algorithms well suited to discrete optimization: branch-and-bound and genetic search.

The branch-and-bound algorithm was able to find the optimal product for all 140 respondents in the MP3 problem using only 29,425 evaluations (f), *i.e.* an average of 210 evaluations per respondent ($\bar{f}$). The genetic search, meanwhile, required 213,210 evaluations (f) to converge for all 140 respondents, *i.e.* an average of 1522 evaluations per respondent ($\bar{f}$). A comparison of the final designs showed that the genetic search was able identify 139 of the 140 optimal designs found by the branch-and-bound algorithm.

A more significant difference between the suboptimization approaches was found in the vehicle case study. The branch-and-bound algorithm required 157,783,100 evaluations (f), with an average of 69,355 evaluations per respondent ($\bar{f}$) and a standard deviation of 181,145 evaluations. The genetic search required only 11,507,160 evaluations, with an average of 5058 evaluations per respondent ($\bar{f}$) and a standard deviation of 952 evaluations. One explanation for the high computational cost of the branch-and-bound algorithm was that 27 of the 2275 respondents required over 1 million function evaluations to identify their most preferred design. Furthermore, solutions from the two approaches were noticeably different. Of the 2275 solutions the optimization algorithms found the same design only 1570 times, while 250 of the unique designs differed on at least four attributes.

The results from this exploration indicate that for problems with smaller design spaces the branch-and-bound approach may be a more computationally effective search strategy. However, the computational cost of the genetic search scales better as the design space increases. For the remainder of this article all candidate designs for the targeted population are identified using a genetic search. The exploration and tailoring of suboptimization algorithms are a source of future work and will be discussed in Section 6.

5.2. *Objective function solution quality of initial starting points*

The expectation of including targeted designs in the initial population is that initial solution quality is improved. Figure 5 shows this improvement (in terms of market share of preference) for both case study problems. The two box plots in each figure are for a fully random population and a fully targeted population. For both problems, a $10x$ initial population size was used. In the MP3 problem, 15 random and 25 targeted populations were generated. From these populations, over

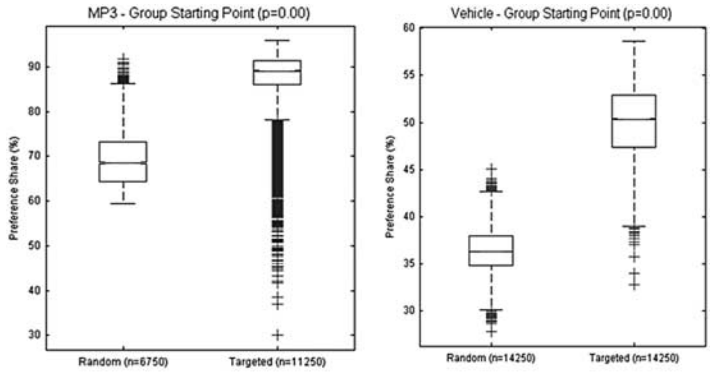

Figure 5. Initial solution quality of all targeted and random designs in the MP3 problem (left) and the vehicle problem (right).

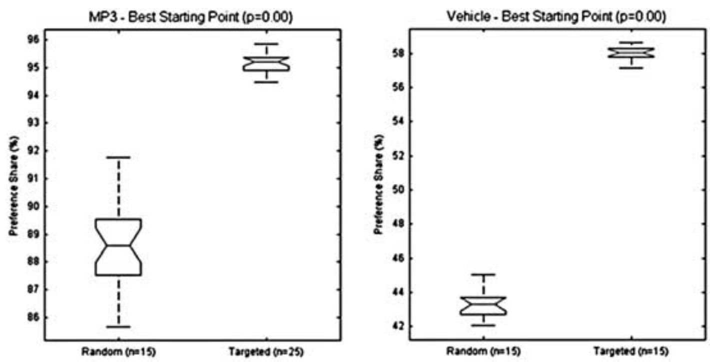

Figure 6. Initial solution quality of best targeted and random designs in the MP3 problem (left) and the vehicle problem (right).

75% of the targeted designs performed better than 75% of the random designs. A similar outcome was seen for the vehicle problem, where 15 populations of each type were used. Both results were statistically significant with p-values $\ll 0.01$.

Figure 6 shows the difference in solution quality when only the best initial scores from each run are considered. The targeted initial populations start with higher median share than the random initial populations. These results were also statistically significant with p-values $\ll 0.01$. Respondents in the MP3 problem were much more likely to purchase than respondents in the vehicle problem, as the best starting points capture over 85% share. Moreover, the best starting points of the targeted initial population are very near the best final solution of 96.37%.

5.3. *Solution quality of final objective function values*

By improving the solution quality of the initial population, the expectation is that the genetic search should converge more rapidly and find a better final solution given a fixed amount of computational effort than a randomly initialized genetic search. To investigate the potential improvement in solution, box plots were created that show the difference between the best final scores for each of the sample runs. Recall that each search was stopped after 20 generations with no improvement in the objective value.

The results in Figure 7 show that for the MP3 problem, the targeted data had less overall spread. This is supported by a p-value of 0.18, which suggests that the two populations have some

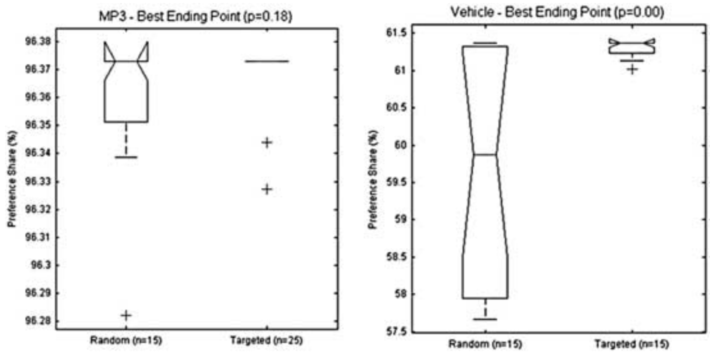

Figure 7. Targeted vs random best ending points in the MP3 problem (left) and the vehicle problem (right).

similarities. A possible explanation for this finding is that the global optimum lies at, or near, 96.37%. This would imply that both approaches were able to converge to the global optimum, with the targeted population converging there more often (22 out of 25 versus 11 out of 15).

The results from the vehicle problem demonstrate that the final solution using a fully targeted population had higher median share than a final solution using the random population. This result was statistically significant with a p-value near 0. Furthermore, the spread of the final solutions when using a targeted initial population was much tighter than the spread of the final solutions when using a randomly generated initial population.

5.4. *Convergence rates with different targeted population settings*

Having demonstrated that a fully targeted initial population can improve final solution quality over a fully random initial population, a study was conducted to explore how the three parameters of a targeted initial population (population size, population fraction and product line fraction) influenced algorithm convergence. The differences in convergence are visualized using several x–y plots. Each plot shows, for all share values, the maximum, median and minimum market share of preference in relation to the number of functional evaluations. The y-axis shows a 25% market share of preference range, while the x-axis shows the number of function calls needed to evaluate 100 generations initialized with the fully random population. These plots have been shifted to include the additional cost required to find the targeted designs using Equation (2). For brevity, only the results from the vehicle problem are presented.

5.4.1. *Effect of population size on convergence rate*

Figure 8 shows the effect of population size on convergence. A 10x population size is used for the randomly generated population. The smaller targeted populations converge to a quality solution in fewer functional evaluations. Moreover, the worst performing targeted population (10x) converges with less computational cost than the randomly generated initial population. This supports the claim of a targeted initial population leading to reduced computational cost.

5.4.2. *Effect of population and product line fractions on convergence rate*

To this point, only fully targeted populations have been compared with fully random populations. However, increasing the diversity within a targeted population may be beneficial. A computationally inexpensive and straightforward way to achieve this diversity is to include random designs

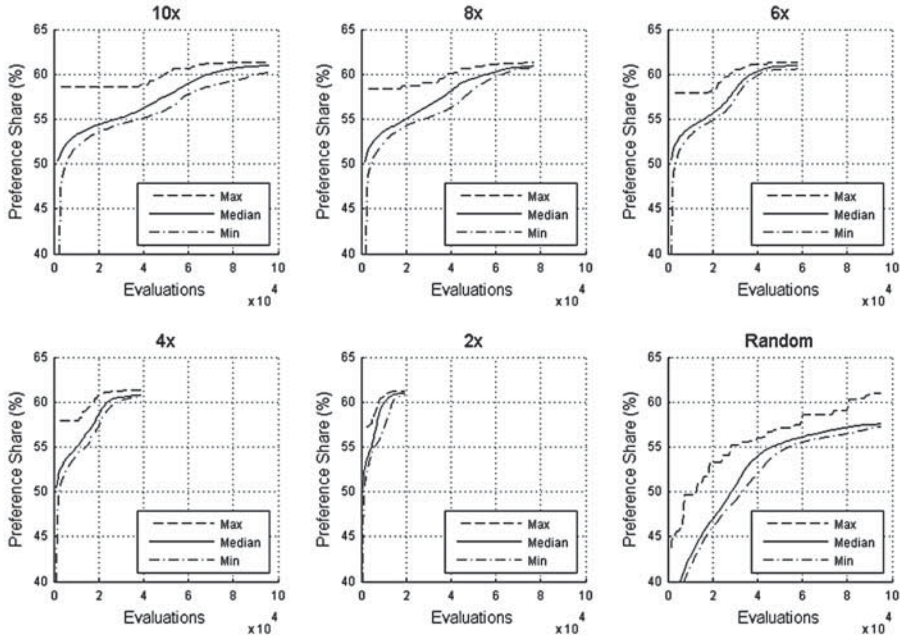

Figure 8. Convergence rate of different population sizes in the vehicle problem.

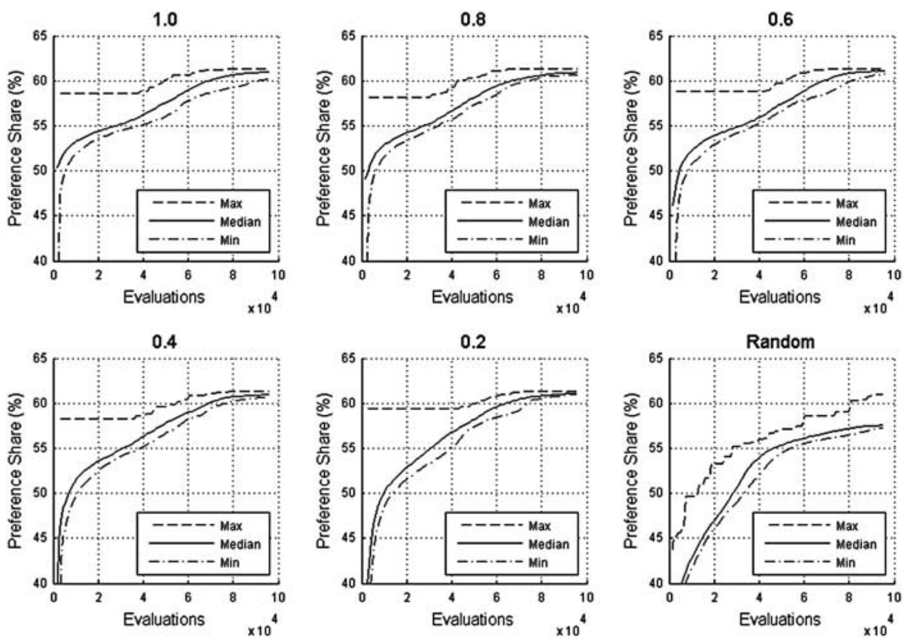

Figure 9. Convergence rate of different population fractions in the vehicle problem.

within the targeted population by using the population fraction and product line fraction settings defined in Subsection 3.2.

Results exploring the impact of the population fraction are shown in Figure 9. Results exploring the impact of the product line fraction are shown in Figure 10. Both figures show that the inclusion

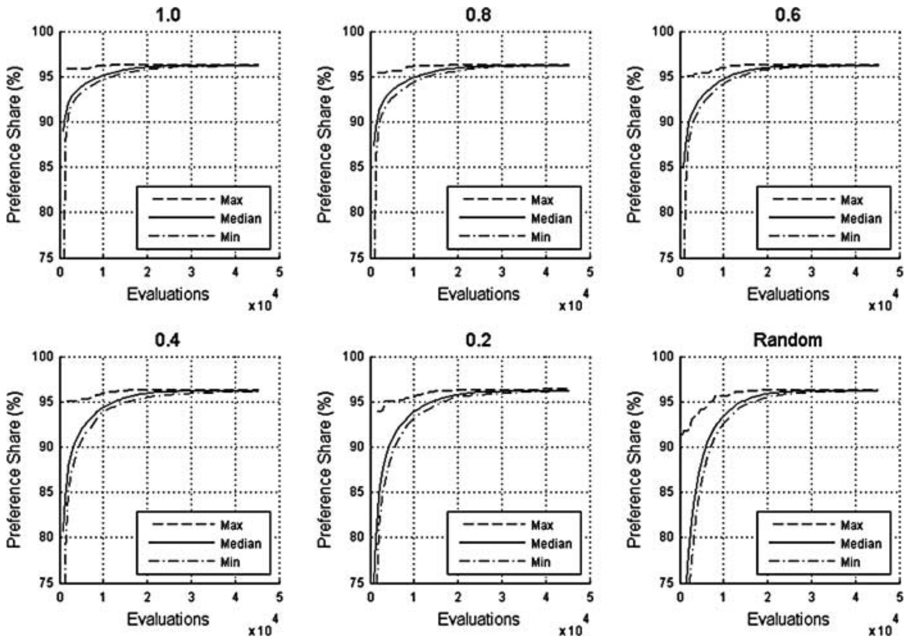

Figure 10. Convergence rate of different product line fractions in the vehicle problem.

of random designs has a relatively small effect on the convergence of a targeted population. Of the two fractions, product line fraction has a slightly larger effect than population fraction. A likely explanation for this is that a random product within a targeted product line lives long enough to impact the solution, while a random product line within the initial population dies off quickly owing to its vastly inferior performance (as shown in Subsection 5.2). An interesting result from these figures is that including a small number of targeted designs in an initial population has a fairly dramatic impact on performance. However, a close look at Figure 10 shows that a larger product line fraction will yield slightly better performance.

Table 4. Summary of MP3 problem results.

Population size	Population fraction	Product line fraction	No. of runs	95% CI on best solution	Average evaluations to market share of preference of 94%	95%	96%
10x	0	0	15	96.36 ± 0.01	6.4e3	8.8e3	15.1e3
10x	1	1	25	96.37 ± 0.00	0.5e3	1.5e3	10.0e3
10x	1	0.8	10	96.35 ± 0.03	0.5e3	2.6e3	11.8e3
10x	1	0.6	10	96.37 ± 0.01	0.7e3	4.1e3	12.4e3
10x	1	0.4	10	96.37 ± 0.01	1.6e3	5.2e3	14.0e3
10x	1	0.2	10	96.36 ± 0.03	4.0e3	7.7e3	15.0e3
10x	0.8	1	5	96.36 ± 0.02	0.5e3	1.5e3	11.2e3
10x	0.6	1	5	96.37 ± 0.02	0.6e3	1.7e3	11.1e3
10x	0.4	1	5	96.36 ± 0.02	0.5e3	3.7e3	11.6e3
10x	0.2	1	5	96.37 ± 0.00	0.7e3	3.2e3	12.2e3
8x	1	1	5	96.37 ± 0.02	0.5e3	0.5e3	8.9e3
6x	1	1	5	96.37 ± 0.00	0.4e3	0.7e3	6.3e3
4x	1	1	5	96.35 ± 0.05	0.4e3	0.9e3	4.8e3
2x	1	1	5	96.31 ± 0.06	0.6e3	1.4e3	2.7e3

Table 5. Summary of vehicle problem results.

Population size	Population fraction	Product line fraction	No. of runs	95% CI on best solution	Average evaluations to market share of preference of 57%	59%	61%
10x	0	0	15	59.64 ± 0.92	60.8e3	118.0e3	138.9e3
10x	1	1	15	61.30 ± 0.06	3.1e3	49.5e3	82.2e3
10x	1	0.8	5	61.37 ± 0.00	3.4e3	52.8e3	83.4e3
10x	1	0.6	5	61.36 ± 0.02	22.4e3	56.4e3	79.8e3
10x	1	0.4	5	61.25 ± 0.13	27.5e3	47.8e3	84.5e3
10x	1	0.2	5	60.97 ± 0.84	42.2e3	59.9e3	85.3e3
10x	0.8	1	5	61.35 ± 0.04	3.0e3	46.5e3	79.2e3
10x	0.6	1	5	61.35 ± 0.04	4.9e3	50.7e3	73.8e3
10x	0.4	1	5	61.31 ± 0.16	2.6e3	48.2e3	81.7e3
10x	0.2	1	5	61.27 ± 0.18	21.7e3	37.3e3	71.3e3
8x	1	1	5	61.33 ± 0.04	3.3e3	36.7e3	64.7e3
6x	1	1	5	61.22 ± 0.13	6.9e3	29.1e3	47.9e3
4x	1	1	5	61.25 ± 0.20	5.6e3	20.1e3	42.7e3
2x	1	1	5	61.22 ± 0.04	7.2e3	11.6e3	18.6e3

5.5. *Summary of results*

A summary of all results can be seen in Table 4 for the MP3 case study and Table 5 for the vehicle case study. In both tables, the first, row corresponds to a random initial population. The remaining rows represent populations with different proportions of targeted elements. Overall, these results demonstrate that the inclusion of targeted designs in the initial population can lead to searches that converge more quickly and find solutions that are as good as or better than a search initialized with a random initial population.

Having explored the effectiveness of using a targeted initial population and the major parameters associated with creating one, the next section of this article revisits the initial claims by analysing these results. It also describes pathways for future development of this approach.

6. Conclusions and future work

This article presented an approach for improving the starting population of a genetic search by strategically selecting products that are desirable for at least a subset of the market. The theory is that product configurations that perform well for a subset of the market have a better chance of performing well for the entire market than a randomly generated product line. Two case study problems of differing difficulty were used to demonstrate the benefits of this approach.

The first investigation in Subsection 5.1 explored the cost associated with finding targeted designs using different search algorithms. The results indicated that as the size of the design space increases, the cost of using branch-and-bound becomes computationally intractable and that a genetic algorithm scales more favourably. A future study will look at how the trade-off between computational cost and design optimality when generating targeted designs influences algorithm performance.

The next investigation in Subsection 5.2 demonstrated that targeted product lines perform better initially than random product lines, which supports the basic theory of this approach. Meanwhile, Subsection 5.3 explored the quality of the final solution. It was shown that for the smaller MP3 problem, both the fully targeted and fully random initial populations led to the discovery of the perceived optimum (96.37% share). However, the solution from the fully targeted initial population more consistently identified the perceived optimum. In the more complex automobile problem, using a fully targeted initial population led to higher median market share of preference values (61.37% versus 59.87%). This finding was statistically significant ($p \ll 0.01$).

The investigation in Subsection 5.4.1 demonstrated that using a fully targeted initial population led to a faster convergence than using a fully random initial population of the same size. Subsection 5.4.1 also demonstrated that a smaller fully targeted population converged to a quality solution in fewer functional evaluations than either a larger fully targeted population or a fully random population. This finding demonstrates that the use of targeted populations can reduce computational cost for genetic product searches in heterogeneous markets.

Initial investigations were also performed regarding the impact of diversity within the targeted product line and the computational cost of finding targeted products. Subsection 5.4.2 showed that adding diversity to the targeted initial population by changing the population fraction and product line fraction had little impact on convergence or final solution quality. This result, combined with the population size study in Subsection 5.4.1, indicates that the genetic search needs only a small percentage of targeted designs to improve algorithm performance. This provides opportunities for additional computational savings. Future work will explore approaches to improve diversity within a targeted population via different product selection strategies (*i.e.* not random). This is necessary as the results in Subsection 5.4.2 suggest that random product lines are eliminated too quickly to affect algorithm performance.

The summarized results shown in Subsection 5.5 indicate that targeting a subset of the respondent population when creating initial designs should ultimately lead to faster searches that find solutions that are as good as, or better than, a search initialized with a random initial population. Furthermore, the approach described in Subsection 3.2 allows targeted designs to be created for subsets of a respondent population or even different system uses such as various loading conditions and missions. The next area of future work involves extending this approach to further guide product line/architecture decisions. First, as multiobjective decision spaces have become more prevalent, this approach must be extended to handle multiple objectives. Next, as the decision space increases in the number of objectives (*i.e.* more than 2 or 3), better methods of visualizing the tradeoffs are needed. Finally, as larger product line design problems can have significant run-times, it would be beneficial to include a run-time analyser that estimates the potential improvements in objectives for additional periods of optimization run-time. This would enable the designer to make informed decisions about terminating the optimization.

Funding

The authors gratefully acknowledge support from the National Science Foundation [NSF grant no. CMMI-0969961] and from General Motors. Any opinions, findings and conclusions presented in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Balakrishnan, P. V., R. Gupta, and V. S. Jacob. 1996. "Genetic Algorithms for Product Design." *Management Science* 42 (8): 1105–1117.

Balakrishnan, P. V., R. Gupta, and V. S. Jacob. 2006. "An Investigation of Mating and Population Maintenance Strategies in Hybrid Genetic Heuristics for Product Line Designs." *Computer & Operations Research* 33 (3): 639–659.

Belloni, A., R. M. Freund, M. Selove, and D. Simester. 2008. "Optimal Product Line Design: Efficient Methods and Comparisons." *Marketing Science* 54 (9): 1544–1552.

Ben-Akiva, M., and S. R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Boston: MIT University Press.

Bruha, I. 2005. "Rule Representation and Initial Population in Genetic Learning." 3rd Industrial Simulation Conference, no. 76: 37–41.

Camm, J., D. Curry, J. Cochran, and S. Kannan. 2006. "Conjoint Optimization: An Exact Algorithm for the Share-of-Choice Problem." *Management Science* 52 (3): 435–447.

Chapman, C., and J. Alford. 2011. "Product Portfolio Evaluation Using Choice Modeling and Genetic Algorithms." Proceedings of the 2010 Sawtooth Software Conference, Newport Beach, CA.

Chong, K. E., M. K. Omar, and N. A. Bakar. 2008. "Solving Assembly Line Balancing Problems using Genetic Algorithm with Heuristics-Treated Initial Population." *World Congress on Engineering: Lecture Notes in Engineering and Computer Science* I–II: 1273–1277.

Cook, H. E., and R. E. DeVor. 1991. "On Competitive Manufacturing. Enterprises I: The S-Model and the Theory of Quality." *Manufacturing Review* 4 (2): 96–105.

Dobson, G., and S. Kalish. 1993. "Heuristics for Pricing and Positioning a Product-Line Using Conjoint and Cost Data." *Management Science* 39 (2): 160–175.

Donndelinger, J. A., J. A. Robinson, and L. A. Wissmann. 2008. "Choice Model Specification in Market-Based Engineering Design." Proceedings of the 2008 IDETC, Brooklyn, NY, DETC2008-50071.

Ferguson, S., P. Cormier, and A. Olewnik. 2011. "Exploring Marketing to Engineering Information Mapping in Mass Customization: A Presentation of Ideas, Challenges and Resulting Questions." Proceedings of the 2011 IDETC, Washington, DC, DETC2011-48742.

Garcia-Arnau, M., D. Manrique, J. Rios, and A. Rodriguez-Paton. 2007. "Initialization Method for Grammar-Guided Genetic Programming." *Knowledge-Based Systems* 20: 127–133.

Gilbride, T. J., and P. J. Lenk. 2010. "Posterior Predictive Model Checking: An Application to Multivariate Normal Heterogeneity." *Journal of Marketing Research* 47 (5): 869–909.

Green, P., and V. Rao. 1971. "Conjoint Measurement for Quantifying Judgmental Data." *Journal of Marketing Research* 8 (3): 355–363.

Green, P., Y. Wind, and A. Jain. 1972. "Consumer Menu Preference: An Application of Additive Conjoint Measurement." Proceedings of the Third Annual Conference of the Association for Consumer Research, pp. 304–315.

Green, P. E., and A. M. Krieger. 1985. "Models and Heuristics for Product Line Selection." *Marketing Science* 4 (1): 1–19.

Hanson, W., and K. Martin. 1996. "Optimizing Multinomial Logit Profit Functions." *Management Science* 42 (7): 992–1003.

Hazelrigg, G. A. 1998. "A Framework for Decision-Based Engineering Design." *Journal of Mechanical Design* 120 (4): 653–658.

Holland, J. 1975. "Adaptation in Natural and Artificial Systems." *SIAM Review* 18 (3): 287–299.

Hoyle, C., W. Chen, B. Ankenman, and N. Wang. 2008. "Optimal Experimental Design of Human Appraisals for Modeling Consumer Preferences in Engineering Design." Proceedings of the 2008 IDETC, Brooklyn, NY, DETC2008–49561.

Huber, J., N. Arora, and R. Johnson. 1998. "Capturing Heterogeneity in Consumer Choices." *ART Forum*, American Marketing Association.

Jaberipour, M., and E. Khorram. 2011. "A New Harmony Search Algorithm for Solving Mixed-Discrete Engineering Optimization Problems." *Engineering Optimization* 43 (5): 507–523.

Karci, A. 2004. "Novelty in the Generation of Initial Population for Genetic Algorithms." *Knowledge-Based Intelligent Information and Engineering Systems*: *Lecture Notes in Computer Science* 3214: 268–275.

Kohli, R., and R. Krishnamurti. 1987. "A Heuristic Approach to Product Design." *Management Science* 33 (12): 1523–1533.

Kouchakpour, P., A. Zaknich, and T. Braunl. 2007. "Population Variation in Genetic Programming." *Information Sciences* 177: 3438–3452.

Kuczapski, A. M., M. V. Micea, L. A. Maniu, and V. I. Cretu. 2010. "Efficient Generation of Near Optimal Initial Populations to Enhance Genetic Algorithms for Job-Shop Scheduling." *Information Technology and Control* 39 (1): 32–37.

Kumar, D., W. Chen, and T. Simpson. 2009. "A Market-Driven Approach to the Design of Platform-Based Product Families." *International Journal of Production Research* 47 (1): 71–104.

Kumar, D., C. Hoyle, W. Chen, N. Wang, G. Gomez-Levi, and F. Koppelman. 2009. "A Hierarchical Choice Modeling Approach for Incorporating Customer Preferences in Vehicle Package Design." *International Journal of Product Development* 8 (3): 228–251.

Kwak, N-S., and J. Lee. 2011. "An Implementation of New Selection Strategies in a Genetic Algorithm—Population Recombination and Elitist Refinement." *Engineering Optimization* 43 (12): 1367–1384.

Lee, K-M., J-T. Tsai, T-K. Liu, and J-H. Chou. 2010. "Improved Genetic Algorithm for Mixed-Discrete-Continuous Design Optimization Problems." *Engineering Optimization* 42 (10): 927–941.

Li, H., and S. Azarm. 2000. "Product Design Selection Under Uncertainty and With Competitive Advantage." *Journal of Mechanical Design* 122 (4): 411–418.

Li, H., and S. Azarm. 2002. "An Approach for Product Line Design Selection Under Uncertainty and Competition." *Journal of Mechanical Design* 124 (3): 385–392.

Lin, C. C. 2012. "Hierarchical Path Planning for Manipulators based on Genetic Algorithm with Non-random Initial Population." *International Journal of Innovative Computing Information and Control* 8 (3A): 1773–1786.

Louviere, J., D. Hensher, and J. Swait. 2001. *Stated Choice Methods: Analysis and Applications in Marketing, Transportation and Environmental Valuation*. Cambridge: Cambridge University Press.

Maaranen, H., K. Miettinen, and M. M. Makela. 2004. "Quasi-random Initial Populations for Genetic Algorithms." *Computers and Mathematics with Applications* 47: 1885–1895.

Maaranen, H., K. Miettinen, and A. Penttinen. 2007. "On Initial Populations of a Genetic Algorithm for Continuous Optimization Problems." *Journal of Global Optimization* 37: 405–436.
Marston, M., and F. Mistree. 1998. "An Implementation of Expected Utility Theory in Decision Based Design." Proceedings of the 1998 DETC, Atlanta, GA, DETC98/DTM-5670.
MathWorks. 2011. MATLAB version 7.12.0.635. March 18, 2011.
McBride, R. D., and F. S. Zufryden. 1988. "An Integer Programming Approach to the Optimal Product Line Selection Problem." *Marketing Science* 7 (2): 126–140.
Michalek, J., O. Ceryan, P. Papalambros, and Y. Koren. 2006. "Balancing Marketing and Manufacturing Objectives in Product Line Design." *Journal of Mechanical Design* 128 (6): 1196–1204.
Michalek, J., P. Evves, F. Adigüzel, F. Feinberg, and P. Papalambros. 2011. "Enhancing Marketing with Engineering: Optimal Product Line Design for Heterogeneous Markets." *International Journal of Research in Marketing* 28 (1): 1–12.
Michalek, J., F. Feinberg, and P. Papalambros. 2005. "Linking Marketing and Engineering Product Design Decisions via Analytical Target Cascading." *Journal of Product Innovation Management* 22 (1): 42–62.
Moore, W., J. Louviere, and R. Verma. 1999. "Using Conjoint Analysis to Help Design Product Platforms." *Journal of Product Innovation Management* 16 (1): 27–39.
Nair, S. K., L. S. Thakur, and K. W. Wen. 1995. "Near Optimal Solutions for Product Line Design and Selection: Beam Search Heuristics." *Management Science* 41 (5): 767–785.
Nakamura, M., N. Yamashiro, Y. Y. Gong, T. Matsumura, and K. Onaga. 2005. "Iterative Parallel Genetic Algorithms Based on Biased Initial Population." *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences* E88A (4): 923–929.
Olewnik, A., and V. G. Hariharan. 2010. "Conjoint-HoQ: Evolving a Methodology to Map Market Needs to Product Profiles." *International Journal of Product Development* 10 (4): 338–368.
Poles, S., Y. Fu, and E. Rigoni. 2009. "The Effect of Initial Population Sampling on the Convergence of Multi-objective Genetic Algorithms." *Multiobjective Programming and Goal Programming* 618: 123–133.
Prahalad, C. K., and R. A. Mashelkar. 2010. "Innovation's Holy Grail." *Harvard Business Review* Jul-Aug: 131–142.
Rafique, A. F., L. S. He, Q. Zeeshan, A. Kamran, and K. Nisar. 2011. "Multidisciplinary Design and Optimization of an Air Launched Satellite Launch Vehicle using a Hybrid Heuristic Search Algorithm." *Engineering Optimization* 43 (3): 305–328.
Sadrzadeh, A. 2012. "A Genetic Algorithm with the Heuristic Procedure to Solve the Multi-Line Layout Problem." *Computers & Industrial Engineering* 62: 1055–1064.
Sawtooth Software. 2008. CBC v.6.0. Sawtooth Software, Sequim, WA. http://www.sawtoothsoftware.com/download/techpap/cbctech.pdf
Sawtooth Software. 2009. The CBC/HB System for Hierarchical Bayes Estimation Version 5.0 Technical Article. Sawtooth Software, Sequim, WA. http://www.sawtoothsoftware.com/download/techpap/hbtech.pdf
Shiau, C.-S., and J. Michalek. 2009. "Should Designers Worry About Market Systems?" *Journal of Mechanical Design* 131 (011011): 1–9.
Shiau, C-S., I. H. Tseng, A. W. Heutchy, and J. Michalek. 2007. "Design Optimization of a Laptop Computer Using Aggregate and Mixed Logit Demand Models with Consumer Survey Data." Proceedings of the 2007 IDETC, Las Vegas, NV, DETC2007-34883.
Simpson, T. W., J. R. A. Maier, and F. Mistree. 2001. "Product Platform Design: Method and Application." *Research in Engineering Design* 13 (1): 2–22.
Smith, G. C., and S. S.-F. Smith. 2003. "Automated Initial Population Generation for Genetic Assembly Planning." *Informational Journal of Computer Integrated Manufacturing* 16 (3): 219–228.
Srivastava, R., and K. Deb. 2013. "An Evolutionary Based Bayesian Design Optimization Approach under Incomplete Information." *Engineering Optimization* 45 (2): 141–165. doi:10.1080/0305215X.2012.661730.
Steiner, W., and H. Hruschka. 2003. "Genetic Algorithms for Product Design: How Well Do They Really Work?" *International Journal of Market Research* 45 (2): 229–240.
Sudharshan, D., J. H. May, and A. D. Shocker. 1987. "A Simulation Comparison of Methods for New Product Location." *Marketing Science* 6 (Spring): 182–201.
Sullivan, E., S. Ferguson, and J. Donndelinger. 2011. "Exploring Difference in Preference Heterogeneity Representation and Their Influence in Product Family Design." Proceedings of the 2011 IDETC, Washington, DC, DETC2011-48636.
Tang, Y., J. Chen, and J. Wei. 2013. "A Surrogate-Based Particle Swarm Optimization Algorithm for Solving Optimization Problems with Expensive Black Box Functions." *Engineering Optimization* 45 (5): 557–576. doi:10.1080/0302215X.2012.690759.
Thakur, L. S., S. K. Nair, K. W. Wen, and P. Tarasewich. 2000. "A New Model and Solution Method for Product Line Design with Pricing." *Journal of the Operational Research Society* 51 (1): 90–101.
Thurston, D. 1991. "A Formal Method for Subjective Design Evaluation with Multiple Attributes." *Research in Engineering Design* 3 (2): 105–122.
Togan, V., and A. T. Daloglu. 2008. "An Improved Genetic Algorithm with Initial Population Strategy and Self-Adaptive Member Grouping." *Computers and Structures* 86: 1204–1218.
Train, K. E. 2003. *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.
Tsafarakis, S., Y. Marinakis, and N. Matsatsinis. 2011. "Particle Swarm Optimization for Optimal Product Line Design." *International Journal of Research in Marketing* 28 (1): 13–22.
Turner, C., S. Ferguson, and J. Donndelinger. 2011. "Exploring Heterogeneity of Customer Preference to Balance Commonality and Market Coverage." Proceedings of the 2011 IDETC, Washington, DC, DETC2011-48581.

Villeneuve, F. J., and D. M. Mavris. 2012. "Aircraft Technology Portfolio Optimization using Ant Colony Optimization." *Engineering Optimization* 44 (11): 1369–1387.
Wang, X., J. Camm, and D. Curry. 2009. "A Branch-and-Price Approach to the Share-of-Choice Product Line Design Problem." *Management Science* 55 (10): 1718–1728.
Wassenaar, H., and W. Chen. 2003. "An Approach to Decision Based Design with Discrete Choice Analysis for Demand Modeling." *Journal of Mechanical Design* 125 (3): 490–497.
Wassenaar, H., W. Chen, J. Cheng, and A. Sudjianto. 2004. "An Integrated Latent Variable Choice Modeling Approach for Enhancing Product Demand Modeling." Proceedings of the 2004 DETC, Salt Lake City, UT, DETC2004/DTM-57487.
Williams, N., P. K. Kannan, and S. Azarm. 2011. "Retail Channel Structure Impact on Strategic Engineering Product Design." *Management Science* 57 (5): 897–914.