PREFERENCE COORDINATION IN ENGINEERING DESIGN DECISION-MAKING

by

Jeremy J. Michalek

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in The University of Michigan
2005

Doctoral Committee:

Professor Panos Y. Papalambros, Chair
Professor Richard D. Gonzalez
Associate Professor Fred M. Feinberg
Assistant Professor Steven J. Skerlos

ACKNOWLEDGMENTS

First I would like to acknowledge direct contributions to the content of the work presented in this dissertation. The efficient choice-based conjoint design for the dial-readout scale used in chapters 4, 5, and 6 was generated by Feray Adigüzel using SAS macros written by Professor Warren Kuhfeld, and Laura Stojan created the graphics to translate the design to an online survey as well as assisting with the engineering model and collecting data for the manufacturing model in Chapter 6. Peter Ebbes also worked with Feray Adigüzel to develop, code, and run the Bayesian MCMC chains for the heterogeneous model in Chapter 5, building off of code written by Professor Peter Lenk, and both of them contributed greatly to my understanding of choice-based conjoint and Bayesian methods. Oben Ceryan contributed to the manufacturing formulation in chapter 6, and Sarah Bahrman and Carl Lenker contributed to the cost and demand models developed in Chapter 7, in addition to providing thoughtful insight to the methodology.

I was extremely fortunate to have a remarkably engaged dissertation committee. Professor Fred Feinberg was instrumental in developing my understanding of current marketing practice and determining how best to coordinate this with engineering. He contributed directly to the publications leading to Chapter 4 and Chapter 5, in addition to spending uncountable hours talking with me about choice theory, conjoint analysis, Bayesian econometrics, marketing research, and mathematics, as well as politics, social theory, feminist theory, post-modernism, multiculturalism, surrealism, consumerism, anti-intellectualism, existentialism, essentialism, social constructionism, hegemony, foreign policy, Noam Chomsky, Howard Zinn, Camille Paglia, Catharine MacKinnon, Ali G, Art Tatum, Rachmaninoff, the dietary preferences of his cat Grazzi, and just about any interesting discussion topic I can think of. The depth of his knowledge and thoughtfulness in such a broad array of subjects would make any Renaissance man blush.

Professor Steve Skerlos was instrumental in developing my understanding of environmental impact and environmental policy, as well broadening and deepening my understanding of products throughout their life-stages and the consequences of design decisions in each of these stages. He contributed directly to the publications leading to Chapter 7, in addition to inspiring me to take a leadership role for educational initiatives in Engineers Without Borders at the University, as well as generally working for a more sustainable world through informed decision-making. I am grateful for a critical pep-talk that he provided at a time when I thought I would never find the right thesis topic. Professor Richard Gonzalez was also involved throughout this work, introducing me to statistical psychology perspectives on human perception, decision-making, and choice, and providing feedback and thoughtful challenges as well as excellent references for this research. He acted to clarify a number of differences in language, background, and philosophies among the disciplines relevant to this work. Finally, Professor Panos Papalambros was involved in every aspect of this dissertation, contributing directly to each chapter; offering a blinding depth of knowledge in design optimization; providing advise and feedback while affording me a tremendous amount of freedom and responsibility; developing my sense of the academic literature and community in engineering design; providing challenges and encouragement; reminding me to take a break now and again; promoting a respect for rigorous work; providing friendly warmth and an impressive breadth of knowledge and etymological expertise: always reminding me of the many Greek origins of Western language and culture. I admire his openness to constantly push the boundaries of his expertise while maintaining extraordinary standards of quality. Panos is the reason that I decided to attend The University of Michigan for graduate school, and his guidance and support throughout the process were invaluable.

In addition to those who contributed directly to the content of this work, a great number of people contributed meaningfully by providing the support, feedback, and perspective that helped to shape my ideas and define my direction. First, I must thank my wonderful loving partner, Molly Bain Frounfelter, for her boundless support, ideas, challenges, patience, and perspective. She has broken through many a barrier in my thinking in a way that has had direct impact on my work, as well my life. Her depth of thought is inspiring, and her critical eye combined with her openness to unconventional ideas makes every conversation with her a learning experience. She provided support for my humanist values and desire to make work meaningful with respect to them, and she deserves credit for a tremendous amount of intellectual growth that influenced this work. I owe my gratitude to a plethora of other individuals for their roles in my work: W. Ross Morrow for his mathematical, conceptual, and artistic genius, and his love of good music, the entire Optimal Design Laboratory, past and present, especially Professor Ruchi Choudhary, Dr. Hosam Fathy, Erin MacDonald, Kuei-Yuan (Miles) Chan, Dr. Michael Kokkolaras, Dr. Panayiotis Georgiopoulos, Professor Hyung Min (Harrison) Kim, and Professor Jean-François Petiot who spent much time plugging through ideas with me. My thanks also goes to Professor Bryan Pardo, Brian Magerko, John Hawkins, Karim Hamza and Professor Brian Jensen, always willing to work through ideas or argue a point, and to Professor Michel Wedel and Professor Kazu Saitou who provided helpful knowledge. Tthe support of the entire National Science Foundation Engineering Research Center for Reconfigurable Manufacturing Systems at The University of Michigan is appreciated, especially Professor Yoram Koren, Dr. Zbigniew Pasek, Dr. Farshid Asl, Professor Theodor Freiheit, and Professor Derek Yip-Hoi. The Antilium research group at The University of Michigan also provided helpful feedback on my work, as well as a broad, interdisciplinary perspective on design. Thanks is due also to the superb faculty, staff, and student leaders at The University of Michigan who truly provide an extraordinary environment for learning, with an abundance of interesting courses, invited lecturers, educational resources, interdisciplinary initiatives, and opportunities for attainment and advancement of broad and deep knowledge and understanding, as well as action. The Engineers Without Borders (BLUElab) group and the Engineering for Community course development committee helped develop my understanding of the role in society of engineering in general, and of my work in particular. I appreciate also the support of my family: my father, for early on teaching me how to learn by entrusting me with many privileges and responsibilities while always making me feel that my ideas and viewpoints were valued and respected; my mother, who remembers the times when she was "too busy to pee" and has been supportive through the most stressful and busy times; my sister, who provides me love, perspective, and friendship; Patti, Jim, and James Frounfelter for their love and the incredible culinary and theatrical skills they share with me to refresh my mind; and my extended family as well for their encouragement throughout my life. I'd also like to thank all of my friends who have provided support and encouragement, as well as Richie Hawtin, Zack de la Roche, John Stewart, Lewis Lapham, Richard James, Darren Emerson, Ludovic Navarre, Sean Booth and Rob Brown for getting me through the tough times.

Finally, this work was supported by the Antilium Interdisciplinary Project of the Rackham Graduate School, the National Science Foundation Engineering Research Center for Reconfigurable Manufacturing Systems, and the Automotive Research Center at the University of Michigan. The support of these organizations is appreciatively acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

The following nomenclature is used consistently in the dissertation wherever possible, with the exception of Chapter 3, which uses its own set of symbols to discuss the ATC convergence proof.

$\lVert \cdot \rVert_2^2$	Square of the l_2 norm of a vector
$\langle \cdot \rangle_\alpha$	Element α of a vector
$\circ$	Element-by-element multiplication of vectors
b	Mixing component index for mixture model
b_τ	Base engine size for topology τ
B	Number of mixing components
c_k	Total cost for producer k
c^{B}	Base unit manufacturing cost
c_j^{E}	Unit material cost for product j achieved by engineering design. Also used as engine cost in Chapter 7.
c^{I}	Investment cost
c_m^{I}	Investment cost to purchase one machine of type m
c^{O}	Operating cost
c_m^{O}	Hourly operating cost of machine type m
c_j^{P}	Total production cost for product j
c_j^{M}	Unit target material cost for product j set by marketing
c_k^{R}	Regulation cost for producer k
c_l^{S}	Unit material cost function for component type l
c^{V}	Product unit variable cost
C^{P}	Cost incurred by manufacturing
C^{M}	Cost target set by marketing
d	Lifetime vehicle miles traveled
D	Deviance Information Criterion (DIC) deviance function
f_τ	ADVISOR vehicle simulation function for engine τ
$\mathbf{g}$	Vector function of engineering design inequality constraints
$\mathbf{g}_{lmn}$	Vector function of machine production inequality constraints for operation n of component l on machine m
$\mathbf{h}$	Vector function of engineering design equality constraints
i	Consumer index
I	Number of individual consumers
I_{D}	Number of individual consumer draws from the Bayesian mixture model

I_S	Number of survey respondents
j	Product index
J_k	Number of products in producer k's product line
$\mathcal{J}_k$	Set of products in producer k's product line
$\mathcal{J}_t$	Set of products in choice set t
k	Producer index
K	Total number of producers
l	Component index
L	Total number of components
m	Machine/equipment type index
M	Total number of machine types
n	Manufacturing process index
N_l	Total number of processes for component l
p_j	Price of product j
P_{ijt}	Probability that individual i chooses product alternative j on choice occasion t
q_j	Demand for product j
r_{lmn}	Rate (min/part) to execute operation n on component l with machine m
$\mathbf{r}$	Engineering design response vector function
S	Size of the market
s_i	Percentage of the market represented by individual/segment i
t	Choice set index for conjoint survey
T	Total amount of machine time available
T_{jm}	Time on machine m allocated to product j
u_{ij}	Utility of alternative j for individual i
v_{ij}	Observable component of utility of product j for individual i
v_{i0}	Utility of the outside good for individual i
V_j	Production volume of product j
V_j^{M}	Production volume target set by marketing for product j
V_j^{E}	Production volume achieved by engineering for product j
V_{jlmn}	Production volume of operation n performed by machine m on component l of design j
$\mathbf{w}_j$	Weighting coefficient vector for product characteristic consistency of product j
x_α	Element α of vector $\mathbf{x}_j$ for any general product
$\mathbf{x}_j$	Engineering design variable vector for product j
y_α	Element α of vector $\mathbf{y}$
$\mathbf{y}$	Vector of engineering model parameters
z_α	Element α of vector $\mathbf{z}_j^{\mathrm{E}}$ for any general product

$\mathbf{z}_j^{\mathrm{M}}$	Vector of product characteristic targets set by marketing for product *j*
$\mathbf{z}_j^{\mathrm{E}}$	Vector of product characteristics achieved by the engineering design of product *j*
α	Generic index
α_τ	Tons of CO_2 produced by a gallon of fuel for engine type τ
$\beta_{i\zeta\omega}$	Part-worth coefficient for consumer segment *i* for attribute ζ at level ω
$\boldsymbol{\beta}_i$	Part-worth coefficient vector compiling $\beta_{i\zeta\omega}$ for all ζ and ω
$\delta_{j\zeta\omega}$	Binary dummy variable for product characteristics and price: 1 if characteristic ζ of product *j* is set to level ω, 0 otherwise.
ε_{ij}	Random error component of utility for individual *i* with respect to product *j*
ζ	Index of product characteristics = {1, 2, ..., *Z*}, ($\zeta = 0$ refers to price)
Z	Number of product characteristics
$\boldsymbol{\theta}_b$	Mean value vector for mixing component *b*
κ_m	Number of machines of type *m*
$\boldsymbol{\Lambda}_b$	Covariance matrix for mixing component *b*
υ	Societal cost valuation per ton of CO_2 in US dollars
Ξ	Aggregation vector of all parameters in the Bayesian model
ξ_l	Number of components of type *l* in the design
Π	Profit
ρ	Penalty parameter for regulation violation
σ^2	Variance of the error terms ε
φ_i	Group membership of individual *i* in the MCMC chain
τ_j	Topology of product *j*
ϕ	Minimum diesel sales percentage required by quota
Φ_{ijt}	Binary dummy variable indicating observed choice of individual *i* with respect to alternative *j* on choice occasion *t*
$\Psi_{i\zeta}$	Spline function of beta for individual *i* and product characteristic/price ζ
ω	Product characteristic level index = {1, 2, ..., Ω_ζ} for product characteristic ζ
Ω_ζ	Number of discrete levels for product characteristic ζ

ABSTRACT

PREFERENCE COORDINATION IN ENGINEERING DESIGN DECISION-MAKING

by

Jeremy J. Michalek

Chair: Panos Y. Papalambros

Whether interested in profit or in social welfare, designers are concerned with the preferences people have and the choices they make. Tools such as Quality Function Deployment have been developed to help designers organize thinking about the relationship between design decisions and stakeholder preferences; however, work incorporating explicit quantitative models of stakeholder preferences into engineering design decision making is still sparse.

Preference coordination draws on theory and methods from the marketing, economics, and psychology literatures to model preference structures and to coordinate them effectively with design models of engineering feasibility and performance for achieving jointly optimal solutions with both technical and market feasibility. This process resolves tradeoffs among competing technical objectives while ensuring that product targets based on market preferences are physically realizable.

Specifically, theory is reviewed and developed for analytical target cascading (ATC), a methodology for decomposing a system into a hierarchy of subsystems and coordinating optimization of each subsystem so as to achieve the joint solution. The ATC methodology is then applied to coordinate marketing and engineering design decision models in a profit-seeking firm. It is demonstrated with a case study that the joint solution obtained through coordination is superior to the solution obtained by treating each discipline independently. The modularity of the framework facilitates extensions, and two

such extensions are pursued: First, the methodology is extended for product line design by coordinating preference models that capture heterogeneity with a set of engineering design models. Second, manufacturing decisions are incorporated by adding a module to coordinate machine investment and allocation decisions. Finally, the scope of preferences is expanded to explore social preferences as expressed through regulation: Game theory is used to predict the design decisions made by profit-seeking producers in a competitive marketplace, and the effects of different regulation scenarios on the resulting decisions are examined.

It is the hope that the methods developed in this dissertation for modeling stakeholder preferences and coordinating with engineering design decision-making will help design engineers and managers to understand the relationship between their decisions and the interests upon which they have impact so that better, more informed decision-making can be realized.

CHAPTER 1

INTRODUCTION

Whether interested in profit or in social welfare, designers are concerned with the preferences people have and the choices they make. Tools such as Quality Function Deployment (QFD) have been developed to help designers organize thinking about the relationship between design decisions and stakeholder preferences; however, much work remains to be done incorporating explicit models of stakeholder preferences in engineering design decision-making.

Disparities between the scopes, perspectives, models and tools of the various disciplines involved in the product development process, both in industry and in academia, can lead to failures in understanding user wants and translating them into viable products, as illustrated by the cartoon in Figure 1.1 (adapted from Dieter, 1991). In particular, Krishnan and Ulrich (2001) outline the different scopes, domains, perspectives, and goals of several academic disciplines that work in product development. Table 1.1 lists those from marketing and engineering design: Marketing researchers model products as "bundles of attributes" from which consumers gain utility, and they use these models to determine appropriate pricing and attribute levels to achieve desirable product positioning, market share, and profit. In contrast, engineering design researchers model products using parametric models of technical performance and use these models to determine appropriate product configurations and dimensions to achieve high performance at low cost. Figure 1.2 summarizes these basic differences, adding also the manufacturing focus on process efficiency.

Table 1.1: Comparison of marketing and engineering design perspectives

	Marketing	Engineering Design
Perspective on Product	"a bundle of attributes"	"a complex assembly of interacting components"
Typical Performance Metrics	"Fit with market", market share, customer satisfaction, profit	"Form and function", technical performance, innovativeness, cost
Representational Paradigm	Customer utility *as a function of* product attributes	Geometric models, parametric models of technical performance
Common Decision Variables	*Consumer-oriented* "attribute levels", price	Product size, shape, configuration, function, dimensions
Critical Success Factors	Positioning, pricing, discerning / meeting customer needs	Creative concept and configuration, performance optimization

Within engineering design, the design optimization paradigm has arisen as a dominant perspective on formal decision-making: A set of design alternatives (typically infinite in number) is described mathematically by parameterizing a design concept into a set of decision variables; design objectives and constraints are written as explicit functions of the variables using physical, geometric, or empirically determined mathematical relationships; and a general analytical or iterative method is used to search the space of the decision variables for a point that satisfies the constraints and maximizes the objective. Decades of research have provided a rich literature of optimization

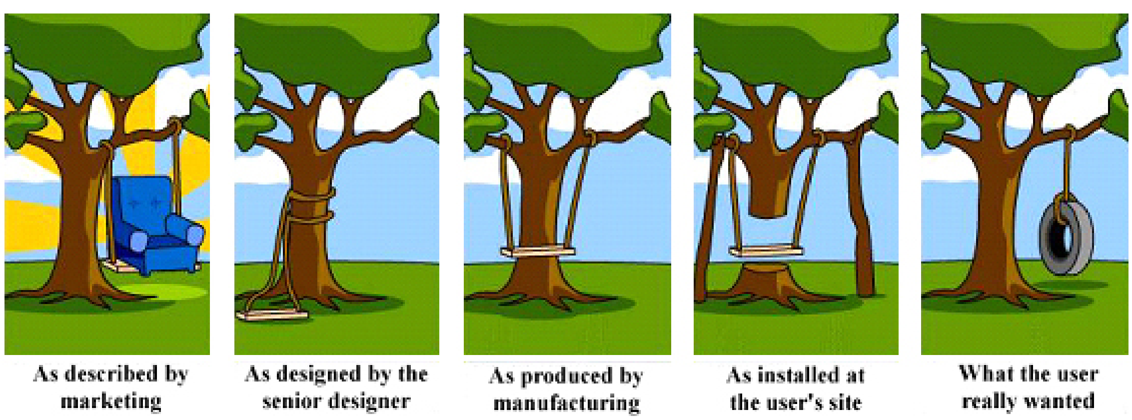


Figure 1.1: Product failures can result from disjoint decision-making

methodologies, tools, and successful applications in a range of engineering disciplines, and many classes of problems can be solved easily with existing methods. One remaining open question is how to deal with design situations that have multiple conflicting objectives. In practice, most design problems have conflicting objectives, and modelers have various techniques for managing them within the framework of the single-objective paradigm. For example, one objective can be selected, and the remaining objectives can be reformulated as constraints ensuring that each is achieved within some user-defined threshold. The design can then be optimized for the remaining objective, and the effects of changing user-defined threshold values can be explored through parametric studies. Alternatively, the modeler can use techniques to approximate and navigate the *Pareto set*: the set of designs that cannot be improved in any one objective without sacrificing another objective (Papalambros and Wilde, 2000). Figure 1.3 shows a Pareto set for two hypothetical minimization objectives.

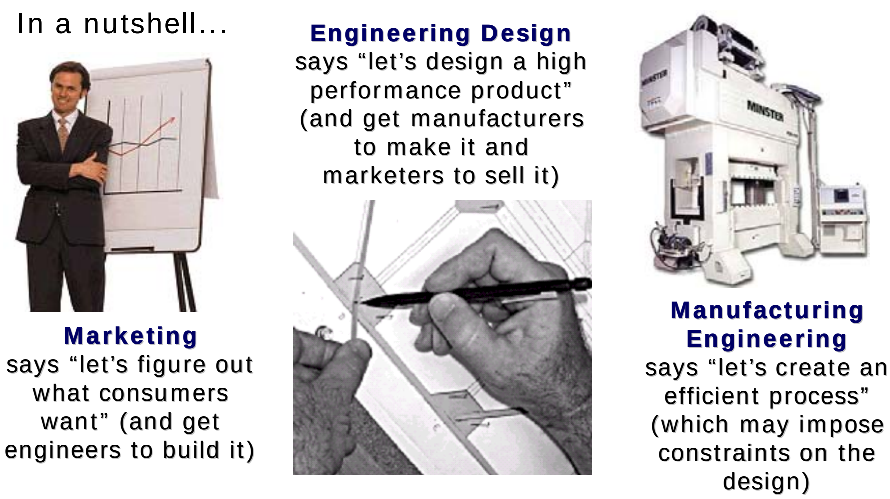


Figure 1.2: Perspectives in product development

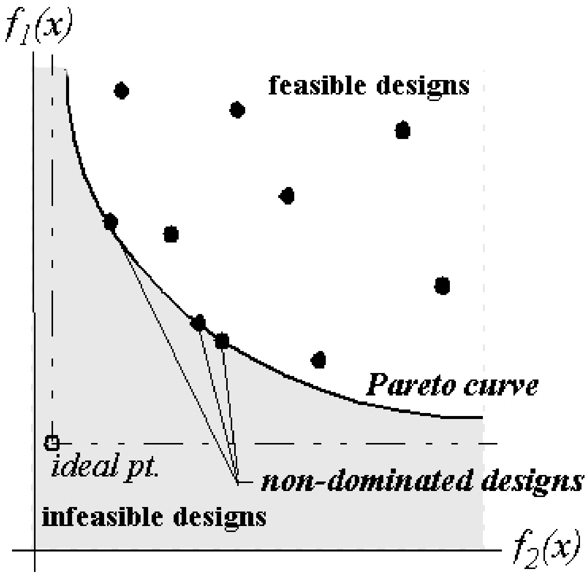


Figure 1.3: A Pareto set for two competing objectives

The Pareto set is typically generated by combining the various objectives into a single aggregate objective function that is a linear weighted sum of the individual objectives. Any specific choice of weighting coefficients yields a specific point on the Pareto set, and if the Pareto set is convex it can be mapped out as the set of points achieved under all possible weighting coefficient combinations. The choice of a single design from the Pareto set requires explicit declaration of the preferences for tradeoffs among objectives, and designers typically make this selection intuitively, without supporting tools or quantitative information. However, a vast literature exists in marketing, psychology, and econometrics for defining models of choice and designing experiments to efficiently elicit preference structures using survey or market data. Yet, when marketers use these methods to make decisions, they optimize with respect to downstream product characteristics, those qualities observed by the consumer, rather than detailed design decisions, and they lack models to describe which combinations of product characteristics can be achieved with a realizable product design and which are unattainable. A method for efficiently and effectively coordinating these two modeling domains would provide engineering designers with quantitative tools to assist in making tradeoff decisions among competing technical objectives through coordination with a higher-level objective as well as indirectly provide marketing with access to models of cost and feasibility constraints so that they can select not just the most desirable product characteristics, but the most desirable, feasible and realizable product design.

Secondly, while engineering designers tend to think of tradeoffs in terms of the Pareto paradigm of resolving competing technical objectives, tradeoffs also exist among the preferences and interests of the various stakeholders who are affected by the designer's decisions. For instance, while the tradeoff between two objectives that are presumed to be desirable by all, such as high vehicle fuel economy vs. high performance, may be adequately described by the Pareto paradigm, the tradeoff between characteristics over which preferences vary across the population, such as vehicle size and ride height, does not fit cleanly into this representation. Furthermore, tradeoffs between private consumer preferences, such as valuing the performance, size, and utility of vehicles, and public social preferences, such as valuing environmental impact reduction, take place even within the preferences of individual stakeholders. When preferences are heterogeneous in the population and the interests of stakeholders conflict, preference coordination can help designers to understand the impact of their decisions toward resolving conflict.

Recently, engineering design researchers, both in academia and in industry, have begun to explore quantitatively the link between engineering design decisions and firm profitability using decision theoretic models incorporating uncertainty. I have taken an approach that differs from the main thrust of this work: Rather than attempting to combine every model into a single decision-making optimization loop solved "all-at-once" (AAO), I modify, develop and apply the analytical target cascading (ATC) methodology to coordinate decision-models from these various disciplines. Figure 1.4 provides a graphical schematic of the decomposition for a dial-readout scale product line case study examined later in the dissertation. ATC allows separation of models by discipline while clearly defining interfaces and interactions through a process of target setting and matching so that each discipline can focus on building models within its area of expertise. ATC then provides a rigorous framework for coordinating these models in order to reach agreement between the disciplines and obtain a jointly-optimal solution. In addition, coordination of models, rather than combining them into a single AAO model, facilitates tractability in optimization because each subproblem is typically solved in a smaller design space and with fewer constraints than the AAO model.

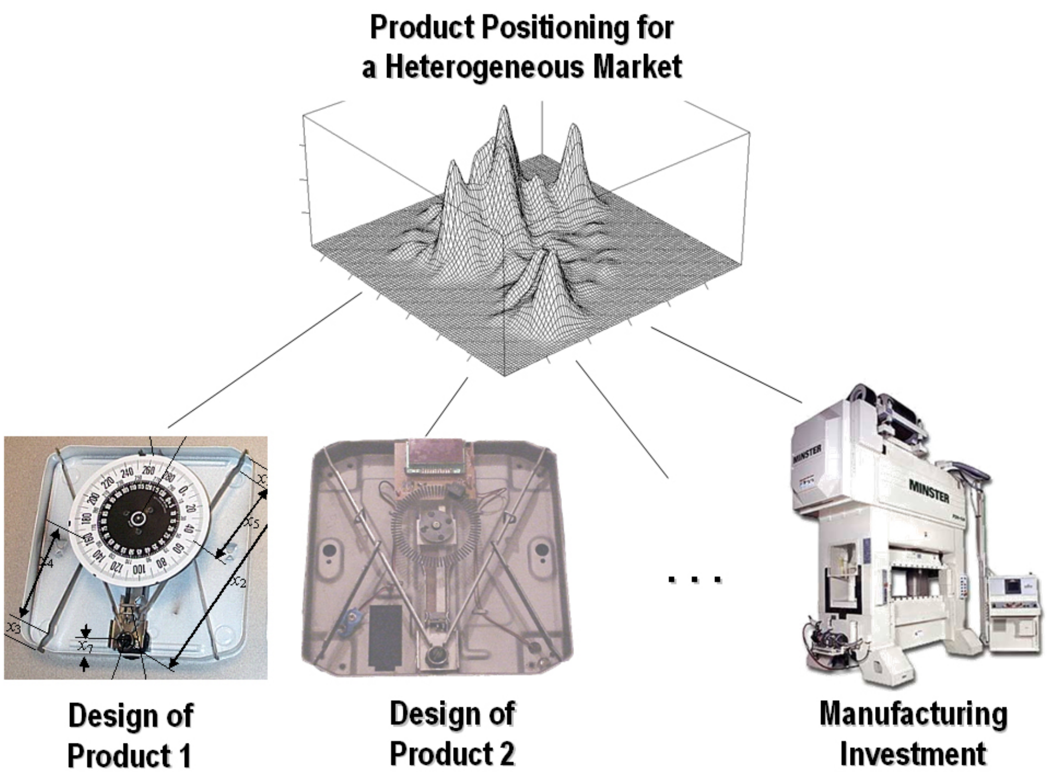


Figure 1.4: Decomposition and coordination of decision-models for product development

The development of concepts for preference coordination proceeds with an overview of the relevant literature in Chapter 2. In Chapter 3 the analytical target cascading methodology and its convergence properties are introduced, the theory is further developed, and a technique is proposed for applying the methodology to the problems with properties like those of interest in this dissertation. In Chapter 4, drawing on established econometric models of choice and efficient experimental design procedures, data-driven models of consumer demand are built and coordinated with parametric engineering models of product performance to achieve a joint solution. A consumer product case study of dial readout scales demonstrates that the methodology provides substantial improvement over disjoint decision-making. The methodology is then extended in Chapter 5 to consider heterogeneity of preferences in the marketplace using hierarchical Bayesian mixture models in order to design an optimal line of products for the heterogeneous market. The methodology is further extended in Chapter 6 to incorporate manufacturing investment decisions and explore tradeoffs between achieving desirable product characteristics to increase market share versus compromising the design to reduce cost. In Chapter 7 the perspective is expanded beyond profit maximization of a single firm to explore the tradeoffs between private consumer preferences and public social preferences. A case study is explored modeling the effects of vehicle emission regulation policy on the design decisions of profit-seeking automobile producers in a competitive marketplace, and the impact on producers, users, and society is examined. Finally, a summary and conclusions are provided in Chapter 8.

1.1 REFERENCES

[1.1] Dieter, G.E. (1991) *Engineering Design*, McGraw-Hill.

[1.2] Krishnan , V., and K. Ulrich (2001) "Product development decisions: a review of the literature," *Management Science* v47 n1 p1-21.

[1.3] Papalambros, P.Y., and D.J. Wilde (2000) *Principals of Optimal Design – Modeling and Computation*, Second Edition, Cambridge University Press.

CHAPTER 2

BACKGROUND

There exist a number of different models in the engineering design, marketing, and economics literature for positioning and designing products and product lines. These various modeling perspectives use overlapping terms (design, attributes, utility, etc.) that are sometimes used ambiguously or differently by different authors, despite efforts to use terms consistently and specifically. This chapter attempts to organize the relevant literature using a simplified reference framework based on Kaul and Rao (1995) and Hauser and Simmie (1981), who offer discussions about integrating the economics perspective, epitomized by Lancaster (1966), with the psychometrics perspective (which uses statistical tools such as multidimensional scaling to understand product perception in an abstract attribute space). To this framework, a design decision vector has been added to represent technical decisions that a designer must make, but which are not necessarily directly observable to the user (for example, bore and stroke size in an engine). In this updated framework, physical, observable characteristics are functions of the design decisions. The resulting framework, shown in Figure 2.1, is intended to be used as a basis to organize and understand existing literature and to describe how the contributions in the proposed research relate to the literature. This framework does not address every relevant aspect; for example, price, cost, production volume, profit, and regulation are not discussed; however, it is effective as a basic structure for understanding the literature.

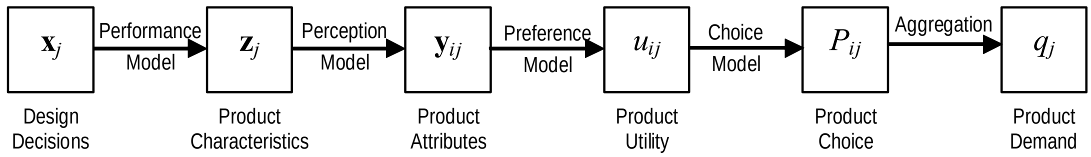


Figure 2.1: General framework

In Figure 2.1, a product j is described parametrically by a vector of design decision variables $\mathbf{x}_j$, which represent the design choices that must be made to specify a product. The vector $\mathbf{x}_j$ is the design variable vector that is commonly used in engineering design optimization models. Examples of design decision variables might include the bore and stroke dimensions of cylinders in a vehicle engine. Additionally, the design decision vector $\mathbf{x}_j$ is generally subject to a set of technical constraint functions, $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, which ensure that the design variable vector maps to a feasible design.

The parametric engineering performance model interprets the design variable vector $\mathbf{x}_j$, mapping it onto objective, measurable, physical characteristics $\mathbf{z}_j$ that are observed by the consumer when making purchasing decisions. The vector $\mathbf{z}_j$ is the product characteristic vector used in marketing product design models to define a product, where each characteristic in the vector is usually restricted to a set of discrete 'levels.' Examples of product characteristics might include the gas mileage and horsepower of a vehicle. The mapping from $\mathbf{x}_j$ to $\mathbf{z}_j$ is typically derived from physics models and physical and geometric relationships.

While the product characteristics $\mathbf{z}_j$ of product j are objective and measurable, perceptions of the product vary among individuals i. The perceptual model for each consumer i maps the product's characteristics $\mathbf{z}_j$ to a point in the perceptual attribute space $\mathbf{y}_{ij}$. The vector $\mathbf{y}_{ij}$ is the perceived product attribute vector used in marketing product positioning models. In some models, the perceptual attribute space is described using dimensions that correspond to perceptual adjectives such as aggressiveness or simplicity. However, in most models, the perceptual attribute space is an abstract space that is generated using factor analysis, principal component analysis, or multidimensional scaling, as discussed later, so that each dimension in the abstract attribute space $\mathbf{y}_{ij}$ represents a combination of the product characteristic dimensions $\mathbf{z}_j$. The resulting space represents a "best fit" reduction of the product characteristic space into a reduced perceptual attribute space, where similar products are close together. Typically, these methods are used primarily to develop intuition about the market and about a large space of characteristics; however, some authors have gone farther to develop mappings for use in optimization.

The preference model for each individual i maps a product with a perceived product attribute vector $\mathbf{y}_{ij}$ into a scalar utility value u_{ij}. The preference model defines which attribute values are desirable and the relative importance of each attribute to individual i such that products with higher utility values are preferred over products of lower utility. Typically preference functions are estimated by observing choices or stated preferences and inferring statistically the parameters of the mapping function that best fit the observed behavior.

Choice of which product to purchase is a function of the utility values of the individual products. P_{ij} represents the probability that consumer i will choose product j. The choice model maps utility values of the set of available products into choice probabilities for each of the products P_{ij} such that products of higher utility are more likely to be chosen[1]. Neglecting issues of product distribution and product availability, if the sample of individuals is representative of the population, the set of consumer choices P_{ij} can be summed over the individuals and scaled by the size of the market to calculate total demand q_j for product j.

[1] Probabilistic choice may be assumed without loss of generality because deterministic choice is a special case where the product of highest utility has choice probability equal to one and all other alternatives have choice probability of zero

2.1 APPROACHES IN THE LITERATURE

2.1.1 Engineering Design Optimization

Engineering design optimization models typically use physics- and geometry-based models to calculate performance related product characteristics $\mathbf{z}_j$ as functions of detailed design variables $\mathbf{x}_j$. The design variables are then optimized to achieve best performance within engineering feasibility constraints, for example, maximizing the value of a particular product characteristic or minimizing deviation from a desired target value. The optimization objective is assumed to be defined by the modeler in a way that is desirable, and target values for $\mathbf{z}_j$ are set exogenously.

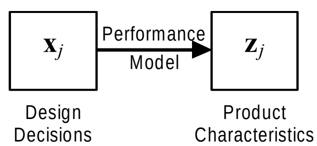


Figure 2.2: Engineering design optimization models

In cases where the product characteristics of interest consist of multiple competing performance objectives, the tradeoffs among the objectives are typically described with a Pareto set, as described in Chapter 1, and preference among objectives is defined using weighting coefficients, which are set intuitively and interactively by the designer without models of how these tradeoffs affect downstream goals. Familiarity with basic design optimization modeling and nonlinear programming methods are assumed in the remainder of this dissertation, although advanced methods such as decomposition and coordination methodologies will be discussed in detail. For a comprehensive overview of design optimization modeling and solution theory with focus on nonlinear programming and variables with continuous domains, see Papalambros and Wilde, 2000.

2.1.2 Decomposition and Coordination Optimization Methods

Optimization problems are typically solved "all-at-once" (AAO), meaning that all design variables $\mathbf{x}_j$ are manipulated simultaneously during search. For large systems and systems involving many disciplines, the $\mathbf{x}_j$ space may have high dimensionality, making optimization difficult or impossible. A group of methods have been proposed for decomposing the full optimization problem into a set of smaller subproblems and coordinating the solutions to those subproblems efficiently and effectively. Most of these methods are described as Multidisciplinary Design Optimization (MDO) because they are intended for decomposition by discipline (Braun, 1996; Sobieski *et al.,* 1985; Cramer *et al.*, 1994; Sobieski and Kodivalam, 1999; Sobieski, 1989). The analytical target cascading (ATC) methodology (Kim, 2001) is similar in that a large problem is decomposed into subproblems; however, the subproblems in ATC are organized into a hierarchy representing the subsystems and components in the design of a large system. The components of this hierarchy are related by passing targets from higher-level systems to be achieved by lower level subsystems and rebalancing those targets based on the responses realized by the subsystems. In this dissertation, ATC is used as a tool for coordinating models from different disciplines in the product development process, including marketing product planning, engineering design, and manufacturing. MDO methods and ATC are discussed in detail in Chapter 3, and specific implementations of ATC are proposed in Chapter 4, Chapter 5, and Chapter 6.

2.1.3 Decision Based Design

Recently, one particular group of researchers in the engineering design optimization community has used decision theory to model product design decisions in the context of product demand and the producer's economic goals, with a specific focus on modeling the effects of uncertainty. This line of work was termed Decision-Based

Design (DBD) by Hazelrigg (1988), although it should not be confused with viewing design as a decision-making process, which is the ubiquitous paradigm in design optimization. In the DBD literature (such as Marston and Mistree, 1998; Li and Azarm, 2000; Gupta and Samuel, 2001; and Wassenaar and Chen, 2003), product characteristics are calculated as functions of the design variables and a vector of stochastic parameters that affect the design. Demand is calculated as a function of the product characteristics and price, although the original framework proposed by Hazelrigg does not specify a method or model for doing this, and different authors have used different techniques. Most work in this area focuses on decision theory and multi-attribute utility theory (Von Neumann and Morgenstern, 1944) rather than on implementation- and data-focused models from marketing and econometrics, although Wassenaar and Chen (2003) have used discrete choice analysis for demand prediction.

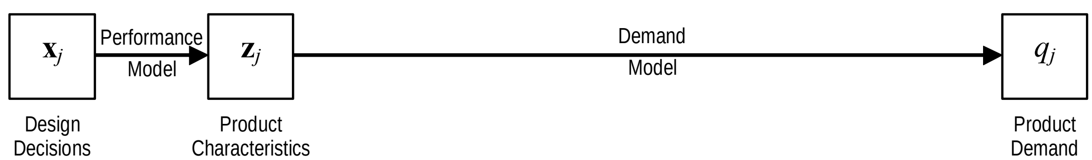


Figure 2.3: Decision-based design models

In the DBD literature, the concept of utility is applied to the utility of the *designer* (the decision-maker), rather than the utility of *consumers* in the market who choose whether or not to purchase. The designer's utility function serves to rank-order all possible outcomes (based on resulting profit, market share, etc.) in order of preference to the designer, and the goal is to optimize design decisions with respect to the designer's utility (i.e., to find the decisions that result in the most preferred outcome). Essentially, the concept of designer's utility is equivalent to enforcing that the designer specify a single objective function to be maximized. Specific DBD applications relevant to design of a single product are detailed in Chapter 4, and those applied to the design of a line of products are detailed in Chapter 5.

2.1.4 Conjoint Analysis and Design of Experiments

Conjoint analysis is a ubiquitous method in marketing for developing efficient experimental designs (what engineers call design of experiments) to estimate the effects of product characteristics (and price) on product utility, expressed through rating, ranking, or choosing among alternatives (Green and Srinivasan, 1978, 1990; Green, Wind and Rao, 1999; Kuhfeld, 2003). The method is called conjoint analysis by marketing and psychology researchers because the effects of various product characteristics are estimated jointly, i.e., comparing full product descriptions involving multiple characteristics rather than asking respondents separately about their preferences for each characteristic. In engineering experimental design, the existence of multiple characteristics is assumed because it is not generally possible to design an experiment to test the effect of one physical parameter in the absence of other relevant physical parameters. Physical quantities will always be present in the experiment, whether or not they are measured. With survey questions, we can choose which characteristics "exist in the experiment" by choosing which characteristics are included in the survey.

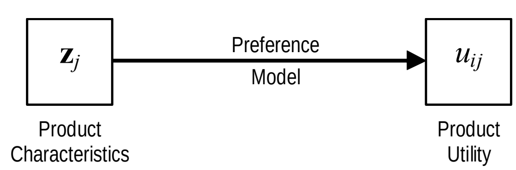


Figure 2.4: Conjoint analysis models

The most common methodology used in design of experiments is to discretize the range of each relevant independent variable into a set of discrete levels and measure the effect of systematically changing these levels on the dependent variable. If possible, a full factorial experiment can be performed measuring the dependent variable under all possible combinations of each independent variable at each level, and the results of the experiment can be used to estimate the main effects and interaction effects of each level of each independent variable on the value of the dependent variable. In practice, full factorial designs often involve many combinations, many experiments, and high cost. Fractional factorial experiments offer an efficient solution. If some of the higher-order interaction effects can be considered negligible, then a fraction of the full factorial experiment can be used to estimate the remaining effects. Without testing all possible combinations, some of the effects will be confounded, i.e., it will be impossible to distinguish one effect from another. Fractional factorial experiments are designed so that the confounding involves higher order effects, which can often be considered negligible. For example, if a particular main effect of one characteristic is confounded with a third-order interaction effect between three characteristics, and if the third-order effect can be assumed negligible, then the estimate of the main effect plus third-order effect is approximately equal to the main effect alone. This is the fundamental principle behind fractional factorial experiments. Wu and Hamada (2000) provide an excellent introduction to theory and concepts.

Typically in conjoint analysis applications the utility is taken to be a linear combination of the main effects of each characteristic at each level, and all interaction terms are neglected (although more advanced designs are also used). Requiring estimates of only main effects allows significant reduction in the size of the fractional factorial experiments, resulting in surveys of manageable size. In the conjoint literature, main effects are referred to as a 'part-worths' because each expresses the component of utility deriving from a specific characteristic at a specific level. The purpose of conjoint analysis is to design an experiment to estimate the preference model mapping product characteristics $\mathbf{z}_j$ to utility u_{ij}; however, the dependent variable, utility, is not directly observable. The observable aspect is some function of utility: either ranking, rating, or choice among alternatives. Given an assumed functional relationship between utility and the observed aspect, the parameters mapping $\mathbf{z}_j$ to u_{ij} can be chosen to best fit the observed data. In particular, choice-based conjoint analysis makes use of discrete choice models as the functional relationship between utility and observed choice, as discussed below. The power of conjoint analysis is that it enables estimation of preferences on all combinations of product attributes by asking for comparisons of a few carefully selected alternatives.

2.1.5 Random Utility Discrete Choice Models

Random utility discrete choice models map the utility values of a set of products onto probabilities of choosing each product from the set (Guadagni and Little, 1983; Gonul and Srinivasan, 1993; Manrai, 1994; Chintagunta, 1994; Berry, 1994; Ben-Akiva and Boccara, 1995).

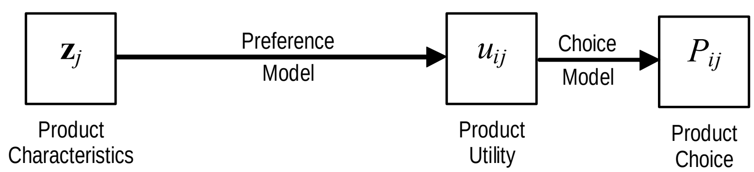


Figure 2.5: Discrete choice analysis models

Random utility choice models, such as logit and probit, are discussed in Chapter 4, Chapter 5, and Chapter 7, and an introduction with detailed derivations is provided in the Appendix. The basic idea is that utility is assumed to be partly observable as a function of objective, measurable characteristics, represented by a deterministic value v_{ij}, and partially unobservable, represented by a stochastic error term ε_{ij}, so that $u_{ij} = v_{ij} + \varepsilon_{ij}$. Different assumptions about the distribution of the unobserved stochastic error term yield different models (such as the logit and probit models). Given an assumed form for the error distribution, an assumed form for the function mapping $\mathbf{z}_j$ to u_{ij}, and data of observed choices in a set of choice situations, the parameters of the assumed forms can be estimated in order to best fit the observed data. While this model fitting can be performed on any observed data, conjoint analysis experimental design techniques offer efficient ways to generate data sets that are orthogonal, balanced, and that reduce bias while extracting maximum information from the minimum number of survey questions.

2.1.6 Marketing Product Design Models

The product design models in marketing (such as Kohli and Sukumar 1990; Nair, Thakur, and Wen, 1995; Krishnan, Singh, and Tirupati, 1999; and Shi, Olafsson and Chen, 2001) assume discrete levels for each product characteristic and choose the characteristic level combinations that maximize profit (or market share) using combinatorial search algorithms. Typically, utility is estimated as a function of the product characteristics using conjoint analysis surveys, and it is assumed that the product with highest utility, as predicted by the conjoint model, is deterministically chosen for each consumer – a "first-choice" model. Detailed design decisions $\mathbf{x}_j$ are ignored, or products are chosen such that $\mathbf{x}_j$ and $\mathbf{z}_j$ are the same. Many commodities may be modeled this way; for example, the price, container size, and container shape of ground coffee could be modeled both as decision variables and as product characteristics observable by the consumer. However, for products with more complex engineering content where consumers do not directly observe the internal engineering design variables, which yet affect the product characteristics observed by consumers, $\mathbf{x}_j$ and $\mathbf{z}_j$ may be quite distinct.

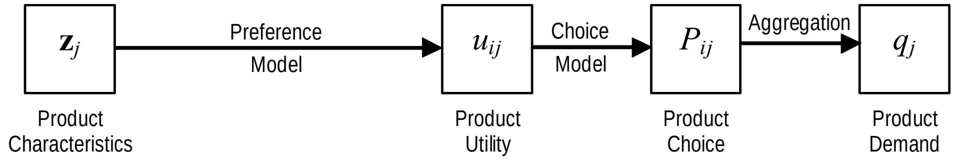


Figure 2.6: Marketing product design models

Strictly speaking, research in this group does not exclusively use objective, measurable aspects for $\mathbf{z}_j$, because in many marketing applications it may not be necessary. Instead words such as "small / medium / large" may be used to describe the levels of a characteristic. These words are perceptual in nature, yet intended to be representative of an objective quality of the product that can be used to make specific changes to the offered products. When coordinating with engineering, strict usage of objective and measurable product characteristics becomes more critical. A detailed discussion of marketing product design models relevant to design of a single product is provided in Chapter 4, and those applied to design of a product line are detailed in Chapter 5.

2.1.7 Product Positioning Models

Product positioning models typically describe a product as a point in a multiattribute space that is derived from existing products using perceptual mapping techniques such as multidimensional scaling (MDS) (for example, Albers, 1979, 1982; Sudharshan, May and Shocker, 1987; Gavish, Horsky and Srikanth, 1983; and Carpenter, 1989). Objective, physical product characteristics are not considered directly. Instead, subjective perceptual attributes such as 'sporty-ness' are used, and it is assumed that once the desired product position is found, the design of a product with the desired product attributes can be pursued separately. In multidimensional scaling (MDS), for example, consumers are asked to rate the degree of similarity between all pairs in a set of products. Then a small number of dimensions (typically two) is proposed to map out the perceptual space of products visually, and the products are placed into an attribute space such that products rated as being similar are close together geometrically. Then, proposed adjectives such as 'sporty-ness' are fit into the space and studied to gain intuition about alignment of products perceived as similar with these perceptual attributes. Typically, product utility is not explicitly mentioned, but ideal point models, which essentially define a utility function as the Euclidian distance of the product attributes from a user's fixed ideal attribute point in the attribute space, are commonly used to model preference, and customer heterogeneity is accounted for by allowing each customer a different ideal point. In practice, each consumer may have a different perceptual space; however, these models generally assume that the perceptual attribute space is constant across consumers, and consumers differ in their preference models, as expressed by different ideal points.

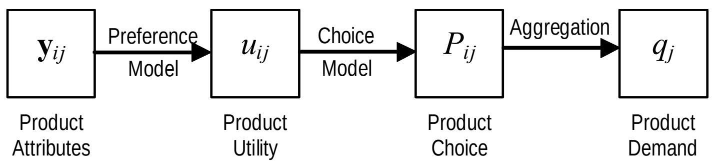


Figure 2.7: Product positioning models

To avoid confusion, the term "product attributes" is reserved in this dissertation for subjective, perceptual aspects of the product $\mathbf{y}_{ij}$, as opposed to objective, measurable product characteristics $\mathbf{z}_j$. Incorporation of perceptual models and attributes in the dissertation methodology is left for future work.

2.1.8 Product Line Selection

The product line selection models are explicitly focused on decisions that account for the competitive interaction between products (for example, Green and Krieger, 1985, 1987; McBride and Zufryden, 1988; Dobson and Kalish, 1988, 1993; and Chen and Hausman, 2000). These models treat the products at a very abstract level: A discrete set of product options is assumed given, and each option is modeled only in terms of its utility and cost (each determined externally). The product itself is not modeled. Customers are assumed to choose the product option with the highest utility. The selection problem is then modeled as a decision of which products from the set to include in the product line.

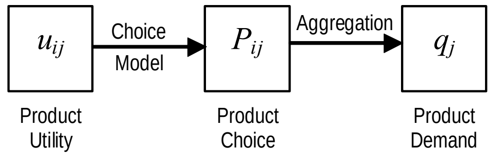


Figure 2.8: Product line selection models

These line selection models can be considered "downstream" models in that they discuss how to select appropriate products for a product line given a set of candidates, but they do not discuss how to generate the candidates. In this dissertation, the set of product candidates is presumed to be an infinite set, so these product line selection methods are not directly applicable. Further discussion of the product line selection models are provided in Chapter 5.

2.2 REFERENCES

[2.1] Albers, S. (1979) "An extended algorithm for optimal product positioning," *European Journal of Operational Research*, v3 p222-231.

[2.2] Albers, S. (1982) "PROPOPP: a program package for optimal positioning of a new product in an attribute space," *Journal of Marketing Research*, p606-608.

[2.3] Ben-Akiva, Moshe and Bruno Boccara (1995) "Discrete choice models with latent choice sets," *International Journal of Research in Marketing*, v12 p9-24.

[2.4] Berry, Steven (1994) "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, v25 n2 p242-262.

[2.5] Braun, R. (1996) *Collaborative Optimization: An Architecture for Large-Scale Distributed Design*, Ph.D. Dissertation, Stanford University.

[2.6] Carpenter, Gregory (1989) "Perceptual position and competitive brand strategy in a two-dimensional, two-brand market," *Management Science*, v35 n9 p1029-1044.

[2.7] Chintagunta, Pradeep (1994) "Heterogeneous logit model implications for brand positioning," *Journal of Marketing Research*, v31 p304-311.

[2.8] Cramer, E., J. Dennis, P. Frank, R. Lewis, and G. Shubin (1994) "Problem formulation for multidisciplinary optimization," *SIAM Journal of Optimization*, v4 p754-776.

[2.9] Gavish, B., Horsky, D., and Srikanth, K. (1983) "An approach to the optimal positioning of a new product," *Management Science*, v29 n11 p1277-1297.

[2.10] Gonul, Fusun and Kannan Srinivasan (1993) "Modeling multiple sources of heterogeneity in multinomial logit models: methodological and managerial issues," *Marketing Science*, v12 n3 p213-229.

[2.11] Green and Srinivasan (1978) "Conjoint analysis in consumer research: issues and outlook," *Journal of Consumer Research*, v5 p103-123.

[2.12] Green and Srinivasan (1990) "Conjoint analysis in marketing: new developments with implications for research and practice," *Journal of Marketing*, v54 n4 p3-19.

[2.13] Green, Paul, Jerry Wind, and Vithala Rao (1999) "Conjoint analysis: methods and applications," taken from *The Technology Management Handbook*, R. C. Dorf (editor), CRC Press: Boca Raton, FL p[12-65]-[12-72].

[2.14] Guadagni and Little (1983) "A logit model of brand choice calibrated on scanner data," *Marketing Science*, v2 n3 p203-238.

[2.15] Gupta, S. and A. Samuel (2001) "Integrating market research with the product development process: a step towards design for profit," *Proceedings of the 2001 ASME DETC*, DFM-21202.

[2.16] Hazelrigg, G.A. (1988) "A framework for decision-based engineering design," *ASME Journal of Mechanical Design*, v120 p653-658.

[2.17] Kim, H.M. (2001) *Target Cascading in Optimal System Design*, Ph.D. Dissertation, Dept. of Mechanical Engineering, University of Michigan, Ann Arbor, MI.

[2.18] Kohli, R., Sukumar, R. (1990) "Heuristics for product-line design using conjoint analysis," *Management Science*, v36 n12 p1464- 1478.

[2.19] Krishnan, V, R. Singh, and D. Tirupati (1999) "A model-based approach for planning and developing a family of technology-based products," *Manufacturing and Service Operations Management*, n1 v2 p132-156.

[2.20] Kuhfeld, W. (2003) *Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques*, accessed on 6-Nov-2004 <http://support.sas.com/techsup/technote/ts689.pdf>

[2.21] Li, H. and S. Azarm (2000) "Product design selection under uncertainty and with competitive advantage" *ASME Journal of Mechanical Design*, v122 p411-418.

[2.22] Manrai, Ajay (1995) "Mathematical models of brand choice behavior," *European Journal of Operations Research*, v82 p1-17.

[2.23] Marton, M. and F. Mistree (1998) "An implementation of expected utility theory in decision based design," *Proceedings of the 1998 ASME DETC*, DTM-5670, p1-11.

[2.24] Nair, S. K., L. S. Thakur, K. Wen, (1995), "Near optimal solutions for product line design and selection: beam search heuristics," *Management Science*, v41 n5 p767–785.

[2.25] Shi, L., S. Ólafsson, and Q. Chen (2001). "An optimization framework for product design," *Management Science*, v47 p1681-1692.

[2.26] Sobieszczanski-Sobieski, J., B.B. James, and A.R. Dovi (1985) "Structural optimization by multilevel decomposition," *AIAA Journal*, v21 p1291-1299.

[2.27] Sobieszczanski-Sobieski, J. (1989) "Optimization by decomposition: a step from hierarchic to non-hierarchic systems," *2nd NASA/Air Force Symposium on Recent Advances in Multidisciplinary Analysis and Optimization*, Hampton, VA, September 28-30, 1988. NASA TM-101494, NASA CP-3031.

[2.28] Sobieszczanski-Sobieski, J. and S. Kodiyalam (1999) "BLISS: A new method for two-level structural optimization," *AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, St. Louis, MO v2 p1274-1286.

[2.29] Sudharshan, D., Jerrold May and Allan Shocker (1987) "A simulation comparison of methods for new product location," *Marketing Science*, v6 n2 p182-201.

[2.30] Von Newmann, J. and O. Morgenstern (1944) *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey.

[2.31] Wassenaar and Chen (2003) "An approach to decision-based design," *ASME Journal of Mechanical Design*, v125 p490-497.

[2.32] Wu, J. and M. Hamada (2000) *Experiments: Planning, Analysis, and Parameter Design Optimization*, Wiley-Interscience.

CHAPTER 3

ANALYTICAL TARGET CASCADING THEORY AND EXTENSIONS

This chapter introduces the general analytical target cascading (ATC) methodology for optimizing complex systems through decomposition into a hierarchy of subsystems and coordination of the solutions to the subsystems in such as way as to obtain a joint system optimum. Convergence characteristics are discussed, and it is shown that the coordination process results in inconsistencies between subsystems in the ATC hierarchy whenever the top level system targets are unattainable. This is important for the applications presented in this dissertation because top level targets will correspond to maximizing profit, rather than matching a fixed, attainable target. A method is then proposed for finding weighting coefficients within the ATC hierarchy to obtain solutions with arbitrarily small inconsistencies so that the designer can define acceptable tolerances for each shared variable in the hierarchy, and the process will automatically find a solution within tolerance. The mathematical notation used in this chapter follows the general ATC notation and is specific to the chapter, deviating from the nomenclature used in the remainder of the dissertation. A summary of the nomenclature specific to this chapter is provided at the end of the chapter. The material in this chapter is based on publications by Michalek and Papalambros (2005a, 2005b).

3.1 INTRODUCTION

Analytical target cascading (ATC) is a model-based, hierarchical optimization methodology for systems design. ATC requires a set of analysis or simulation models that predict responses (the characteristics) of each system, subsystem, and component as a function of the design variables (the decisions) (Kim, 2003). The analysis models are organized using design optimization models that are the elements or building blocks of the hierarchy, as shown in Figure 3.1 with the standard index notation. The top level represents the overall system and each lower level represents a subsystem or component of its parent element. In the ATC process, top-level system design targets are propagated down to lower subsystem and component level targets that are then optimized to meet the targets as closely as possible. The resulting responses are rebalanced at higher levels by iteratively adjusting targets and designs to achieve consistency.

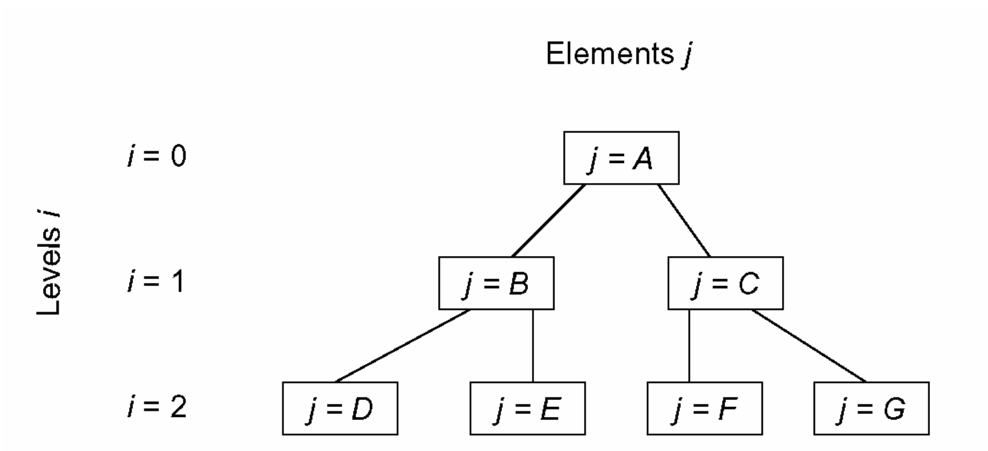


Figure 3.1: Example of index notation for a hierarchically partitioned problem

Following Michelena *et al*. (2003), and using the general notation introduced by Michalek and Papalambros (2005a), the *original design target problem* is:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \quad \left\| \mathbf{r}(\mathbf{x}) - \mathbf{T} \right\|_2^2 \\ &\text{subject to } \mathbf{g}(\mathbf{x}) \le \mathbf{0},\ \mathbf{h}(\mathbf{x}) = \mathbf{0}, \end{aligned} \tag{3.1}$$

where $\mathbf{T}$ is the vector of targets, $\mathbf{r}$ is the vector-valued response function, $\mathbf{x}$ is the complete vector of design variables, $\mathbf{g}$ and $\mathbf{h}$ are vectors of design constraint functions, and $\| \|_2^2$ denotes the square of the l_2 norm. Eq.(3.1) represents the entire large-scale system, and it is solved all-at-once (AAO); i.e., all variables and functions are evaluated together during search. Given that the system has an implied hierarchical structure of N+1 levels, as in Figure 3.1, the formulation (still solved AAO) can be equivalently represented by designating response variables and linking variables, creating copies of these variables at parent and child levels, and adding constraints forcing the copies to be equal:

$$
\begin{aligned}
&\underset{\overline{\mathbf{x}}_{ij},\mathbf{y}^i_{(i+1)j}}{\text{minimize}} \quad \left\|\mathbf{R}^0_{0l}-\mathbf{T}\right\|_2^2 \\
&\text{subject to} \sum_{k\in C_{ij}} \left\|\mathbf{R}^i_{(i+1)k}-\mathbf{R}^{i+1}_{(i+1)k}\right\|_2^2=0, \\
&\qquad\qquad \sum_{k\in C_{ij}} \left\|\mathbf{S}_k\mathbf{y}^i_{(i+1)j}-\mathbf{y}^{(i+1)}_{(i+1)k}\right\|_2^2=0, \\
&\qquad\qquad \mathbf{g}_{ij}\left(\overline{\mathbf{x}}_{ij}\right)\leq\mathbf{0},\ \mathbf{h}_{ij}\left(\overline{\mathbf{x}}_{ij}\right)=\mathbf{0}, \\
&\text{where} \qquad \mathbf{R}^i_{ij}=\mathbf{r}_{ij}\left(\overline{\mathbf{x}}_{ij}\right), \\
&\qquad\qquad \overline{\mathbf{x}}_{ij}=\left[\mathbf{x}^i_{ij},\mathbf{y}^i_{ij},\mathbf{R}^i_{(i+1)k_1},\ldots,\mathbf{R}^i_{(i+1)k_{c_{ij}}}\right]^T, \\
&\forall j\in E_i, i=0,1,\ldots,N,
\end{aligned}
\tag{3.2}
$$

where $\mathbf{x}^i_{ij}$ is the vector of local variables for element j at level i, $\mathbf{y}^i_{ij}$ is the vector of linking variables for element j at level i, $\mathbf{y}^{i-1}_{ij}$ is the copy of the vector of linking variables at element j level i coordinated by the parent element at level (i-1), $\mathbf{S}_j$ is the selection matrix indicating which terms of the parent coordinating linking variable vector $\mathbf{y}^{i-1}_{ij}$ are relevant to the linking variable vector $\mathbf{y}^i_{ij}$ at element j, $\mathbf{R}^i_{ij}$ is the vector of responses at element j level i, $\mathbf{R}^{i-1}_{ij}$ is the vector of response targets for element j at level i that are set by the parent element at level (i-1), $\mathbf{r}_{ij}$ is the vector-valued response function of element j at level i, $\mathbf{g}_{ij}$ is the vector of inequality constraints at element j level i, $\mathbf{h}_{ij}$ is the vector of equality constraints at element j level i, E_i is the set of elements at level i, C_{ij} is the set of element j's children numbered 1 through c_{ij}, and l designates the top level element, as in Michalek and Papalambros (2005a). Note that $\mathbf{y}^i{}_{ij}$ drops out for elements that do not have linking variables, such as element l, and $\mathbf{R}^i{}_{(i+1)k}$ terms drop out for leaf elements (elements that do not have children).

Following Michelena *et al*. (2003), the formulation in Eq.(3.2) is relaxed by allowing deviation between linking variable and response variable copies to be within a tolerance ε and minimizing ε. Additionally, vectors of weighting coefficients $\mathbf{w}$ are introduced for linking and response variables to specify the relative importance of matching each target at each level. This yields the *relaxed AAO formulation*, which is set up to be, but has not yet been, decomposed:

$$\begin{aligned}
&\underset{\overline{\mathbf{x}}_{ij},\mathbf{y}^i_{(i+1)j},\varepsilon^R_{ij},\varepsilon^y_{ij}}{\text{minimize}} && \left\|\mathbf{R}^0_{0l}-\mathbf{T}\right\|_2^2+\sum_{i=0}^{N-1}\sum_{j\in E_i}\varepsilon^R_{ij}+\sum_{i=0}^{N-1}\sum_{j\in E_i}\varepsilon^y_{ij} \\
&\text{subject to} && \sum_{k\in C_{ij}}\left\|\mathbf{w}^R_{(i+1)k}\circ\left(\mathbf{R}^i_{(i+1)k}-\mathbf{R}^{i+1}_{(i+1)k}\right)\right\|_2^2\le\varepsilon^R_{ij}, \\
& && \sum_{k\in C_{ij}}\left\|\mathbf{S}_k\mathbf{w}^y_{(i+1)j}\circ\left(\mathbf{S}_k\mathbf{y}^i_{(i+1)j}-\mathbf{y}^{(i+1)}_{(i+1)k}\right)\right\|_2^2\le\varepsilon^y_{ij}, \\
& && \mathbf{g}_{ij}\left(\overline{\mathbf{x}}_{ij}\right)\le\mathbf{0},\ \mathbf{h}_{ij}\left(\overline{\mathbf{x}}_{ij}\right)=\mathbf{0}, \\
&\text{where} && \mathbf{R}^i_{ij}=\mathbf{r}_{ij}\left(\overline{\mathbf{x}}_{ij}\right), \\
& && \overline{\mathbf{x}}_{ij}=\left[\mathbf{x}^i_{ij},\mathbf{y}^i_{ij},\mathbf{R}^i_{(i+1)k_1},\ldots,\mathbf{R}^i_{(i+1)k_{c_{ij}}}\right]^T, \\
&\forall j\in E_i, i=0,1,\ldots,N,
\end{aligned}\tag{3.3}$$

where ε^R_{ij} is the response deviation tolerance variable for element j level i, ε^y_{ij} is the linking deviation tolerance variable for element j level i, $\mathbf{w}^R_{ij}$ is the response deviation weighting coefficient vector for element j at level i, $\mathbf{w}^y_{ij}$ is the linking variable deviation weighting coefficient vector for element j at level i, and the ○ symbol is used to indicate term-by-term multiplication of vectors such that $[a_1\ a_2\ \ldots\ a_n]^T \circ [b_1\ b_2\ \ldots\ b_n]^T = [ab_1\ ab_2\ \ldots\ ab_n]^T$.

Finally, the problem is decomposed into separate elements P_{ij}, and monotonicity analysis (Papalambros and Wilde, 2000) is used to show that the ε-bound constraints of each element are active, allowing them to be solved for ε and moved into the objective function. The general notation for a *single ATC element* P_{ij} in the hierarchy then contains only the terms of Eq.(3.3) relevant to element P_{ij}, i.e., only terms that are not constant with respect to the decision variables of element P_{ij}. The notation for element P_{ij} is then:

$$\begin{aligned}
\underset{\bar{\mathbf{x}}_{ij},\mathbf{y}^i_{(i+1)j}}{\text{minimize}}\ & \left\|\mathbf{w}^R_{ij} \circ \left(\mathbf{R}^i_{ij} - \mathbf{R}^{i-1}_{ij}\right)\right\|_2^2 + \left\|\mathbf{S}_j\mathbf{w}^y_{ip} \circ \left(\mathbf{S}_j\mathbf{y}^{i-1}_{ip} - \mathbf{y}^i_{ij}\right)\right\|_2^2 \\
& + \sum_{k\in C_{ij}} \left\|\mathbf{w}^R_{(i+1)k} \circ \left(\mathbf{R}^i_{(i+1)k} - \mathbf{R}^{i+1}_{(i+1)k}\right)\right\|_2^2 + \sum_{k\in C_{ij}} \left\|\mathbf{S}_k\mathbf{w}^y_{(i+1)j} \circ \left(\mathbf{S}_k\mathbf{y}^i_{(i+1)j} - \mathbf{y}^{(i+1)}_{(i+1)k}\right)\right\|_2^2 \\
\text{subject to}\ & \mathbf{g}_{ij}\left(\bar{\mathbf{x}}_{ij}\right) \le \mathbf{0}, \quad \mathbf{h}_{ij}\left(\bar{\mathbf{x}}_{ij}\right) = \mathbf{0}, \\
\text{where}\ & \mathbf{R}^i_{ij} = \mathbf{r}_{ij}\left(\bar{\mathbf{x}}_{ij}\right), \\
& \bar{\mathbf{x}}_{ij} = \left[\mathbf{x}^i_{ij}, \mathbf{y}^i_{ij}, \mathbf{R}^i_{(i+1)k_1}, \ldots, \mathbf{R}^i_{(i+1)k_{c_{ij}}}\right]^T,
\end{aligned} \tag{3.4}$$

The sequence of solving each optimization problem element P_{ij} and passing its solution to the rest of the hierarchy is called a coordination strategy. Michelena *et al.* (2003) proved that using certain classes of coordination strategies to manage elements of the ATC formulation in Eq.(3.4), will result in convergence to the same solution as that of the relaxed AAO formulation in Eq.(3.3). Under these specific coordination strategies, managing the ATC hierarchy can be viewed as solving a series of Hierarchical Overlapping Coordination (HOC) problems, which have been shown to have non-ascent, global convergence properties (Shima and Haimes, 1984; Park *et al.,* 2001; Michelena *et al.,* 1999).

ATC has been applied to automotive applications (Kim *et al.,* 2002; Kim *et al.,* 2003; Kokkolaras *et al.,* 2004), including the design of product families (Kokkolaras *et al.,* 2002), as well as to the design of building systems (Choudhary *et al.,* 2003). Decomposing large-scale problems can be advantageous because it organizes and separates models and information by focus or discipline, provides communication only where necessary, and facilitates concurrent design. Moreover, ATC can solve some problems that are computationally difficult or impossible to solve all-at-once. Occasionally decomposition can also result in improved computational efficiency because the formulation of each element typically has fewer degrees of freedom and fewer constraints than the AAO formulation. However, computational efficiency of ATC is not yet well understood, and empirical evidence shows that it can vary dramatically depending on the choice of weighting coefficients (Tzevelekos *et al.,* 2003).

Several other systems have been proposed for multidisciplinary design optimization (MDO) of complex systems. In particular, Collaborative Optimization (CO) (Braun, 1996), based on concepts introduced by Sobieski (Sobieszczanski-Sobieski *et al.,* 1985), contains a similar form of minimizing deviations between targets and responses using the square of the l_2 norm. CO formulations so far have dealt only with bilevel problems, although multilevel extensions seem possible. Moreover, it has been observed by Alexandrov and Lewis (2000) and reemphasized by Kim (2001) that CO cannot, in general, produce KKT points because of constraint qualification failures, whereas ATC has proven convergence properties. ATC is different from MDO frameworks such as multidisciplinary feasible (MDF) and individual discipline feasible (IDF) (Cramer *et al.,* 1994), or the Bi-Level Integrated System Synthesis (BLISS) approach (Sobieszczanski-Sobieski and Kodiyalam, 1999), where analysis models at a single level are integrated under a master problem introduced as an authority to achieve the overall design goal. Furthermore, ATC should not be confused with strategies for nonhierarchical systems, such as Concurrent Subspace Optimization (CSSO) (Sobieszczanski-Sobieski, 1989), or formulation choices for design optimization statements at individual problem elements, such as simultaneous analysis and design (SAND) or nested analysis and design (NAND) (Balling and Sobieszczanski-Sobieski, 1994). In contrast, ATC represents a multilevel decision-making hierarchy for the design of systems consisting of an arbitrarily large hierarchy of levels of analysis and design models representing systems, subsystems, and components.

The global convergence theory of ATC (Michelena *et al.,* 2003) asserts that weighting coefficients can be found such that consistency deviation terms converge to zero. However, it will be shown that for problems with attainable targets, strictly consistent designs can be found with any positive finite weighting coefficients, but for problems with unattainable targets, strict design consistency cannot be achieved with finite weighting coefficients. Thus, the selection of proper weighting coefficients is necessary to achieve a solution within acceptable inconsistency tolerances. This result is particularly relevant when intentionally using "stretch targets" or "stretch goals*,"* terms used in management communities to describe setting very high, usually unattainable, goals in order to motivate employees (Lingle and Schiemann, 1999). In the applications of this dissertation, the top level goal is maximization of profit, and no attainable target is set.

In this chapter the issue of consistency for unattainable targets is discussed, and an iterative approach is proposed to find weighting coefficients that achieve solutions with user-specified inconsistency tolerances. The method is then generalized and demonstrated with several examples.

3.2 CONSISTENCY FOR UNATTAINABLE TARGETS

In the ATC global convergence proof Michelena *et al*. (2003) proved that when elements of the ATC hierarchy (Eq.(3.4)) are solved separately and iteratively using certain coordination strategies, the system will converge to the solution of the relaxed AAO formulation, Eq.(3.3). They go further to assert: “given that consistency and feasibility are assumed for the original design target problem, it is possible to find weights $w^R_{(i+1)k}$ and $w^y_{(i+1)k}$ such that $\varepsilon^R_{(i+1)k}$ and $\varepsilon^y_{(i+1)k}$ ... converge to zero. ... This implies that the ATC process, recursively applied to the problem hierarchy, produces an optimum solution of the original design target problem.”

The concepts of feasibility and consistency deserve further discussion here. Feasibility of the original design target problem means that a design exists that satisfies all constraints. Feasibility of the ATC elements means a local design exists at each ATC element P_{ij} that satisfies all of the constraints at that element. Consistency of the ATC formulation further supposes a solution exists such that $\mathbf{R}^i_{(i+1)k} = \mathbf{R}^{i+1}_{(i+1)k}$ and $\mathbf{y}^i_{(i+1)k} = \mathbf{y}^{i+1}_{(i+1)k}$ for all i, $j \in E_i$, $k \in C_{ij}$, which implies that $\varepsilon^R = 0$ and $\varepsilon^y = 0$ for all elements. Feasibility of the original design target formulation implies that a design exists in the decomposed ATC formulation that is feasible at all elements and consistent among elements.

In this section it is demonstrated that despite existence of a feasible, consistent design, the ATC formulation will not find this design with finite weighting coefficients unless the design meets the top level targets exactly. Specifically, if a feasible solution to the original problem exists that meets the top level targets exactly, then any choice of positive, finite weighting coefficients in the ATC formulation will yield a consistent solution. If such a solution does not exist, the ATC formulation will not yield a consistent solution for any finite weighting coefficients. However, an ATC solution can be found with arbitrarily small inconsistency deviations if weights are chosen appropriately.

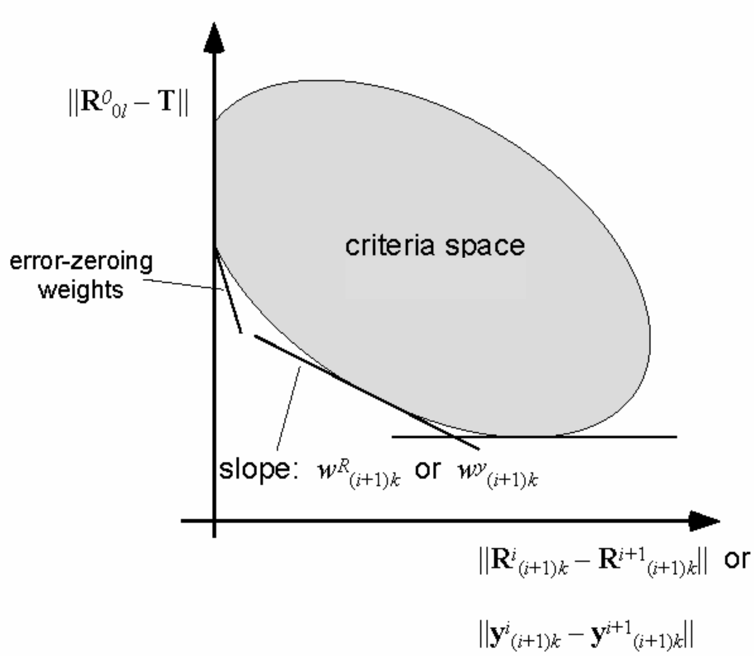


Figure 3.2: Existence of error-zeroing weights proposed by Michelena *et al.* (2003)

Michelena *et al.* (2003) proposed a Pareto optimization analogy to illustrate the existence of error-zeroing weights, as shown in Figure 3.2. They observed that Eq.(3.4) contains a weighted sum of deviation metric terms, and they visualized the solution as a Pareto set between terms in the objective function, showing how larger weighting coefficients for parent-child deviation terms yield points with lower consistency deviation between parent and child at the expense of minimizing deviation from the top level target. However, this figure could be misleading. Note that if a consistent, feasible design exists that meets the top level targets, then the design would map to the origin in Figure 3.2, and any other design would be either dominated by or equivalent to it in this space. Therefore, in this case the Pareto surface degenerates to a single point, the origin, which can be achieved with any positive weighting coefficients. If such a design does not exist, then it will be shown in Eq.(3.7), Eq.(3.13), and Eq.(3.16) that in general the consistency deviation approaches zero only as the weighting coefficients for consistency approach infinity. So, in this case the vertical axis is tangent to the Pareto surface, and there are no finite error-zeroing weights. This is important for applications where unattainable targets are used purposefully or when the designer is uncertain if targets can be achieved.

A simple example will demonstrate this situation. Let us examine an unconstrained level-0 element with a single level-1 child. The level-0 element is called l, and the level-1 element is called k. There are no linking variables and only a single top level target T is considered. Following Eq.(3.4), the level-0 problem (P_{0l}) is written as:

$$\underset{\left\{\mathbf{x}_{0l}^{0},\mathbf{R}_{1k}^{0}\right\}}{\text{minimize}} \; \left\|T - r_{0l}\left(\mathbf{x}_{0l}^{0},\mathbf{R}_{1k}^{0}\right)\right\|_2^2 + \left\|\mathbf{w}_{1k}^{R}\circ\left(\mathbf{R}_{1k}^{0}-\mathbf{R}_{1k}^{1}\right)\right\|_2^2. \tag{3.5}$$

Writing out the squared l_2 norm in terms of vector elements by using the angle bracket symbol < > to denote vector elements indexed with α, and dropping the functional dependency notation for r_{0l}, the objective function f_{0l} at level-0 is

$$f_{0l} = \left(T - r_{0l}\right)^2 + \sum_{\alpha}\left(\left\langle \mathbf{w}_{1k}^{R}\right\rangle_{\alpha}\left\langle \mathbf{R}_{1k}^{0}-\mathbf{R}_{1k}^{1}\right\rangle_{\alpha}\right)^2. \tag{3.6}$$

The first order necessary conditions for optimality of an unconstrained problem require that if a (local) solution to Eq.(3.6) exists, than the gradient of the objective function with respect to the response targets $\mathbf{R}_{1k}^{0}$ at that point must be zero:

$$
\begin{aligned}
\frac{\partial f_{0l}}{\partial \mathbf{R}_{1k}^0} &= 2\left(r_{0l}-T\right)\frac{\partial r_{0l}}{\partial \mathbf{R}_{1k}^0}+2\sum_{\alpha}\left\langle \mathbf{w}_{1k}^R\right\rangle_{\alpha}^2\left\langle \mathbf{R}_{1k}^0-\mathbf{R}_{1k}^1\right\rangle_{\alpha}\frac{\partial\left\langle \mathbf{R}_{1k}^0\right\rangle_{\alpha}}{\partial \mathbf{R}_{1k}^0} \\
&= 2\left(r_{0l}-T\right)\frac{\partial r_{0l}}{\partial \mathbf{R}_{1k}^0}+2\begin{pmatrix} \left\langle \mathbf{w}_{1k}^R\right\rangle_1^2\left\langle \mathbf{R}_{1k}^0-\mathbf{R}_{1k}^1\right\rangle_1\left[1,0,\ldots,0\right]^T \\ +\left\langle \mathbf{w}_{1k}^R\right\rangle_2^2\left\langle \mathbf{R}_{1k}^0-\mathbf{R}_{1k}^1\right\rangle_2\left[0,1,\ldots,0\right]^T \\ +\ldots \\ +\left\langle \mathbf{w}_{1k}^R\right\rangle_n^2\left\langle \mathbf{R}_{1k}^0-\mathbf{R}_{1k}^1\right\rangle_n\left[0,0,\ldots,1\right]^T \end{pmatrix} \\
&= 2\left(r_{0l}-T\right)\frac{\partial r_{0l}}{\partial \mathbf{R}_{1k}^0}+2\mathbf{w}_{1k}^R\circ\mathbf{w}_{1k}^R\circ\left(\mathbf{R}_{1k}^0-\mathbf{R}_{1k}^1\right)=0 \\
\therefore \left\langle \mathbf{R}_{1k}^0-\mathbf{R}_{1k}^1\right\rangle_{\alpha} &= \left(\frac{T-r_{0l}}{\left(\left\langle \mathbf{w}_{1k}^R\right\rangle_{\alpha}\right)^2}\right)\frac{\partial r_{0l}}{\partial\left\langle \mathbf{R}_{1k}^0\right\rangle_{\alpha}}.
\end{aligned}
\tag{3.7}
$$

This last equation shows that the optimal design will not be strictly consistent ($\mathbf{R}_{1k}^0 \neq \mathbf{R}_{1k}^1$) for positive, finite weights unless the top level target is met exactly or the derivative of the response function with respect to $\mathbf{R}_{1k}^0$ happens to be zero at the optimum. If top level targets are unattainable ($T - r_{0l}$) ≠ 0, then the inconsistency deviation error ($\mathbf{R}_{1k}^0 - \mathbf{R}_{1k}^1$) will be nonzero, except in the special case where the derivative of the response function is zero at the optimum, which can happen mostly by coincidence. Thus, in general ($\mathbf{R}_{1k}^0 - \mathbf{R}_{1k}^1$) approaches zero only as the terms of $\mathbf{w}_{1k}^R$ approach infinity.

At this point one is tempted to simply set large weights. However, apart from the ATC convergence requirement, the size of the weights will also have a scaling effect on the nonlinear programming algorithm used to solve the element problem. Adverse scaling will increase computational time or altogether prevent solution of the element problem. Additionally as will be shown later, in multilevel hierarchies the resulting deviations at any particular element depend on ratios of the weights at that element to weights at the parent element, and there are interactions between weights for linking variables $\mathbf{w}^y$ and for response variables $\mathbf{w}^R$. So, simply setting all weighting coefficients to large values will not necessarily result in small inconsistency deviation values. The task then is to find appropriate weights such that the resulting inconsistency deviation is acceptable. One way to approach this task is to use the results of Eq.(3.7) to calculate estimates of the weighting terms $\mathbf{w}_{1k}^R$ required to achieve acceptable consistency errors $\boldsymbol{\theta}_{1k}^R$ for each of the response targets $\mathbf{R}_{1k}^0$. To do this, we set the left hand side of the equation to the desired inconsistency $\boldsymbol{\theta}_{1k}^R$ and solve for the weights:

$$\left\langle \mathbf{w}_{1k}^R \right\rangle_\alpha = \left| \left(\frac{r_{0l} - T}{\left\langle \boldsymbol{\theta}_{1k}^R \right\rangle_\alpha} \right) \frac{\partial r_{0l}}{\partial \left\langle \mathbf{R}_{1k}^0 \right\rangle_\alpha} \right|^{\frac{1}{2}} \tag{3.8}$$

Thus, in this example the weighting update method for finding appropriate weights to achieve consistency error tolerances $\boldsymbol{\theta}_{1k}^R$ would follow these steps:

1. Set initial-guess weights (say, $\mathbf{w}_{1k}^R$ = [1, 1, ..., 1]^T).
2. Solve the ATC problem and calculate the top-level target deviation and the derivative of the response function at the solution.
3. If the deviation tolerance is not satisfied at the solution, use Eq.(3.8) to find new weighting terms, and return to step 2.

3.3 GENERALIZATION OF THE WEIGHTING UPDATE METHOD

The goal of the weighting update method is to automatically identify appropriate weighting coefficients that achieve designs with acceptable deviation tolerance values for the response variables at each element $\boldsymbol{\theta}_{ij}^R$ and for the linking variables at each parent coordinating element $\boldsymbol{\theta}_{(i+1)j}^y$. The problem is first solved using starting values for weighting coefficients. The solution to that problem is used to calculate a linear approximation of the weighting coefficients needed to achieve the desired tolerances.

Weights are updated with this approximation, and the problem is solved again. This process is repeated until the inconsistency deviation tolerance is achieved, namely, the final solution satisfies the conditions

$$
\begin{aligned}
&\left|\mathbf{R}_{ij}^{i-1} - \mathbf{R}_{ij}^{i}\right| \le \boldsymbol{\theta}_{ij}^{R}, \\
&\left|\mathbf{y}_{(i+1)k}^{(i+1)} - \mathbf{y}_{(i+1)k'}^{(i+1)}\right| \le \boldsymbol{\theta}_{(i+1)j}^{y}, \\
&\quad \forall k, k' \in C_{ij},\ j \in E_i,\ i = 0,1,\ldots,N.
\end{aligned}
\tag{3.9}
$$

To generalize the method presented in the previous section, one of the KKT first order necessary conditions for optimality of constrained nonlinear problems, which involves the Lagrangian, is examined. From Eq.(3.4), the Lagrangian L_{ij} of element j at level i is

$$
\begin{aligned}
L_{ij} = {} & \left\| \mathbf{w}_{ij}^{R} \circ \left(\mathbf{r}_{ij} - \mathbf{R}_{ij}^{i-1} \right) \right\|_2^2 + \left\| \mathbf{S}_j \mathbf{w}_{ip}^{y} \circ \left(\mathbf{S}_j \mathbf{y}_{ip}^{i-1} - \mathbf{y}_{ij}^{i} \right) \right\|_2^2 \\
& + \sum_{k \in C_{ij}} \left\| \mathbf{w}_{(i+1)k}^{R} \circ \left(\mathbf{R}_{(i+1)k}^{i} - \mathbf{R}_{(i+1)k}^{i+1} \right) \right\|_2^2 \\
& + \sum_{k \in C_{ij}} \left\| \mathbf{S}_k \mathbf{w}_{(i+1)j}^{y} \circ \left(\mathbf{S}_k \mathbf{y}_{(i+1)j}^{i} - \mathbf{y}_{(i+1)k}^{(i+1)} \right) \right\|_2^2 + \boldsymbol{\mu}_{ij}^{T} \mathbf{g}_{ij} + \boldsymbol{\lambda}_{ij}^{T} \mathbf{h}_{ij}
\end{aligned}
\tag{3.10}
$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are the vectors of Lagrange multipliers for the inequality and equality constraints respectively. Expressing the norms using vector terms indexed with the symbol α, we have

$$
\begin{aligned}
L_{ij} = &\sum_{\alpha_1}\left(\left\langle \mathbf{w}_{ij}^{R}\right\rangle_{\alpha_1}\left\langle \mathbf{r}_{ij}-\mathbf{R}_{ij}^{i-1}\right\rangle_{\alpha_1}\right)^2+\sum_{\alpha_2}\left(\left\langle \mathbf{S}_j\mathbf{w}_{ij}^{y}\right\rangle_{\alpha_2}\left\langle \mathbf{S}_j\mathbf{y}_{ip}^{i-1}-\mathbf{y}_{ij}^{i}\right\rangle_{\alpha_2}\right)^2\\
&+\sum_{k\in C_{ij}}\sum_{\alpha_3}\left(\left\langle \mathbf{w}_{(i+1)k}^{R}\right\rangle_{\alpha_3}\left\langle \mathbf{R}_{(i+1)k}^{i}-\mathbf{R}_{(i+1)k}^{i+1}\right\rangle_{\alpha_3}\right)^2\\
&+\sum_{k\in C_{ij}}\sum_{\alpha_4}\left(\left\langle \mathbf{S}_k\mathbf{w}_{(i+1)j}^{y}\right\rangle_{\alpha_4}\left\langle \mathbf{S}_k\mathbf{y}_{(i+1)j}^{i}-\mathbf{y}_{(i+1)k}^{(i+1)}\right\rangle_{\alpha_4}\right)^2+\boldsymbol{\mu}_{ij}^T\mathbf{g}_{ij}+\boldsymbol{\lambda}_{ij}^T\mathbf{h}_{ij}
\end{aligned}
\tag{3.11}
$$

If a feasible solution to Eq.(3.4) exists, one property of the KKT first order necessary conditions states that at the (local) solution the gradient of the Lagrangian with respect to each term β of the response target vector $\mathbf{R}^i_{(i+1)\gamma}$ for element γ is zero.

$$
\begin{aligned}
\frac{\partial L_{ij}}{\partial\left\langle \mathbf{R}_{(i+1)\gamma}^{i}\right\rangle_\beta} = &\, 2\sum_{\alpha_1}\left[\left(\left\langle \mathbf{w}_{ij}^{R}\right\rangle_{\alpha_1}\right)^2\left\langle \mathbf{r}_{ij}-\mathbf{R}_{ij}^{i-1}\right\rangle_{\alpha_1}\frac{\partial\left\langle \mathbf{r}_{ij}\right\rangle_{\alpha_1}}{\partial\left\langle \mathbf{R}_{(i+1)\gamma}^{i}\right\rangle_\beta}\right]\\
&+2\left(\left\langle \mathbf{w}_{(i+1)\gamma}^{R}\right\rangle_\beta\right)^2\left\langle \mathbf{R}_{(i+1)\gamma}^{i}-\mathbf{R}_{(i+1)\gamma}^{i+1}\right\rangle_\beta\\
&+\boldsymbol{\mu}_{ij}^T\frac{\partial \mathbf{g}_{ij}}{\partial\left\langle \mathbf{R}_{(i+1)\gamma}^{i}\right\rangle_\beta}+\boldsymbol{\lambda}_{ij}^T\frac{\partial \mathbf{h}_{ij}}{\partial\left\langle \mathbf{R}_{(i+1)\gamma}^{i}\right\rangle_\beta}=\mathbf{0}
\end{aligned}
\tag{3.12}
$$

Therefore, at the solution the deviation between response variable copies at parent and child level is

$$
\begin{aligned}
\left\langle \mathbf{R}_{(i+1)\gamma}^{i}-\mathbf{R}_{(i+1)\gamma}^{i+1}\right\rangle_\beta = &\,\frac{1}{\left(\left\langle \mathbf{w}_{(i+1)\gamma}^{R}\right\rangle_\beta\right)^2}\sum_{\alpha_1}\left(\left(\left\langle \mathbf{w}_{ij}^{R}\right\rangle_{\alpha_1}\right)^2\left\langle \mathbf{R}_{ij}^{i-1}-\mathbf{r}_{ij}\right\rangle_{\alpha_1}\frac{\partial\left\langle \mathbf{r}_{ij}\right\rangle_{\alpha_1}}{\partial\left\langle \mathbf{R}_{(i+1)\gamma}^{i}\right\rangle_\beta}\right)\\
&-\frac{1}{2\left(\left\langle \mathbf{w}_{(i+1)\gamma}^{R}\right\rangle_\beta\right)^2}\left(\boldsymbol{\mu}_{ij}^T\frac{\partial \mathbf{g}_{ij}}{\partial\left\langle \mathbf{R}_{(i+1)\gamma}^{i}\right\rangle_\beta}+\boldsymbol{\lambda}_{ij}^T\frac{\partial \mathbf{h}_{ij}}{\partial\left\langle \mathbf{R}_{(i+1)\gamma}^{i}\right\rangle_\beta}\right)
\end{aligned}
\tag{3.13}
$$

Note that this equation holds for all elements except the top level element. To achieve desired response variable deviation tolerances within $\boldsymbol{\theta}^{R}_{(i+1)\gamma}$ for each element in $\mathbf{R}^{i}_{(i+1)\gamma}$, each weighting term β in $\mathbf{w}^{R}_{(i+1)\gamma}$ should be updated as

$$\left\langle \mathbf{w}^{R}_{(i+1)\gamma} \right\rangle_{\beta} = \left| \frac{\Psi_{\gamma\beta}}{\left\langle \boldsymbol{\theta}^{R}_{(i+1)\gamma} \right\rangle_{\beta}} \right|^{\frac{1}{2}}$$

$$\text{where } \Psi_{\gamma\beta} = \sum_{\alpha_1} \left(\left(\left\langle \mathbf{w}^{R}_{ij} \right\rangle_{\alpha_1} \right)^2 \left\langle \mathbf{R}^{i-1}_{ij} - \mathbf{r}_{ij} \right\rangle_{\alpha_1} \frac{\partial \left\langle \mathbf{r}_{ij} \right\rangle_{\alpha_1}}{\partial \left\langle \mathbf{R}^{i}_{(i+1)\gamma} \right\rangle_{\beta}} \right) - \frac{1}{2} \left(\boldsymbol{\mu}^{T}_{ij} \frac{\partial \mathbf{g}_{ij}}{\partial \left\langle \mathbf{R}^{i}_{(i+1)\gamma} \right\rangle_{\beta}} + \boldsymbol{\lambda}^{T}_{ij} \frac{\partial \mathbf{h}_{ij}}{\partial \left\langle \mathbf{R}^{i}_{(i+1)\gamma} \right\rangle_{\beta}} \right) \tag{3.14}$$

Note again that this equation holds for all levels except the top level, where the weighting coefficient vector is not updated. Top level response deviations reflect failure of the design to meet the top level targets, rather than inconsistencies in the design, and the top level weighting coefficient vector is set by the modeler to express the relative importance of matching each top level target; it is not updated. While all weighting coefficient vectors reflect the relative importance of matching variable copies, the lower-level vectors are updated such that the final preference reflects that which is needed to achieve user-defined inconsistency tolerances.

Additionally, at the solution the gradient of the Lagrangian with respect to the linking variables of element j is zero.

$$
\begin{aligned}
\frac{\partial L_{ij}}{\partial \left\langle \mathbf{y}_{ij}^{i} \right\rangle_{\beta}} &= 2\sum_{\alpha_1} \left(\left\langle \mathbf{w}_{ij}^{R} \right\rangle_{\alpha_1} \right)^2 \left\langle \mathbf{r}_{ij} - \mathbf{R}_{ij}^{i-1} \right\rangle_{\alpha_1} \frac{\partial \left\langle \mathbf{r}_{ij} \right\rangle_{\alpha_1}}{\partial \left\langle \mathbf{y}_{ij}^{i} \right\rangle_{\beta}} \\
&\quad + 2\left(\left\langle \mathbf{S}_j \mathbf{w}_{ip}^{y} \right\rangle_{\beta} \right)^2 \left\langle \mathbf{y}_{ij}^{i} - \mathbf{S}_j \mathbf{y}_{ip}^{i-1} \right\rangle_{\beta} \\
&\quad + \boldsymbol{\mu}_{ij}^{T} \frac{\partial \mathbf{g}_{ij}}{\partial \left\langle \mathbf{y}_{ij}^{i} \right\rangle_{\beta}} + \boldsymbol{\lambda}_{ij}^{T} \frac{\partial \mathbf{h}_{ij}}{\partial \left\langle \mathbf{y}_{ij}^{i} \right\rangle_{\beta}} = \mathbf{0}
\end{aligned}
\tag{3.15}
$$

Therefore, the deviation between linking variable term β in $\mathbf{y}_{ij}^{i}$ and the parent coordination copy in $\mathbf{y}_{ip}^{i-1}$ is

$$
\begin{aligned}
\left\langle \mathbf{y}_{ij}^{i} - \mathbf{S}_j \mathbf{y}_{ip}^{i-1} \right\rangle_{\beta} &= \frac{1}{\left(\left\langle \mathbf{S}_j \mathbf{w}_{ip}^{y} \right\rangle_{\beta} \right)^2} \sum_{\alpha_1} \left(\left(\left\langle \mathbf{w}_{ij}^{R} \right\rangle_{\alpha_1} \right)^2 \left\langle \mathbf{R}_{ij}^{i-1} - \mathbf{r}_{ij} \right\rangle_{\alpha_1} \frac{\partial \left\langle \mathbf{r}_{ij} \right\rangle_{\alpha_1}}{\partial \left\langle \mathbf{y}_{ij}^{i} \right\rangle_{\beta}} \right) \\
&\quad - \frac{1}{2\left(\left\langle \mathbf{S}_j \mathbf{w}_{ip}^{y} \right\rangle_{\beta} \right)^2} \left(\boldsymbol{\mu}_{ij}^{T} \frac{\partial \mathbf{g}_{ij}}{\partial \left\langle \mathbf{y}_{ij}^{i} \right\rangle_{\beta}} + \boldsymbol{\lambda}_{ij}^{T} \frac{\partial \mathbf{h}_{ij}}{\partial \left\langle \mathbf{y}_{ij}^{i} \right\rangle_{\beta}} \right)
\end{aligned}
\tag{3.16}
$$

This term represents deviation between linking variable copies at element j and the parent coordination copy. Recall that linking variables are shared by elements at the same level and coordinated at the parent level. To achieve a desired deviation tolerance between elements at the same level, the weight for each term β must be set high enough so that the difference between copies at any two child elements is less than or equal to the tolerance. The updating calculation for the linking variable weighting coefficients is then

$$\left\langle \mathbf{w}_{ip}^{y} \right\rangle_{\beta} = \underset{\forall j, j' \in C_{(i-1)p}^{\beta}}{\text{maximum}} \left| \frac{\Psi_{j\beta} - \Psi_{j'\beta}}{\left\langle \boldsymbol{\theta}_{ip}^{y} \right\rangle_{\beta}} \right|^{\frac{1}{2}}$$

$$\text{where } \Psi_{j\beta} = \sum_{\alpha_1} \left(\left(\left\langle \mathbf{w}_{ij}^{R} \right\rangle_{\alpha_1} \right)^2 \left\langle \mathbf{R}_{ij}^{i-1} - \mathbf{r}_{ij} \right\rangle_{\alpha_1} \frac{\partial \left\langle \mathbf{r}_{ij} \right\rangle_{\alpha_1}}{\partial \left\langle \mathbf{S}_j^T \mathbf{y}_{ij}^i \right\rangle_{\beta}} \right) - \frac{1}{2} \left(\boldsymbol{\mu}_{ij}^T \frac{\partial \mathbf{g}_{ij}}{\partial \left\langle \mathbf{S}_j^T \mathbf{y}_{ij}^i \right\rangle_{\beta}} + \boldsymbol{\lambda}_{ij}^T \frac{\partial \mathbf{h}_{ij}}{\partial \left\langle \mathbf{S}_j^T \mathbf{y}_{ij}^i \right\rangle_{\beta}} \right) \tag{3.17}$$

and where $C^{\beta}_{(i-1)p}$ is the set of children of parent element p that contain linking variable β (i.e., Ψ drops out for children where $<\mathbf{S}_j^T \mathbf{y}^i_{ij}>_{\beta} = 0$).

In summary, the generalized weighting update method involves iteratively solving the ATC formulation and updating the weighting coefficient vectors of each element (which express relative preferences for meeting each target) to achieve a solution with user-specified inconsistency deviation tolerances for each response variable $\boldsymbol{\theta}^R$ and each linking variable $\boldsymbol{\theta}^y$. The method is implemented with the following steps:

4. Set an acceptable inconsistency deviation tolerance for each response variable and each linking variable, and set initial weights (for example, set all weights to 1).
5. Solve the ATC problem.
6. If the inconsistency deviation tolerance is not satisfied at the solution, update each term in each weighting coefficient vector using Eq.(3.14) and Eq.(3.17), and return to step 2.

3.4 DEMONSTRATION

To illustrate the topic of strict consistency for unattainable targets a simple example is used where the target (zero in this case) is unattainable:

$$\begin{aligned} &\underset{z_1}{\text{minimize}} \quad \|z_1\|_2^2 \\ &\text{subject to } z_1 \geq 1 \end{aligned} \tag{3.18}$$

The solution to this problem is $z_1 = 1$. In the relaxed formulation of this problem, copies of z_1 are made at level-0 element l and at level-1 element k, using the R notation to designate responses (there are no local variables or linking variables), and the weighted deviation between the copies is constrained less than or equal to ε. The positive, finite weight w is used as the weighting term. The relaxed AAO problem (before decomposition) is then:

$$\begin{aligned} &\underset{R_{1k}^0, R_{1k}^1, \varepsilon}{\text{minimize}} \quad \left\|R_{0k}^0\right\|_2^2 + \varepsilon \\ &\text{subject to } g_1 = \left\|w\left(R_{1k}^0 - R_{1k}^1\right)\right\|_2^2 - \varepsilon \leq 0, \\ &\qquad\qquad\quad g_2 = 1 - R_{1k}^1 \leq 0, \\ &\text{where} \qquad R_{0l}^0 = r_{0l}\left(R_{1k}^0\right) = R_{1k}^0. \end{aligned} \tag{3.19}$$

Note that the relaxed AAO problem is used in the remainder of this example, and the problem is not decomposed for ATC, since Michelena *et al*. (2003) showed that these formulations yield equivalent solutions. At a KKT point, the gradient of the Lagrangian is zero.

$$\nabla f + \mu_1 \nabla g_1 + \mu_2 \nabla g_2 = \mathbf{0}$$

$$\begin{bmatrix} 2R_{1k}^0 \\ 0 \\ 1 \end{bmatrix} + \mu_1 \begin{bmatrix} 2w^2\left(R_{1k}^0 - R_{1k}^1\right) \\ 2w^2\left(R_{1k}^1 - R_{1k}^0\right) \\ -1 \end{bmatrix} + \mu_2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\therefore \mu_1 = 1, \; (g_1 \text{ is active}) \quad \therefore R_{1k}^0 = \frac{w^2}{1+w^2} R_{1k}^1 \tag{3.20}$$

$$\therefore \mu_2 = \left(\frac{2w^2}{1+w^2}\right) R_{1k}^1 \; (g_2 \text{ is active})$$

$$\therefore R_{1k}^1 = 1, \quad R_{1k}^0 = \frac{w^2}{1+w^2}, \quad R_{1k}^1 - R_{1k}^0 = \frac{1}{1+w^2} \neq 0$$

$$\therefore \varepsilon = \left(\frac{w}{1+w^2}\right)^2 \neq 0$$

This shows that ε is nonzero at the KKT point for any finite weight; however, w can be found to achieve ε arbitrarily close to zero. It is important to note that ε approaches zero as w approaches infinity *or* zero, and the goal is to ensure that the inconsistencies between the responses at each level are within an acceptable tolerance, rather than focusing on the value of ε. The inconsistency ($R_{1k}^1 - R_{1k}^0$) approaches zero only as w approaches infinity.

In addition, to demonstrate the need to avoid setting arbitrarily large weights, this problem was implemented in Matlab© 6.5.0 (MathWorks, 2004) using the *fmincon* function with the feasible starting point $[R_1^0 \; R_1^1 \; \varepsilon]^T$ = $[2 \; 5 \; 5]^T$ and the following parameters: TolCon = TolFun = TolX = 10^{-10}. The algorithm and parameters are specified here because the algorithm behavior depends on the parameters and starting point; however, this example serves to show the basic trends. Figure 3.3 shows the number of function evaluations needed to converge to a solution for each value of w. The figure shows an upward trend, emphasizing the need to avoid large weighting terms when possible.

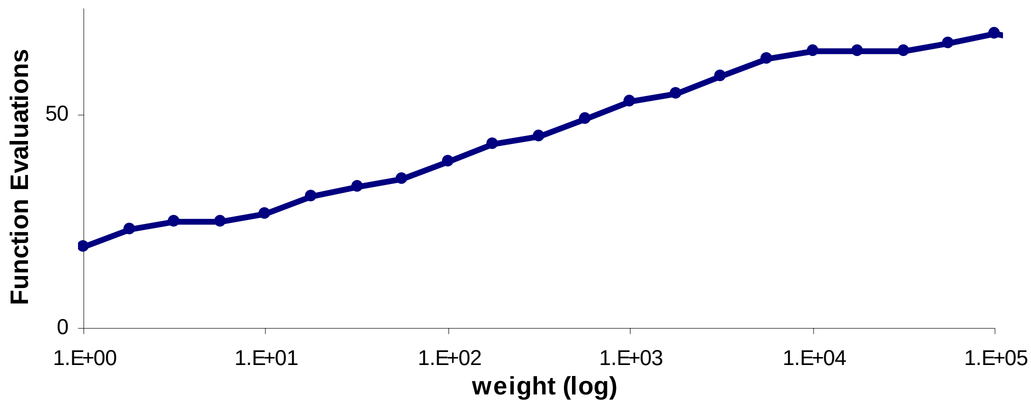


Figure 3.3: Number of function evaluations required to find the solution as a function of the weighting term

Figure 3.4 shows the resulting inconsistency deviation ($R_1^1 - R_1^0$) at the optimum for each value of *w*. The graph shows a trend of reduced error as the weighting term is increased, although the error never reaches zero.

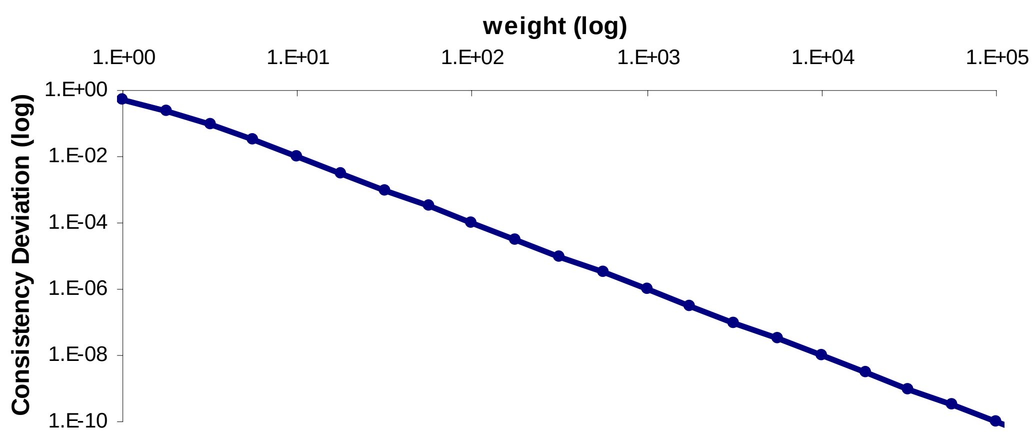


Figure 3.4: Inconsistency at the optimum as a function of the weighting term

In general, it is difficult to set appropriate weights simply by guessing. The weighting update method is applied to this example to show how appropriate weights are found. In this example, the response function r_{0l} is a linear function of R_1^0, so the derivative of the response function is a constant (=1), therefore, the use of the weighting update method to find appropriate weights yields,

$$w = \left| \left(\frac{r_0 - T}{\theta} \right) \frac{\partial r_0}{\partial R_1^0} \right|^{\frac{1}{2}} = \left| \frac{R_1^0}{\theta} \right|^{\frac{1}{2}} \tag{3.21}$$

The update procedure was implemented for this example with an inconsistency tolerance goal of $\theta = 10^{-2}$, and a starting weight of $w = 1$. The proper weight needed to achieve this inconsistency tolerance, $w = 9.95$, was found after three weighting update iterations and a total of 89 function evaluations.

3.5 GEOMETRIC PROGRAMMING EXAMPLE

The geometric programming example, proposed by Kim (2003), is used here as a multi-level example with linking variables to demonstrate the weighting update method. The original design target problem is

$$\begin{aligned}
&\underset{z_1, z_2, \ldots, z_{14}}{\text{minimize}}\ f = z_1^2 + z_2^2 \\
&\text{subject to } g_1 = z_3^{-2} + z_4^2 - z_5^2 \le 0 \\
&\qquad g_2 = z_5^2 + z_6^{-2} - z_7^2 \le 0 \\
&\qquad g_3 = z_8^2 + z_9^2 - z_{11}^2 \le 0 \\
&\qquad g_4 = z_8^{-2} + z_{10}^2 - z_{11}^2 \le 0 \\
&\qquad g_5 = z_{11}^2 + z_{12}^{-2} - z_{13}^2 \le 0 \\
&\qquad g_6 = z_{11}^2 + z_{12}^2 - z_{14}^2 \le 0 \\
&\qquad h_1 = z_1^2 - z_3^2 - z_4^{-2} - z_5^2 = 0 \\
&\qquad h_2 = z_2^2 - z_5^2 - z_6^2 - z_7^2 = 0 \\
&\qquad h_3 = z_3^2 - z_8^2 - z_9^{-2} - z_{10}^{-2} + z_{11}^2 = 0 \\
&\qquad h_4 = z_6^2 - z_{11}^2 - z_{12}^2 - z_{13}^2 + z_{14}^2 = 0 \\
&\qquad z_1, z_2, \ldots, z_{14} \ge 0
\end{aligned} \tag{3.22}$$

The original problem will be decomposed first as a two-level ATC hierarchy with three elements, as proposed by Kim (2003), and secondly as a three-level ATC hierarchy with five elements, as proposed by Tzevelekos *et al*. (2003). The feasible starting point $\mathbf{z} = [5, 5, 2.76, 0.25, 1.26, 4.64, 1.39, 0.67, 0.76, 1.70, 2.26, 1.41, 2.71, 2.66]^T$ is used for all trials, and the acceptable inconsistency tolerance value of 10^{-2} is used for all response variables and linking variables.

3.5.2 Two-Level Decomposition

In the two-level decomposition, following Kim (2003), the problem is partitioned into three elements: one level-0 element, A, with two level-1 children, B and C. The equality constraints of the original problem h_1, h_2, h_3, and h_4 are solved for z_1, z_2, z_3, and z_6 respectively and used as response functions of elements *A, A, B,* and C respectively. The objective function of each element is then to minimize deviation between targets and responses at that element, as in Eq.(3.4), where the top level targets are both zero. The variable z_{11} is treated as a linking variable between elements B and C, variables z_4, z_5, and z_7, are local variables of element A, variables z_8, z_9, and z_{10} are local variables of element B, and variables z_{12}, z_{13}, and z_{14} are local variables of element C. The constraints g_1, g_2, g_3, g_4, g_5, and g_6 are associated with elements A, A, B, B, C, and C respectively. Kim (2003) provides a picture of this decomposition for reference.

The problem was first solved with default weights $\mathbf{w}_{1B}^{R} = \mathbf{w}_{1C}^{R} = \mathbf{w}_{1A}^{y} = [1]$. At the solution, resulting inconsistency deviations are 0.688, 0.649, and 0.961 for z_3, z_6, and z_{11} respectively, all of which are larger than the acceptable tolerance value of 10^{-2}. Using the weighting update method, the weights are updated with Eq.(3.14) and Eq.(3.17), and the new problem is solved. This process of updating and solving is repeated four times before converging. The final weights, $\mathbf{w}_{1B}^{R}$ = 14.534, $\mathbf{w}_{1C}^{R}$ = 16.561, and $\mathbf{w}_{1A}^{y}$ = 27.572, yield inconsistencies of 10^{-2} for z_3, z_6, and z_{11}. The weighting update method successfully found the weighting coefficients that yield a solution with the desired inconsistency tolerance. These results are summarized in Table 3.1.

3.5.3 Three-Level Decomposition

In the three-level decomposition, following Tzevelekos *et al*. (2003), the problem is partitioned into five elements: one level-0 element A with two level-1 children, B and C, and two level-2 elements, D and E, which are children of B and C respectively. In the formulation z_5 is treated as a linking variable between elements B and C, z_{11} is set as a parameter with known (optimal) value 1.30, the equality constraints of the original problem h_1, h_2, h_3, and h_4 are used to calculate z_1, z_2, z_3, and z_4 as response functions of elements B, C, D, and E respectively, and the response function of element A is $f = (z_1^2 + z_2^2)$, with the top level target set to zero. The variable z_4 is a local variable of element B, variable z_7 is a local variable of element C, variables z_8, z_9, and z_{10} are local variables of element D, and variables z_{12}, z_{13}, and z_{14} are local variables of element E. The constraints g_1, g_2, g_3, g_4, g_5, and g_6 are associated with elements B, C, D, D, E and E respectively.

The problem was first solved with default weights $\mathbf{w}_{1A}^{y} = \mathbf{w}_{1B}^{R} = \mathbf{w}_{1C}^{R} = \mathbf{w}_{2D}^{R} = \mathbf{w}_{2E}^{R} = [1]$. At the solution, resulting inconsistencies are 1.47, 1.26, 0.78, 0.80, and 1.05 for z_1, z_2, z_3, z_5, and z_6 respectively, all of which are larger than the acceptable tolerance value of 10^{-2}. Using the weighting update method, the weights are updated with Eq.(3.14) and Eq.(3.17), and the new problem is solved. This process of updating and solving is repeated five times before converging. The final weights, $\mathbf{w}_{1A}^{y}$ = 109.70, $\mathbf{w}_{1B}^{R}$ = 99.34, $\mathbf{w}_{1C}^{R}$ = 103.59, $\mathbf{w}_{2D}^{R}$ = 85.96, and $\mathbf{w}_{2E}^{R}$ = 98.05, yield inconsistencies of 10^{-2} for z_1, z_2, z_3, z_5, and z_6. The weighting update method successfully found weights that yield a solution with the desired inconsistency tolerance. These results are summarized in Table 3.1.

Table 3.1: Results of the two-level and three-level geometric programming examples

		Two-Level		Three-Level	
		Default weights	Weighting update method	Default weights	Weighting update method
Weighting Coefficients (at soln.)	$\mathbf{w}^{y}_{1A}$	1	27.72	1	109.70
	$\mathbf{w}^{R}_{1B}$	1	14.61	1	99.34
	$\mathbf{w}^{R}_{1C}$	1	16.64	1	103.59
	$\mathbf{w}^{R}_{2D}$	- -	- -	1	85.96
	$\mathbf{w}^{R}_{2E}$	- -	- -	1	98.05
Resulting Inconsistency	z_1	- -	- -	1.47	0.01
	z_2	- -	- -	1.26	0.01
	z_3	0.69	0.01	0.78	0.01
	z_5	- -	- -	0.80	0.01
	z_6	0.65	0.01	1.05	0.01
	z_{11}	0.96	0.01	- -	- -

3.6 ACCURACY

It is important to stress that inconsistencies in response and linking variables affect the entire solution, not only the copied variables themselves. Table 3.2 summarizes the solutions to the original, 2-level ATC, and 3-level ATC formulations. For the 2-level and 3-level formulations, results are shown for default weights (all weights = 1) and for the weighting update method (WUM) with inconsistency tolerances of 10^{-2} for all variables. In the table, the * symbol indicates that the variable has nonzero inconsistency at the solution, and the value of the variable copy at the parent level is reported. The † symbol indicates that the variable was treated as a static parameter. Notice that solutions using the default weights are far from the solution to the original problem, whereas solutions using the weighting update method are close for all variables. Smaller inconsistency tolerances result in solutions closer to the solution of the original problem.

Table 3.2: Optimal solution to AAO, 2-level ATC, and 3-level ATC formulations

	Original	2-Level ATC		3-Level ATC	
	AAO	Default Weights	WUM (10^{-2})	Default Weights	WUM (10^{-2})
z_1	2.84	2.25	2.83	0.75^*	2.82^*
z_2	3.09	2.04	3.07	0.64^*	3.07^*
z_3	2.36	1.53^*	2.35^*	1.58^*	2.35^*
z_4	0.76	0.76	0.76	0.90	0.76
z_5	0.87	1.00	0.87	0.70^*	0.87^*
z_6	2.81	1.21^*	2.79^*	1.76^*	2.80^*
z_7	0.94	1.30	0.94	0.64	0.93
z_8	0.97	0.93	0.97	0.97	0.97
z_9	0.87	1.07	0.87	0.86	0.86
z_{10}	0.79	0.93	0.80	0.80	0.80
z_{11}	1.30	0.94^*	1.30^*	$1.30^\dagger$	$1.30^\dagger$
z_{12}	0.84	0.84	0.84	0.84	0.84
z_{13}	1.77	1.27	1.76	1.76	1.76
z_{14}	1.55	0.96	1.54	1.55	1.55

3.7 LOCAL CONVERGENCE

One purpose of using the weighting update method is to avoid setting weights arbitrarily high to avoid costly iterations; however, the weighting update method requires an outer loop of additional update iterations to converge on the desired weights, so it is worthwhile examining and comparing the convergence efficiency. The two-level geometric programming problem was resolved using the required weights found by the weighting update method directly as starting weights, thus achieving the desired tolerance without any weighting update iterations. This represents the best possible case that could be attained by guessing weights. Still, in this case the algorithm required almost twice as many function evaluations per element to converge as did the weighting update method. These results are summarized in Table 3.3. Note that the Matlab function *fmincon*, based on SQP, was used in all cases.

Table 3.3: Speed of convergence statistics for the geometric programming problem

	Number of Function Evaluations	Num. Weight Updates
Original Problem AAO	25,173	- -
2 Level ATC Default Weights	***A*:** 241 ***B*:** 110 ***C*:** 115	- -
2 Level ATC Weighting Update	***A*:** 18,002 ***B*:** 8,639 ***C*:** 8,517	4
2 Level ATC Required Weights	***A*:** 31,777 ***B*:** 15,316 ***C*:** 16,297	- -
3 Level ATC Default Weights	***A*:** 195 ***B*:** 158 ***C*:** 152 ***D*:** 24 ***E*:** 19	- -
3 Level ATC Weighting Update	***A*:** 45,092 ***B*:** 34,087 ***C*:** 35,449 ***D*:** 984 ***E*:** 905	5

At first this may appear counterintuitive, but there is an explanation: It took longer to converge when starting with the required weights because the starting point is not close to the solution. Large weighting coefficients act to slow progress of the algorithm by restricting the deviation between parent and child elements at each ATC iteration. Conceptually, this can be thought of as an effect similar to that of a trust-region algorithm, where high weighting coefficients have the effect of tight trust regions, preventing large moves at each iteration. In contrast, the weighting update method first solves the problem with small weighting coefficients, allowing the algorithm to move quickly in the design space and converge to a point close to the final solution. The weights are then updated (increased), and the new problem is solved starting at the solution to the problem with the previous weighting coefficients. In this way, the weighting update method first moves quickly to the proximity of the solution, then tightens tolerances and closes in precisely on the final solution. Results vary based on the problem, acceptable inconsistency tolerance, and the starting point; however, this example shows that using the weighting update method can sometimes be substantially more efficient than even best-case scenario guessing.

Further study on local convergence properties of ATC and the weighting update method is needed before these results can be generalized. Note that in contrast with notions of asymptotic local convergence developed for AAO algorithms, e.g., standard nonlinear programming, local convergence concepts have not been rigorously defined for any system optimization method relying on decomposition, including ATC.

Table 3.3 also shows that the ATC decomposition can be solved with fewer function evaluations per element than the original AAO formulation. It is difficult to compare these cases directly since the objective function of each element is different than the objective function of the AAO formulation; however, generally, each decomposed element will take less computational time per function evaluation than the AAO formulation, and decomposition allows additional possibilities of parallel computing. These results are encouraging because they show that in some situations the decomposed formulation can be solved in less time than the AAO formulation.

3.8 CONCLUSIONS

This chapter showed that it is important to set ATC weighting coefficients appropriately to achieve inconsistency deviations within an acceptable tolerance when top level targets are unattainable. Setting appropriate weights is nontrivial, particularly for multilevel hierarchies where weights at various levels influence each other in complex ways. Setting weights too small can result in solutions far from the solution of the original problem, and setting weights too large can result in excessive computational cost and numerical problems. The weighting update method can automatically find weighting coefficients required for generating a solution with user-specified inconsistency tolerances. This method can help ATC users to achieve acceptable solutions without the burden of trial-and-error searching for appropriate weighting coefficients, which can be intractable for multilevel problems. Despite the added computation involved in iteratively updating the weights, the total computational cost can sometimes be lower than solving the problem directly with the required weights or solving the problem AAO. Future work is needed to define and understand local convergence properties of coordination strategies for hierarchical partitioned systems and bring more rigor to solution efficiency definitions for decomposed optimization strategies.

3.9 NOMENCLATURE

$\Vert\cdot\Vert$	Vector norm
$\langle\cdot\rangle_\alpha$	Vector element α (where α indexes the vector elements, ranging from 1 to the length of the vector)
$\circ$	Element-by-element vector multiplication, for example $[a_1, a_2]^T \circ [b_1, b_2]^T = [a_1b_1, a_2b_2]^T$ and $\mathbf{a} \circ \mathbf{b} = \mathrm{diag}(\mathbf{ab}^T)$
c_{ij}	Number of elements that are children of element j at level i
C_{ij}	The set of elements that are children of element j at level i
E_i	The set of elements at level i of the hierarchy
f_{ij}	Objective function for element j at level i
$\mathbf{g}_{ij}$	Vector function of inequality constraints for element j at level i in negative null form
$\mathbf{h}_{ij}$	Vector function of equality constraints for element j at level i in null form
i	ATC hierarchy level index (starts at level 0)
j	ATC element index
k	ATC element index, used to designate children of element j
l	ATC element index designating the top level element
L_{ij}	The Lagrangian for the formulation of element j at level i
p	ATC element index, used to designate the parent of element j
P_{ij}	Problem formulation of element j at level i
$\mathbf{r}_{ij}$	Vector function that calculates responses for element j at level i
$\mathbf{R}_{ij}^{i}$	Vector of response variable copies at level i for element j
$\mathbf{R}_{ij}^{i-1}$	The $(i–1)th$ level parent-copy of the vector of responses that function as targets for element j at level i
$\mathbf{S}_j$	Binary selection matrix for element j specifying which terms in the parent coordination vector are relevant to element j

$\mathbf{T}$	Vector of top level targets (= $\mathbf{R}^{-1}{}_{0l}$)
$\bar{\mathbf{x}}_{ij}^{i}$	Aggregation vector for all input variables to the response function of element *j* at level *i*
$\mathbf{x}_{ij}^{i}$	Vector of local decision variables for element *j* at level *i*
$\mathbf{y}_{ij}^{i}$	Vector of linking variables for element *j* at level *i*
$\mathbf{y}_{(i+1)j}^{i}$	Vector of coordinating variables for the linking variables in the children of element *j* at level *i*. This vector includes one copy of each linking variable from all of element *j*'s children
$\mathbf{w}_{ij}^{R}$	The weighting coefficient vector for the deviation of responses between element *j* at level *i* and its parent
$\mathbf{w}_{(i+1)j}^{y}$	The weighting coefficient vector for the deviation of linking variables coordinated at element *j* level *i*
α	Index for terms in a vector
β	Index for a specific term in a vector
ε_{ij}^{R}	The tolerance variable for consistency of targets set at element *j* level *i* and the responses of *j*'s children
ε_{ij}^{y}	The tolerance variable for consistency of linking variables coordinated at element *j* level *i* for child elements at the (*i*+1)*th* level
γ	ATC element index, used to designate a specific child of element *j*
$\boldsymbol{\mu}_{ij}$	Vector of Lagrange multipliers for inequality constraints at element *j* level *i*
$\boldsymbol{\lambda}_{ij}$	Vector of Lagrange multipliers for equality constraints at element *j* level *i*
$\boldsymbol{\theta}_{ij}^{R}$	Vector of user specified tolerances for inconsistency deviation between response variables of element *j* at level *i* and targets set by *j*'s parent
$\boldsymbol{\theta}_{(i+1)j}^{y}$	Vector of user specified tolerances for inconsistency deviation between linking variables at level *i*+1 that are coordinated at parent element *j* at level *i*

3.10 REFERENCES

[3.1] Alexandrov, N.M. and R.M. Lewis (2000) "Analytical and computational aspects of collaborative optimization," NASA/TM-2000-210104.

[3.2] Balling, R.J. and J. Sobieszczanski-Sobieski (1994) "Optimization of coupled systems: a critical overview of approaches," *Proceedings of the 5th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA-94-4330-CP, Panama City, p753-773.

[3.3] Braun, R. (1996) *Collaborative Optimization: An Architecture for Large-Scale Distributed Design*, Ph.D. Dissertation, Stanford University.

[3.4] Choudhary, R., A. Malkawi, P.Y. Papalambros (2003) "A hierarchical design optimization framework for building performance analysis," *Proceedings of the*

Eight International Building Performance Simulation Association (IBPSA) Conference, Eindhoven, Netherlands, August 11-14.

[3.5] Cramer, E., J. Dennis, P. Frank, R. Lewis, and G. Shubin, (1994) "Problem formulation for multidisciplinary optimization," *SIAM Journal of Optimization*, v4 p754-776.

[3.6] Kim, H.M. (2001) *Target Cascading in Optimal System Design*, Ph.D. Dissertation, Dept. of Mechanical Engineering, University of Michigan, Ann Arbor, MI.

[3.7] Kim, H.M. (2003) "Target cascading in optimal system design," *ASME Journal of Mechanical Design*, v125 n3 p474-480.

[3.8] Kim, H.M., M. Kokkolaras, L. Louca, G. Delagrammatikas, N. Michelena, Z. Filipi, P.Y. Papalambros, and D. Assanis (2002) "Target cascading in vehicle redesign: a class VI truck study," *International Journal of Vehicle Design*, v29 n3 p199-225.

[3.9] Kim, H.M., D.G. Rideout, P.Y. Papalambros, and J.L. Stein (2003) "Analytical target cascading in automotive design," *ASME Journal of Mechanical Design*, v125 p481-489.

[3.10] Kokkolaras, M., R. Fellini, H.M. Kim, N. Michelena, and P. Papalambros (2002) "Extension of the target cascading formulation to the design of product families," *Journal of Structural and Multidisciplinary Optimization*, v14 n4 p293-301.

[3.11] Kokkolaras, M., L. Louca, G. Delagrammatikas, N. Michelena, Z. Filipi, P. Papalambros, J. Stein, and D. Assanis (2004) "Simulation-based optimal design of heavy trucks by model-based decomposition: an extensive analytical target cascading case study," *International Journal of Heavy Vehicle Systems*, (forthcoming).

[3.12] Lingle, J. and W. Schiemann (1999) *Bullseye! : Hitting Your Strategic Targets Through High-Impact Measurement*, Free Press, New York.

[3.13] MathWorks, The (2004) *MATLAB and Simulink for Technical Computing*, 16-Dec-2004 at *http://www.mathworks.com/*.

[3.14] Michalek, J.J. and P.Y. Papalambros (2005a) "Technical brief: weights, norms, and notation in analytical target cascading," *ASME Journal of Mechanical Design*, (forthcoming).

[3.15] Michalek, J.J. and P.Y. Papalambros (2005b) "An efficient weighting update method to achieve acceptable consistency deviation in analytical target cascading," *ASME Journal of Mechanical Design*, (forthcoming).

[3.16] Michelena, N., P. Papalambros, H.A. Park, and D. Kulkarni (1999) "Hierarchical overlapping coordination for large-scale optimization by decomposition," *AIAA Journal* v37 n7 p890-896.

[3.17] Michelena, N., H. Park, and P. Papalambros (2003) "Convergence properties of analytical target cascading," *AIAA Journal*, v41 n5 p897-905.

[3.18] Papalambros, P.Y. and D.J. Wilde (2000) *Principles of Optimal Design: Modeling and Computation*, 2nd ed., Cambridge Univ. Press, New York.

[3.19] Park, H., Michelena, N., Kulkarni, D. and P. Y. Papalambros (2001) "Convergence criteria for hierarchical overlapping coordination under linear constraints," *Computational Optimization and Applications*, v18 n3 p273-293.

[3.20] Shima, T. and Y.Y. Haimes, (1984) "The convergence properties of hierarchical overlapping coordination," *IEEE Transactions on Systems, Man, and Cybernetics* vSMC-14 n1 p74-87.

[3.21] Sobieszczanski-Sobieski, J. (1989) "Optimization by decomposition: a step from hierarchic to non-hierarchic systems," *2nd NASA/Air Force Symposium on Recent Advances in Multidisciplinary Analysis and Optimization*, Hampton, VA, September 28-30. NASA TM-101494, NASA CP-3031.

[3.22] Sobieszczanski-Sobieski, J., B.B. James, and A.R. Dovi (1985) "Structural optimization by multilevel decomposition," *AIAA Journal*, v21 p1291-1299.

[3.23] Sobieszczanski-Sobieski, J. and S. Kodiyalam (1999) "BLISS: a new method for two-level structural optimization," *AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, St. Louis, MO v2 p1274-1286.

[3.24] Tzevelekos, N., M. Kokkolaras, P.Y. Papalambros, M.F. Hulshof, L.F.P. Etman, and J.E. Rooda (2003) "An empirical local convergence study of alternative coordination schemes in analytical target cascading," *Proceedings of the 5th World Congress on Structural and Multidisciplinary Optimization*, Lido di Jesolo-Venice, Italy, May 19-23.

CHAPTER 4

COORDINATION OF MARKETING PREFERENCE MODELS WITH ENGINEERING PERFORMANCE MODELS

In this chapter the analytical target cascading methodology is applied to coordinate marketing profit maximization goals, built upon econometric choice models of consumer preference, with engineering design models of product feasibility and performance to achieve joint solutions with both technical and market feasibility. Here design of a single product is pursued from the perspective of a single profit-seeking producer. It is shown that the solution obtained through coordination is superior to that achieved by treating the two disciplines independently. This chapter serves to introduce a number of concepts that will be expanded upon in Chapter 5 and Chapter 6. The material in this chapter is based on publications by Michalek, Feinberg, and Papalambros (2004, 2005).

4.1 INTRODUCTION

Product development, as a costly and time-consuming prelude to the introduction of new products, has been the object of intense study by practitioners and academics in both marketing and engineering design. The academic literature proposes a number of models to help guide product planners in assessing consumer needs or "value systems," as well as to capitalize on synergies in the production process itself. As such, the entire process is typically broken down into a number of stages which, for parsimoniousness and reasons of disciplinary boundaries, are addressed separately in product optimization. For example, in marketing one may ask, *given* a set of known characteristics and levels –

and, presumably, an expedient method for delivering them in one product at an attractive price point – which combination would most appeal to consumers, or which combination would be most profitable. Engineering design faces the converse problem of delivering an optimal, feasible product *given* a set of desired performance targets, features and costs. That is, each discipline works under constraints and guidelines set by the other.

In marketing conjoint studies, for example, product characteristic 'levels' are chosen to be in line with engineering guidelines and so, in a sense, are conditional upon them. If engineering cannot deliver a specific product characteristic or some particular value of it, consumers are not asked for their reactions to it. So, what we learn about consumer preferences, and thereby the set of products produced, is contingent on knowing in advance which targets are technically infeasible or unrealistic. Similarly, engineering design models aim to maximize or achieve target levels of performance characteristics, subject to physical and production constraints, without knowing if consumers would want and pay for them.

A turnkey system formalizing product optimization through coordinated communication between established marketing and engineering design models has not emerged for a number of reasons. First and foremost are reasons of historical development and disciplinary boundaries: Research on product development in marketing has long differed from that in engineering design in terms of product representation and choice of performance and success metrics (Krishnan and Ulrich 2001). For example, in marketing a product is often modeled as a "bundle of attributes" (e.g., McAlister 1982), over which consumers have preferences represented by utilities, so that firms can manipulate the former to maximize the latter. In engineering design, by contrast, products may be described as complex assemblies of interacting components, for which parametric models are built to represent design decisions, such as shape, size, and configuration, which then are manipulated to maximize performance objectives. Measurements of "success" also differ between the two disciplines, with marketing assessing degree of market fit, consumer satisfaction, overall share and ultimately profit, and engineering design concerned with technical performance, innovativeness and cost effectiveness. The two disciplines even point to different critical success factors external to the design process itself: Marketers stress the importance of positioning, advertising messages, choosing the right price tier and understanding "consumer needs" using data, while engineers generally use intuition when dealing with customer needs, emphasizing the creativeness and functionality of the product concept and working toward technical objectives such as reliability, durability, environmental impact, energy use, heat generation, manufacturability and cost reduction, among others. In short, the two communities do not so much disagree on the product development process as have different scopes, perspectives, languages and notions of drivers of success.

Second, marketing and engineering design models differ in terms of domains and control variables, and so their corresponding models of the product development process do not easily speak to one another. For example, in marketing, a chief goal is simply figuring out what consumers want, addressed through such methods as focus groups, test markets, surveys and measurement models like conjoint. Consumer preferences are taken as primitives, and one optimizes over levels of product characteristics, looking for 'sweet spots' in the product characteristic space. Engineering design, by contrast, seeks to meet specific performance goals, conditional on existing production processes and the realities of physics and geometry. In short, both optimization variables and the nature of constraints are very different. Formal models attempting to combine them would therefore have to span an unusually broad domain.

Among the main problems identified by Krishnan and Ulrich (2001) is effective communication between marketing and engineering design. Even with full information and broadly-validated modeling frameworks, miscommunication can lead to sub-optimal product designs, a problem particularly pronounced for high-tech products where the marketing and engineering design domains are widely separated. Such claims have broad antecedents in prior empirical and theoretical research. Gupta *et al*. (1986) studied the interface between marketing and R&D in American hi-tech firms, and found that other factors (besides standard market uncertainty and firm strategy explanations) exert strong effects, most notably organizational design and "sociocultural" differences between marketing managers and their R&D counterparts. This is hardly confined to US firms: Song and Parry (1992) confirmed these findings for over 200 analogous firms in Japan, while also cataloguing subtle points of difference between the two nations' development cultures. Souder (1988), in analyzing a vast database of product development projects, found a variety of consistent problem types between marketing and R&D managers, and thus formulated a model to improve integration between the two. Griffin and Hauser (1992) further considered the multiple interfaces between marketing, engineering and manufacturing, comparing the effects of Quality Function Deployment (QFD) to more traditional project development approaches. They consistently underscore the critical role and nature of information flow, and how QFD uniquely allows enhanced "horizontal" flow through the development team.

In this chapter, a new approach is proposed to link marketing and engineering product design decision-making formally. In doing so, the intent is not to merge the two, but to make use of their respective strengths and to capitalize on models that are especially well-suited to joint optimization. From marketing, methods from discrete choice analysis are adopted, as applied to efficient conjoint designs; both have deep theoretical roots and have been validated in hundreds of disparate empirical studies throughout the world (Cattin and Wittink 1982, Wittink and Cattin 1989, Wittink *et al.* 1994). From engineering design, the analytical target cascading methodology is adopted, in addition to general design optimization methodologies. It is demonstrated how these pre-existing methods can work in tandem to converge on optimal product designs, avoiding the time-consuming, error-prone and costly iterations that often characterize complex product development processes.

Several key ideas pervade this approach. The first is a departure from assumptions made in certain marketing models, which often hold that design problems are primarily ones of capital: With sufficient funds, any desired *combination* of product characteristics can be achieved. Although marketers are aware that, strictly speaking, some such combinations are quite difficult to attain (Urban and Hauser, 1993) it is fairly common in conjoint studies, for example, to allow product characteristics to be paired according to the needs of the experimenter (e.g., Haaijer 1999). Such a premise is less commonly enacted by engineers, who deal more directly with feasibility constraints: designs that cannot exist under present technology or that are physically impossible with any technology. The methodology presented is particularly suited to support the study of complex durables where feasibility constraints prohibit some combinations of characteristics from being achievable at any cost. The empirical application in this chapter considers such constraints, where some product characteristic combinations literally are impossible to achieve, not just difficult or costly.

The second idea is the often-underestimated complexity of interrelations among engineering constraints, even for simple artifacts such as the dial-readout scale studied later in this chapter. When this complexity is further mapped onto the product characteristics space, simple strategies, like restricting product characteristic levels to feasible combinations in a conjoint study, has little value and will *not* lead to optimal solutions. The simple case study of this chapter is an empirical demonstration of such a situation, and illustrates the need for caution when studying design domains of higher complexity.

The final idea is the key role of iteration: Marketing and engineering design decisions must be iteratively updated, preserving the individuality of each discipline but converging to the optimum for the product, not for the discipline. If marketing sets targets without iterative interaction with engineering design, results can be substantially inferior than under the proposed ATC coordination method. This is especially intriguing, given that the model used to account for consumer preferences can call upon the wealth of methods developed for conjoint measurement, methods that offer optimal product designs from a marketing viewpoint – so long as they are checked against their technical realization.

In the next sections a short review of the relevant literature in marketing and engineering design is provided, ATC is introduced as a formal linking mechanism between marketing and engineering design, and its use is demonstrated on a simple durable product: a dial-readout bathroom scale.

4.2 MARKETING AND ENGINEERING IN PRODUCT DESIGN

4.2.1 Marketing Product Planning Models

Kaul and Rao (1995) provide an integrative review of product positioning and design models in the marketing literature. They differentiate between product positioning models, which involve decisions about abstract perceptual attributes, and product design models, which involve choosing optimal levels for a set of physical, measurable product characteristics. In this dissertation only measurable product characteristics are used; however, a comprehensive framework similar to the one proposed by Kaul and Rao could be used to include perceptual attributes, product positioning, and consumer heterogeneity. In this dissertation, conjoint-based product design models from a marketing perspective will be referred to as *product planning models*.

Optimal product planning in the marketing literature is typically posed as selection of optimal price and product characteristic levels that achieve maximum profit or market share. For complex products, where engineering constraints may prevent some combinations of product characteristic levels from being technically attainable, it is difficult to define explicitly which combinations of characteristics are feasible. Even if these combinations can be defined and eliminated from a conjoint study, the optimal solution using the conjoint data may still contain infeasible combinations of product characteristics. For such products, planning decisions made without engineering input may well yield inferior solutions.

4.2.2 Engineering Design Models

The engineering design optimization literature focuses on methods for choosing values of design variables that maximize product performance objectives. Papalambros and Wilde (2000) provide an introduction to engineering design optimization modeling techniques, strategies and examples. When multiple conflicting optimization objectives exist, the solution is a Pareto set of optimal products, and the choice of a single product from that set requires explicit expression of preferences among objectives. Such preferences are notoriously difficult to define in practice. Some methods use interactive, iterative searches to elicit preferences, relying on intuition in navigating the Pareto surface and choosing an appropriate design (Diaz, 1987).

As discussed in Chapter 2, recent efforts in the design literature take the approach of resolving tradeoffs among technical objectives by proposing models of the producer's financial objective (Hazelrigg, 1988; Li and Azarm, 2000; Gupta and Samuel, 2001; Wassenaar and Chen, 2003; Georgiopoulos, 2003; Georgiopoulos *et al.,* 2004). Gu *et al*. (2002) build on this work using the collaborative optimization MDO framework to coordinate decision models in the engineering and business disciplines. In the approach here, by contrast, a hierarchy of product planning and engineering design models are coordinated to design a product using ATC, which is proven to converge to joint solutions for arbitrarily large hierarchies, as discussed in Chapter 3. Unlike prior approaches in engineering design, explicit econometric models of consumer preference distributions are developed conditional on actual survey choice data for coordination with engineering design decisions.

4.2.3 Prior Approaches to Integrating Engineering into the Marketing Product Design Literature

Over the last decade, the marketing literature has increasingly turned to questions of integration with engineering design and production, usually noting the difficulty of doing so. Early discussions of problems integrating information from various sources in the product development process include Griffin (1997) and Wind and Mahajan (1997). Griffin was among the first to highlight the use of cross-functional teams to shorten – and purportedly optimize – the product development process, a process especially lengthy for innovative and for highly complex products. Wind and Mahajan went so far as to issue a warning about the "inadequacy" of modeling techniques in marketing to encompass the entire new product development process, particularly as design incorporates information from multiple sources. Certain externalities can come into play as well; for example, Moorman and Slotegraaf (1999) highlight how information in the external environment can stimulate firms to deploy their technology and marketing capabilities so as to influence the level and speed of relevant product development activities. They conclude that the most valuable characteristic of firm capabilities may be their ability to serve as "flexible strategic options." How they might accomplish this, in practical terms, is still largely an open question.

Part of the 'integration problem' is certainly one of terminology and conceptualization. Garcia and Calantone (2002), for example, detail the often contradictory ways in which notions of innovation are used in the new product development literature, particularly in marketing and engineering design. They emphasize the importance of maintaining both marketing and technological perspectives when discussing innovations and the relative lack of empirical work directed at "really new" innovations, and offer a set of measures to help classify innovations across the domains of practitioners and academics.

Researchers in marketing have often pointed to engineering design and production as the key contributors to product success or failure. Sethi *et al*. (2001) stress that multiple studies have found that the primary determinant of new product success is innovativeness, the extent to which a new product provides meaningfully unique benefits, rather than the ability to satisfy pre-existing wants of the type uncovered in a typical conjoint study. Srinivasan *et al*. (1997), in addressing the concept selection stage of a new product development process, emphasize the importance of utilizing both product characteristic-based customer preference and product cost models, and they offer empirical evidence for the need to push beyond such models to more complete "customer-ready" prototypes. Halman *et al*. (2003) additionally consider the advantages of a platform-based approach to product development, showing how economies-of-scale enhance both marketing and physical production. Although they are primarily interested in product lines (as opposed to individual products), they underscore the paucity of literature linking marketing with engineering practice in product management.

Hauser (2001), by contrast, emphasizes the sheer complexity of the development process, in terms of coordination of resources and agents with multiple criteria for success (for example, speed to market, customer satisfaction and product quality). He applies agency theory to formulate a set of metrics and a weighting method to help firms balance and optimize such complex development processes, with variables spanning concerns from both marketing and engineering design.

A great majority of research on new product success has focused on product characteristics and the product development process, rather than interactive and ancillary factors. In their study of product launch support, Hultink *et al.* (2000) examine data on many hundreds of product introductions and identify divergent product success criteria for what they distinguish as "consumer goods" and "industrial goods." For example, the former seem to benefit from strategies that defend market positions, while the latter benefit from those that leverage technological innovations to penetrate new markets. Furthermore, the optimal marketing 'mix' – the relative emphasis on consumer-oriented variables like promotion, display and advertising, and more nuts-and-bolts technological dimensions – differs systematically between the two product types, so it is unsurprising to see them receiving different degrees of emphasis in the marketing and engineering design communities. However, Kahn (2002) finds that marketing assumes primary responsibility for making market forecasts across both types of goods, with a considerably shorter horizon for "consumer goods."

Pullman *et al.* (2002) present one of the few studies on the relative effectiveness of marketing-based and engineering-based approaches to optimal product design; specifically, they consider conjoint analysis (a marketing-based method, and one used in the present study) and QFD (a more engineering-oriented approach). They found that the two approaches converged on many of the most important features, but that the engineering approach was better able to highlight those characteristics which had *both* positive and negative aspects. Further, they found that the marketing approach better identified *current* consumer preferences, while the engineering approach better identified *core* consumer needs. A major conclusion of their study is that the two approaches should be pursued in tandem. Still, no formal system combining them is presently available to the design community. While this dissertation does not directly address tools such as QFD, which offer help to designers in the absence of mathematical product models, it instead presents a formal methodology to coordinate results of conjoint analysis with product models when such models can be called upon.

Finally, Leenders and Wierenga (2002) offer a major review of extant approaches to the interplay and integration of marketing and engineering design with a particular emphasis on relative effectiveness. They found that one of the most effective methods is simply locating marketing and product development team members closer together to facilitate the interchange of information and, presumably, to encourage a form of joint optimization. They caution, however, that while encouraging this sort of information exchange does indeed enhance new product performance, it nevertheless carries substantial costs, mainly in the great deal of complexity intrinsic in formalizing such relationships. Although it does not formally address industrial design (aesthetics) or manufacturing (physical production), the present study serves to initiate development of the very formalization between marketing and engineering design which prior authors have underscored as problematic.

4.2.4 New Product Development.

Although the NPD literature is too vast to summarize here competently, other authors have done so in articles devoted to the subject. Brown and Eisenhardt (1995), in a broadly integrative survey, present a snapshot of the burgeoning product development literature, distinguishing three major themes, development as rational planning, as a web of communication, and as "disciplined" problem-solving. Based on these broad distinctions, they fashion a model of critical success factors in product development, paying unusual attention to the distinct roles of various actors – senior management, project leaders, suppliers, purchasers – and the vital interplay afforded by communication among them at various stages of the development process. Meta-analyses of the product performance literature are provided by Montoya-Weiss and Calantone (1994) and by Henard and Szymanski (2001); both synthesize decades of prior research in the area with an eye towards generalizations, though the former does report a large number of points of divergence, despite commonalities in methodological approach in the surveyed literatures. Henard and Szymanski (2001) focus specifically on the key determinants of relative product success. Of two dozen such factors identified by prior authors, they find that the most broadly critical include product advantage, market potential, the ability to match customer needs, and pre-existing firm proficiencies. Interestingly, the role of communication between various firm entities is largely that of mediator, in that many of the critical success factors can be directly affected by it.

For the purposes here, among the key conclusion of all these prior lines of research is this: Fostering effective, ongoing communication between marketing and engineering design (among other entities) is a critical factor in the eventual success of a product development project. The methodology presented next is uniquely suited to accomplishing exactly that goal in a rigorous, mathematical design system.

4.3 METHODOLOGY

Here ATC, introduced in Chapter 3, is proposed as a joint system for product development that calls upon methodologies from both engineering design and marketing. The ATC framework allows a joint planning and engineering design problem to be formally decomposed into a (marketing) product planning subproblem and an engineering design subproblem. Each of these separately has been the object of intense study, and among the felicities of ATC is the ability to call on methods for optimizing each of these subproblems in order to attack the much more difficult joint problem and to prove that obtained solutions are identical. The product planning subproblem is well-known to marketers, as it involves choosing product characteristics and price (e.g., as warranted by a conjoint or similar model) that will maximize some firm-level objective function. Here, a simple model of expected firm profit is used, contingent on an estimated discrete choice model for demand, such as logit or probit. The ATC formalism allows flexibility in this regard, and profit is just one among any number of possible firm-level objectives. The engineering design subproblem is quite different, and involves choosing a *feasible* design that achieves known target product characteristics as closely as possible. Using ATC, the subproblems are iteratively solved until the joint system converges upon a consistent optimal product design. Although the model is presented formally in the following sections, the main ideas are depicted in Figure 4.1.

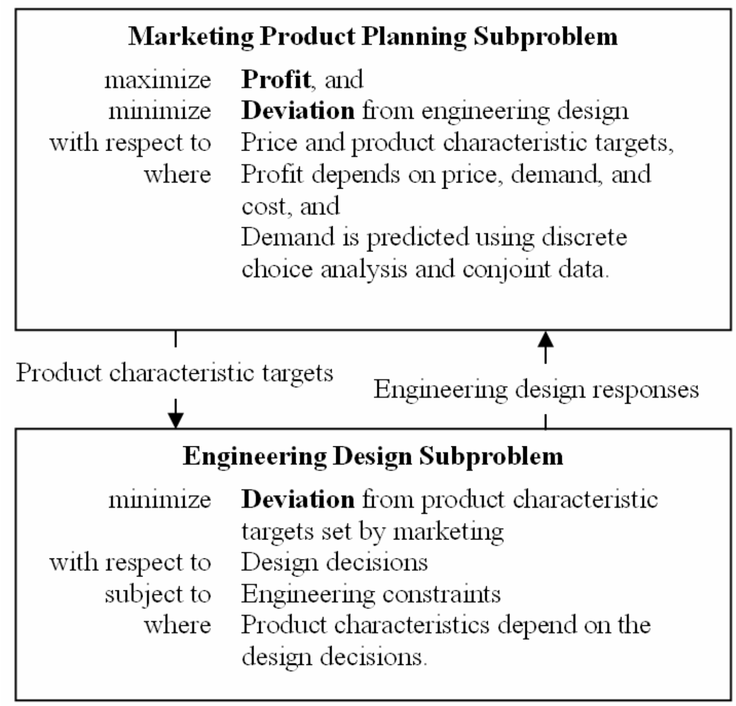


Figure 4.1: ATC formulation of the product planning and engineering design subproblems

In this setup the marketing product planning subproblem requires that profit be maximized with respect to product characteristics and price, but also stipulates minimal deviation from an achievable engineering design; it is in this last requirement that the formulation differs from the one typical in marketing applications. The engineering design subproblem is in some sense the dual: It sets design decisions to minimize deviations from the product characteristics requested by marketers, but must respect engineering constraints, which often are exogenous in the sense of being dictated by, for example, geometry and physics. The two problems ‘speak’ to one another in a very natural sense. In real organizations, it is typical for one group, either marketing or engineering design, to deliver an initial set of specifications, which the other attempts to meet, starting off an actual iterative process between the two. Here models are used to perform iterations and to reach a desirable, feasible, and consistent solution. Potential savings in reducing the 'physical' iteration between groups of people are not addressed directly, but these likely will be considerable if the models required for ATC are available.

4.3.1 Analytical Target Cascading

As discussed in Chapter 3, analytical target cascading is a methodology for systems optimization. It works by viewing a complex system as a decomposable hierarchy of interrelated subsystems, each of which can be analyzed and optimized separately and then coordinated (Kim 2001). In order to apply ATC, one must have a mathematical model for each of the subsystems – which in general can be numerous, although this chapter refers to only one engineering design subsystem – so that one can compute subsystem response as a function of decisions made for that subsystem. Given the various mathematical models for the subsystems, the modeler organizes them into a *hierarchy*, as in the computer example shown in Figure 4.2; note that the top level represents the overall system and each lower level represents a subsystem of its parent element. The process would be similar for even small durables, although the number of subsystems and their potential interactions would be smaller. For example, in the dial-readout scales studied here, the hierarchy consists only of one marketing subsystem (parent) and one engineering design subsystem (child); however, in general both the marketing and engineering models could consist of a hierarchy of any number of submodels.

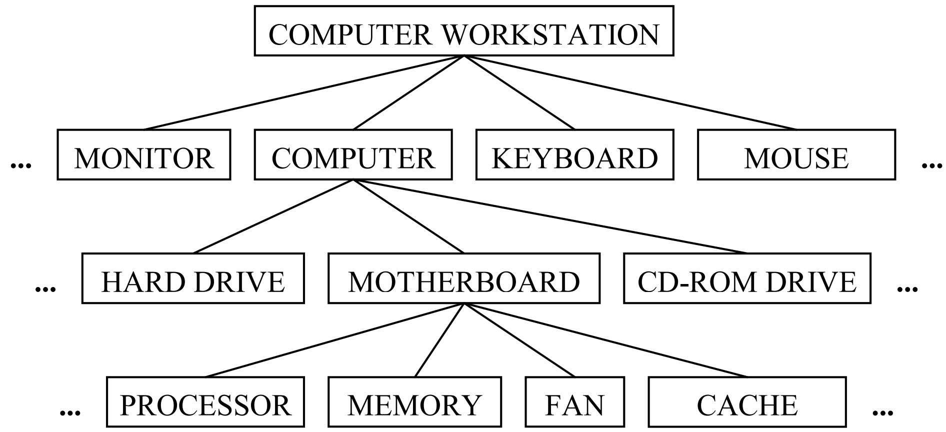


Figure 4.2: Hypothetical ATC hierarchy for a computer workstation

In the ATC process, top-level system design targets are propagated down to subsystems, which then are optimized to match the targets as closely as possible. The resulting responses are rebalanced at higher levels by iteratively adjusting targets and designs throughout the hierarchy to achieve consistency, the latter process called the coordination strategy. Chapter 3 showed that, using certain classes of coordination strategies, the ATC formulation will converge to the same solution as the un-decomposed (or "all-at-once") problem, within a user-specified tolerance.

Using ATC can be advantageous because it organizes and separates models and information by focus or discipline, providing communication only where necessary. Some problems that are computationally difficult or impossible to solve all-at-once can be solved using ATC, and in some cases ATC can result in improved computational efficiency because the formulation of each individual element typically has fewer degrees of freedom and fewer constraints than the all-at-once formulation.

As mentioned earlier, the formulation and example presented in this chapter contains a hierarchy of only two elements: the marketing product planning subproblem *M* and the engineering design subproblem *E*, which is the child (sub-level) of *M*. However, for complex systems ATC allows the flexibility to model the engineering design subproblem as a hierarchy of subsystems and components rather than with a single element. It is possible to conceive of a formulation where marketing tasks are also modeled as a hierarchy such that the product planning subsystem interacts with engineering design while other subsystems represent other aspects of the marketing mix such as promotion, packaging, pricing and positioning.

In the following subsections, the engineering design model and product planning model used in this chapter are described in detail.

4.3.2 ATC Engineering Design Subproblem

In the engineering design subproblem, design characteristics $\mathbf{z}$ are calculated as functions of the design variables $\mathbf{x}$ using the response functions $\mathbf{r}(\mathbf{x})$, where the variables $\mathbf{x}$ are constrained to feasible values by constraint functions $\mathbf{g}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$. General procedures for defining design variables $\mathbf{x}$, response functions $\mathbf{r}(\mathbf{x})$, and constraint functions $\mathbf{g}(\mathbf{x})$, and $\mathbf{h}(\mathbf{x})$ to define a product design space are well established in the design optimization literature (Papalambros and Wilde, 2000); however, modeling specifics are entirely product dependent. The objective function of the engineering subproblem is to minimize deviation between the product characteristics achieved by the design $\mathbf{z}_E$ and the targets set by marketing $\mathbf{z}_M$. Using ATC notation introduced in Chapter 3, this objective function is written as

$$\left\|\mathbf{w} \circ \left(\mathbf{z}_M - \mathbf{z}_E\right)\right\|_2^2, \tag{4.1}$$

where $\| \|_2^2$ denotes the square of the l_2 norm, $\mathbf{w}$ is a weighting coefficient vector, and $\circ$ indicates term-by-term multiplication, such that $[a_1\ a_2\ ...\ a_n] \circ [b_1\ b_2\ ...\ b_n] = [ab_1\ ab_2\ ...\ ab_n]$. For complex products, engineering constraints typically restrict the ability to meet some combinations of product characteristic targets, and the ATC process acts to guide marketing in setting achievable targets while designing feasible products that meet those targets.

4.3.3 ATC Marketing Product Planning Subproblem

In the marketing planning subproblem, a fairly simple model of profit, Π, is adopted, which in the standard way is taken to be revenue minus cost,

$$\Pi = q\left(p - c_V\right) - c_I \,. \tag{4.2}$$

Here, q is the quantity of the product produced and sold (product demand), p is the selling price, c_V is the variable cost per product, and c_I is the investment cost. It would be further possible to augment this model in any number of ways popular in the extant literature, for example, a discount function to capture the time value of money, a separate term to account for fixed costs or salvage value, or a concave loss-like function for risk and uncertainty. Nevertheless, Eq.(4.2) captures the main forces at work, and can be readily modified. Among the firm's decision variables are pricing and product characteristics. For simplicity here the variable and investment costs c_V and c_I are considered constant across all possible product designs; however, they could also be written as functions of the engineering design decisions, as will be discussed in Chapter 6. Note that overall demand q depends on price p as well as on product characteristics $\mathbf{z}$.

A straightforward version of choice-based conjoint using the standard logit model is called upon to establish a plausible demand function q as a function of the decision variables $\mathbf{z}$ and p; see the Appendix for a detailed introduction to discrete choice models, and see, for example, Louviere and Woodworth (1983) for an early application of a similar model. Only the design of a single product is considered here, and so the types of heterogeneity corrections allowed by more recent latent class and hierarchical Bayes approaches are less relevant here than they would be in the case of the product *lines* discussed in Chapter 5, thus simplifying implementation considerably. Andrews *et al.* (2002) provide a full discussion of these issues specifically in the context of conjoint. Finally, demand is formulated with the producer operating as a monopolist or at least in a market where the firm's decision variables do not result in predictable systematic variation in the actions of other firms (i.e., in a so-called "zero conjectural variations" setting). It is possible to adopt a game theoretic setting to account for potential oligopoly, and a version of such a set-up applied in a similar production-based context is discussed in Chapter 7 and in Michalek, Papalambros and Skerlos (2005).

A vast body of work in discrete choice analysis has enabled the modeling of choices made in uncertain environments (Train, 2003). As is typical in marketing applications, a random utility formulation is used to link observed covariates – here, price and product characteristics – to observed individual-level choices. Formally, there is a set $\mathcal{J}$ of product alternatives numbered 1, 2, ..., J with deterministic components $\{v_1, v_2, \ldots, v_J\}$ and associated errors $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_J\}$; to account for the possibility of no alternative being acceptable, there is also an 'outside good,' indexed as alternative 0, with error ε_0 and attraction value v_0 normalized to zero ($v_0 = 0$). The probability P_j that we observe a choice of alternative j is equal to the probability that alternative j has the highest utility:

$$P_j = \Pr\left[v_j + \varepsilon_j \geq v_{j'} + \varepsilon_{j'}, \forall j' \in \mathcal{J} \right] \tag{4.3}$$

Computational efficiency depends critically on the distribution assumed for the ε random error terms in Eq.(4.3). Errors can take several forms, and it generally requires extremely large samples for assumptions about distributional error to have any substantive impact; consequently, researchers often work with error specifications allowing the most tractability. For example, if errors are assumed to be normally distributed, then the form of P_j is called the *multinomial probit model*, which does not admit of closed-form expressions for choice probabilities in terms of underlying attractions. However, if ε terms are assumed to be Type II extreme-value (or Gumbel) distributed (i.e., $\Pr[\xi < x] = \exp[-\exp(-x)]$), as in Guadagni and Little (1983), then it can be shown that

$$\begin{aligned} P_j &= e^{v_j}\left[1+\sum_{j'\in\mathcal{J}} e^{v_{j'}}\right]^{-1} \\ P_0 &= \Pr\left[\text{No Choice}\right] = \left[1+\sum_{j'\in\mathcal{J}} e^{v_{j'}}\right]^{-1} \end{aligned} \tag{4.4}$$

where the "1" in the denominator accounts for the outside good, with $v_0 = 0$ (see Train 2003, Chapter 3 for proof). This form is called the *multinomial logit model* (MNL). Note that, even in a monopolist setting, the presence of an outside good ensures that the probability of "no choice" is always non-zero, and so choice probabilities for undesirable or overly expensive products will be low.

It is assumed that v can be measured as a function of observable quantities such as price, product characteristics, consumer characteristics, etc. In this dissertation only price and product characteristics are considered. A rule is needed for mapping prices and product characteristics into the deterministic component of utility v. A good deal of recent work examines non-parametric methods for accommodating individual-level (Kalyanam and Shively 1998; Kim *et al*. 2003c) or latent utility functions (Andrews *et al*. 2002). These are, however, computationally intensive and difficult to embed within an iterative optimization scheme. Instead, a standard linear mapping of product characteristic *levels* (conjoint part-worths) is used, fitting natural splines to interpolate intermediate values. The observable component of utility v_j for product j is then written as

$$v_j = \sum_{\zeta=1}^{Z} \sum_{\omega=1}^{\Omega_\zeta} \beta_{\zeta\omega} \delta_{j\zeta\omega} \,, \tag{4.5}$$

where $\delta_{j\zeta\omega}$ is a binary dummy variable such that $\delta_{j\zeta\omega} = 1$ indicates alternative j possesses characteristic or price ζ at level ω, and $\beta_{\zeta\omega}$ is the "part-worth" coefficient of characteristic or price ζ at level ω. In $\delta_{j\zeta\omega}$ the elements of the product characteristic vector **z** are enumerated as $\zeta = \{1, 2, \ldots Z\}$, and price p included as the term $\zeta = 0$. Each product characteristic or price ζ is discretized into Ω_ζ levels, $\omega = \{1, 2, \ldots \Omega_\zeta\}$. One advantage of using discrete levels is that it does not presume linearity with respect to the continuous variables. For example, we cannot assume that a \$5 price increase has the same effect for a \$10 product as it does for a \$25 product.

Given a set of observed choice data, values can be found for the β parameters such that the likelihood of the model predicting the observed data is maximized. A great deal of research in marketing is devoted to recovering model parameters through latent classes, finite mixtures or using hierarchical Bayes methods (Andrews, Ainslie, and Currim, 2002); however, here the standard maximum likelihood formulation (Louviere *et al*., 2000) is used. The log of the sample likelihood for a particular individual on a particular choice occasion t is:

$$\sum_{j \in \mathcal{J}_t} \Phi_{jt} \ln \left[\frac{\exp\left(\sum_{\zeta=1}^{Z} \sum_{\omega=1}^{\Omega_\zeta} \beta_{\zeta\omega} \delta_{j\zeta\omega} \right)}{1 + \sum_{j' \in \mathcal{J}_t} \exp\left(\sum_{\zeta=1}^{Z} \sum_{\omega=1}^{\Omega_\zeta} \beta_{\zeta\omega} \delta_{j\zeta\omega} \right)} \right], \tag{4.6}$$

where $\Phi_{jt} = 1$ if the observed choice on choice occasion n is alternative j and $\Phi_{jt} = 0$ if j is not the observed choice. Here $\mathcal{J}_t$ is the set of alternatives available on choice occasion t. Eq.(4.6) is maximized with respect to the $\beta_{\zeta\omega}$ terms after summing across all individuals and choice occasions. In this way, the part-worth coefficients $\beta_{\zeta\omega}$ are obtained for each level ω of each product characteristic or price ζ.

In all random utility models, such as the logit used here, one must be careful about *model identification*; for example, adding a constant term to all attraction values v shifts them upward to the same extent and does not change choice probabilities predicted by the logit model. Thus, in using Eq.(4.5), there is an infinite number of solutions for optimal β values that predict equivalent choice probabilities and therefore have identical likelihood values. Standard practice is to impose an *identification constraint* on the system of coefficients, which unambiguously chooses just one among all possible 'optimal' solutions. Such constraints typically set a linear combination of the coefficients to zero. For clarity, we select from the infinity of equivalent solutions the one solution where the mean coefficient value $\sum_{\zeta=1}^{\Omega_\zeta} \frac{\beta_{\zeta\omega}}{\Omega_\zeta}$ is the same for all ζ. By adding this constraint the model has $1+\sum_{\zeta=1}^{Z}\left(\Omega_\zeta - 1\right)$ degrees of freedom, and the solution is uniquely defined (i.e., "identified").

The $\beta_{\zeta\omega}$ terms represent part-worths of discrete values but have no information about intermediate values. To optimize over continuously valued product characteristics and price, it is necessary to estimate utilities for such intermediate values. To this end, polynomial splines can be used, because linear splines are not differentiable at knots (the estimated values). In the case study to follow natural cubic splines are used although with a greater number of characteristic levels higher-order splines would be possible. Lastly, then, the deterministic component of utility can be written as a function of the continuous-valued product characteristics values $\mathbf{z}$ and price p using the spline function

Ψ_ζ of the discrete level part-worths $\beta_{\zeta\omega}$ for each characteristic or price ζ. If price is indexed as $\zeta = 0$, the attraction value is written as

$$v_j = \Psi\left(\mathbf{z}_j, p\right) = \Psi_0\left(p_j\right) + \sum_{\zeta=1}^{Z} \Psi_\zeta\left(\left\langle \mathbf{z}_j \right\rangle_\zeta\right), \tag{4.7}$$

where the angle bracket notation $<\mathbf{z}_j>_\zeta$ indicates the ζ^{th} element of the vector z_j.

Thus far, only the question of relative preference among the alternatives has been addressed, as embodied by choice probability. The model specification is completed by invoking a known market potential, *S*. This is reasonable, given the quasi-monopolist setting, although it is acknowledged that markets with some degree of category expansion – as a function of price and product characteristics – would need to have market potential measured as a function of those quantities, after which optimization could be carried out. Given market potential *S*, demand q_j for product *j* is linearly related to choice probabilities:

$$q_j = sP_j = Se^{v_j}\left[1 + \sum_{j' \in \mathcal{J}} e^{v_{j'}}\right]^{-1} \tag{4.8}$$

Such market potentials can be given exogenously at the outset or estimated through a variety of techniques based on historical data or test markets (see Lilien *et al*. (1992) for a full review of such methods). The ATC methodology requires only representation of demand as a function of price and product characteristics, not necessarily one related to the form chosen for this or any particular study.

Maximum likelihood estimation can be used to fit β parameters to any set of observed choice data; however, collinearities in the characteristics and price of the choice sets can make accurate parameter estimation difficult and can cause problems generalizing to new choice sets (Louviere *et al*., 2000). Conjoint analysis (CA) has been widely used to develop efficient, orthogonal and balanced survey designs (experimental designs) to determine which product characteristics are important to consumers, and appropriate levels for each characteristic. There is a vast literature on conjoint analysis and appropriate experimental designs, and the reader is directed to any of the classic or recent articles, notably Louviere's (1988) expository article, the review by Green and Srinivasan (1990) or Kuhfeld's (2003) exhaustive account.

Conjoint studies present subjects with a series of products or product descriptions, which they evaluate. Products can be presented in various ways, but characteristic levels are always made clear, either in list form, pictorially, or both. Subjects can indicate their preferences among products by ranking (i.e., putting in an ordered list), rating (for example, on a 1-10 scale) or choosing their favorite from a set. Choice-based conjoint is used for data collection because it is more natural for respondents (who choose products, rather than rating or ranking them in their daily lives). Concordant with standard practice (Kuhfeld, 2003), efficient designs are generated to collect maximum information about preferences with a minimum number of questions, offering successive sets of products and asking which is most preferred in each, or whether none is acceptable (the "no choice" option).

4.3.4 Complete Formulation

Figure 4.3 provides an overview of the modeling process. The demand model is constructed by collecting consumer choice data using a choice-based conjoint survey. The logit demand model part-worth coefficients for the discrete product characteristic levels are then fit to the data using maximum likelihood techniques, and splines are fit through these discrete levels. These splines, fixed during optimization, then act to calculate demand in the marketing subproblem. The marketing subproblem and engineering design subproblem are coordinated by passing target and response product characteristics back and forth until convergence.

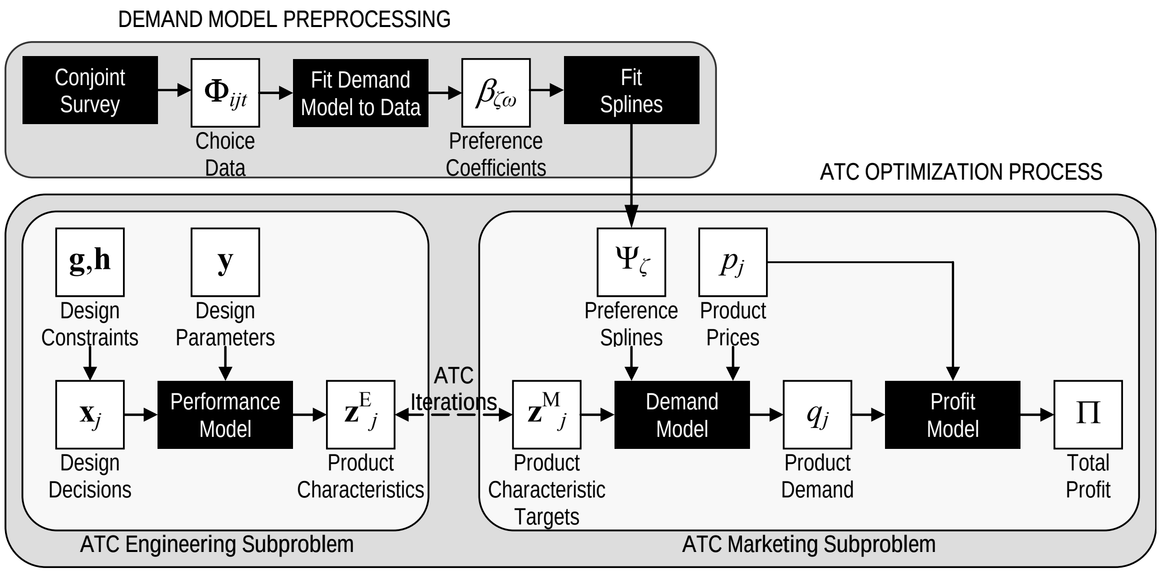


Figure 4.3: Diagram of the modeling process

Figure 4.4 depicts a schematic of the complete ATC formulation of the product development problem for a single-product-producing monopolist, conditional on the preference splines of the demand model. In this formulation there is only one product, so the product index j is dropped. In the product planning subproblem, price p and product characteristic targets $\mathbf{z}_M$ are chosen to maximize profit Π while minimizing the deviation between the product characteristic targets set by marketing $\mathbf{z}_M$ and those achieved by the engineering design $\mathbf{z}_E$ using weighting coefficients $\mathbf{w}$ to specify the tradeoff between the two objectives. Profit Π is calculated as revenue minus cost as in Eq.(4.2), and demand q is calculated using the logit model in Eq.(4.7)-(4.8) with known market potential s. In the engineering design subproblem, design variables $\mathbf{x}$ are chosen to minimize the deviation between characteristics achieved by the design $\mathbf{z}_E$ and targets set by marketing $\mathbf{z}_M$ using Eq.(4.1) subject to engineering constraints $\mathbf{g}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$. These two subproblems are solved iteratively, each using standard nonlinear programming techniques (Papalambros and Wilde, 2000) to solve each subproblem until the system converges. The weighting update method from Chapter 3 may then be used to find weighting coefficient values **w** that produce a solution satisfying user-specified tolerances for inconsistency between marketing and engineering for each term in **z**. This method is important for producing consistent solutions in cases where the top level subproblem does not have an attainable target: In this case profit is maximized rather than setting an attainable profit target.

Marketing Planning Subproblem

$$
\begin{aligned}
&\text{maximize } \Pi - \left\| \mathbf{w} \circ \left(\mathbf{z}_{\mathrm{M}} - \mathbf{z}_{\mathrm{E}} \right) \right\|_2^2 \\
&\text{with respect to } \mathbf{z}_{\mathrm{M}}, p \\
&\text{subject to } \mathbf{z}_{\min} \leq \mathbf{z}_{\mathrm{M}} \leq \mathbf{z}_{\max} \\
&\qquad\qquad p_{\min} \leq p \leq p_{\max} \\
&\text{where } \Pi = q\left(p - c_{\mathrm{V}} \right) - c_{\mathrm{I}} \\
&\qquad\quad q = s e^{v} \left(1 + e^{v} \right)^{-1} \\
&\qquad\quad v = \Psi_0\left(p \right) + \sum_{\zeta=1}^{Z} \Psi_\zeta \left(\langle \mathbf{z} \rangle_\zeta \right)
\end{aligned}
$$

$\mathbf{z}_{\mathrm{M}}$ ↓ ↑ $\mathbf{z}_{\mathrm{E}}$

Engineering Design Subproblem

$$
\begin{aligned}
&\text{minimize } \left\| \mathbf{w} \circ \left(\mathbf{z}_{\mathrm{M}} - \mathbf{z}_{\mathrm{E}} \right) \right\|_2^2 \\
&\text{with respect to } \mathbf{x} \\
&\text{subject to } \mathbf{g}\left(\mathbf{x} \right) \leq \mathbf{0},\ \mathbf{h}\left(\mathbf{x} \right) \leq \mathbf{0} \\
&\qquad\qquad \mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max} \\
&\text{where } \mathbf{z}_{\mathrm{M}} = \mathbf{r}\left(\mathbf{x} \right)
\end{aligned}
$$

Figure 4.4: ATC formulation of the product planning and engineering design problem

4.4 EMPIRICAL APPLICATION: DIAL-READOUT SCALES

The potential of the joint marketing and engineering design model is demonstrated in the design of a standard household dial-readout bathroom scale. Scales possess a number of attractive features for illustrative purposes: Consumers are nearly uniformly familiar with them; the 'range' of bathroom scales in the marketplace is relatively small, even among inexpensive consumer durables, and as such would have to be considered moderately differentiated at most; even the best scales are not very costly, so one could potentially measure price effects well; the number of characteristics consumers value is not large, the number of 'levels' (part-worths) within each characteristic is reasonable, and characteristics and levels can be known in advance through prior studies and on-line data; finally, lack of mechanical complexity makes it possible to formulate a small, explicit set of geometric and physical constraints for the engineering design subproblem. These simplifications are convenient, but the methodology presented here can be used for most durable products, even if certain modeling aspects may vary considerably (in terms of arduousness) among products.

4.4.1 Marketing Planning Subproblem

Marketers must first identify which product characteristics under their control are of interest to consumers, and which levels they can distinguish. A great deal of information on bathroom scales was made available for this work in a proprietary report indicating which characteristics figured high in consumer preferences. Some, like "color," could be interchanged or manipulated on the fly without interaction with other scale components, and were thus left out of the experimental design. It was also considered which characteristics would be especially important to convey in an on-line purchase environment, the environment the experiment was meant to loosely simulate, given that the study was itself conducted on the web. Finally, it is important that the chosen characteristics can be directly quantified and easily conveyed to respondents in an unambiguous manner; thus, nebulous descriptors such as “nicely proportioned” were eschewed in favor of actual proportions. Five product characteristics – weight capacity, aspect ratio, platform area, tick mark gap and number size, plus price – were adopted because of their relevance to consumers and designers, as well as their prevalence in on-line purchase descriptions and pictures; these appear, along with the levels for each, in Table 4.1. Each characteristic was discretized into five levels, which allowed adequate spline interpolation, as shown in Figure 4.5. These levels were chosen to span the range of values of products in the market based on a sample of 32 different scales sold on the internet, to ensure realism and to capture realistic anticipated trade-offs. Characteristics such as brand name were deliberately avoided because of great differential familiarity and lack of a direct tie-in with the design of the underlying product.

Table 4.1: Product characteristic and price levels

k	Description	Metric	Units	Levels				
z_1	Weight capacity	Weight causing a 360° dial turn	lbs	200	250	300	350	400
z_2	Aspect ratio	Platform length divided by width	--	6/8	7/8	8/8	8/7	8/6
z_3	Platform area	Platform length times width	in^2	100	110	120	130	140
z_4	Tick mark gap	Distance between 1-lb tick marks	in.	2/32	3/32	4/32	5/32	6/32
z_5	Number size	Length of readout number	in.	0.75	1.00	1.25	1.50	1.75
p	Price	US Dollars	$	10	15	20	25	30

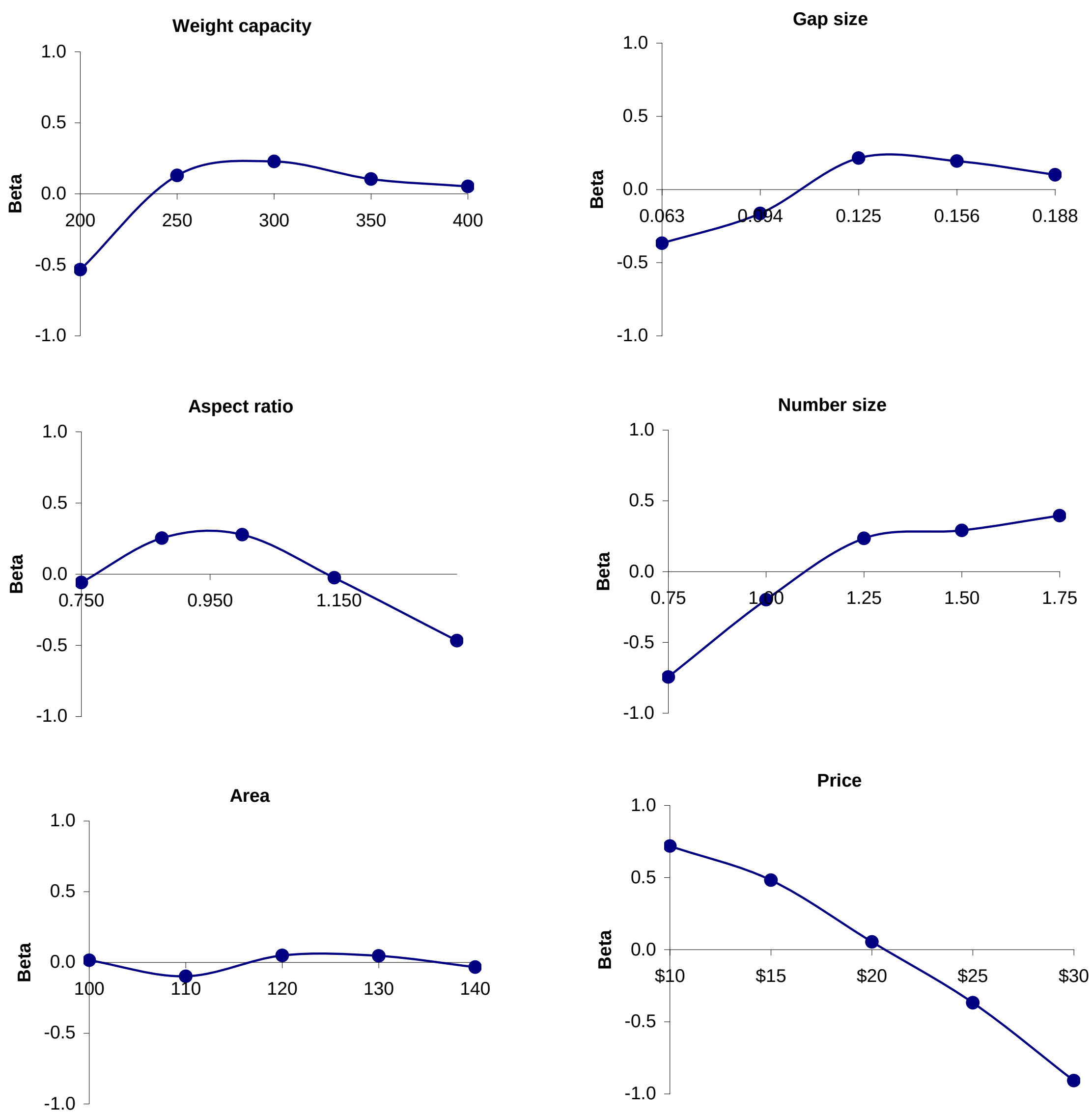


Figure 4.5: Interpolated spline functions

An efficient choice-based conjoint design (50 sets of three-products, plus "no choice") was used and implemented on the Internet. Respondents were solicited through announcements on numerous internet newsgroups as well as through two classes at the University of Michigan, one in marketing research, the other in engineering design. All respondents[2], 184 in total, were offered incentives in the form of sweepstakes for gift certificates in various amounts. A great deal of effort was put into having the choice task correspond to the sort found at on-line shopping sites. To that end, scales were presented in terms of their underlying product characteristic information in list format and pictorially, including a close-up of the dial to facilitate comparison across the last two characteristics; a screen capture is provided in Figure 4.6.

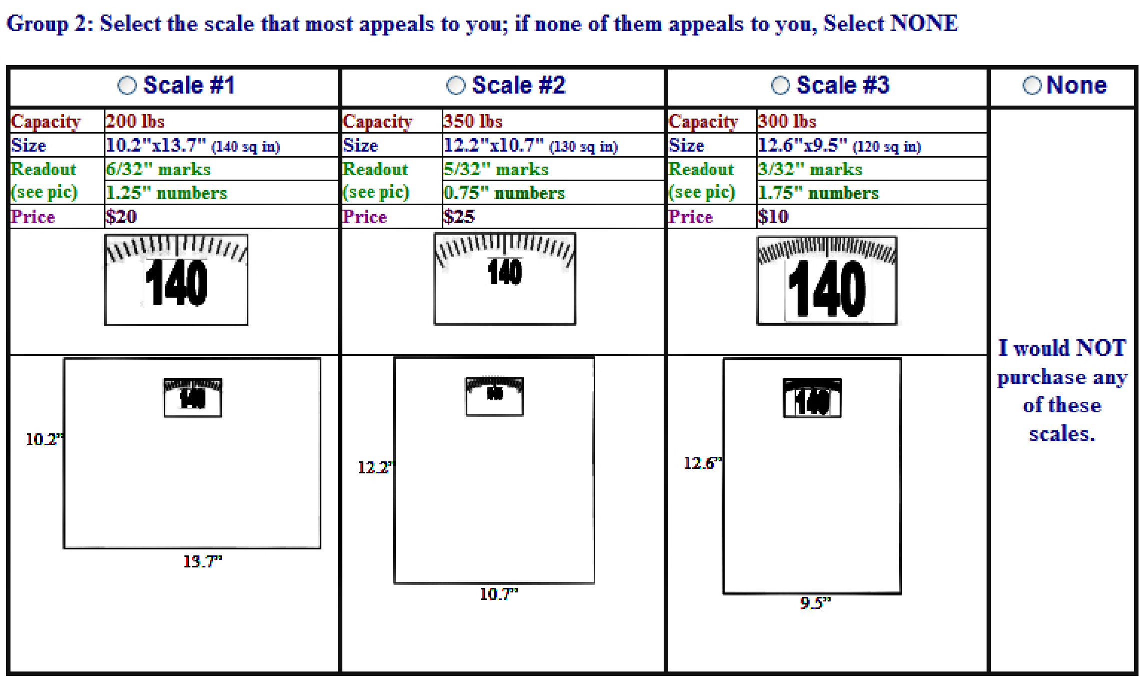


Figure 4.6: Online conjoint scale choice task

[2] Demographic data were also solicited from respondents. The sample is 58% male. For men, mean height, weight and age are 70.7 inches, 177 lbs. and 28.2, respectively; corresponding values for women were 64.6 inches, 129 lbs. and 26.5. Respondents were also asked three questions relevant to scale purchase behavior: (1) Do you need vision correction to see clearly at a distance of 6 feet?; (2) Have you tried (deliberately) to lose at least ten pounds in the last year?; (3) Have you purchased a scale in the past two years? Female and male affirmative proportions were {0.49, 0.40, 0.22} and {0.48, 0.36, 0.22}, respectively, and were not statistically distinguishable. No significant systematic relationships between these variables and preference patterns in the conjoint task were noted.

As dictated by the conjoint design, options involving physically or geometrically infeasible product characteristic level combinations were included, because responses were used to measure consumer value systems (part-worth utilities) and trade-offs, *not* to be restricted to feasible designs in the engineering design subproblem. Not requiring such a feasible set to be explicitly delineated in advance is among the main strengths of the ATC approach.

Model parameters were estimated, as described earlier, using maximum likelihood and a Newton type algorithm[3]. The resulting β values are listed in Table 4.2; these values have been scaled so that the mean in each set of characteristics is the same. The average β value for each characteristic is -0.004, corresponding to the relative attractiveness of scales with respect to the "no choice" option, which for identification purposes has $v_0 = 0$ identically. These values show reasonable trends: Respondents monotonically prefer larger numbers and lower cost, but have interior preferences for weight capacity, platform area, shape (aspect ratio) and interval mark gap. One might argue that more weight capacity is always better, and so body weight data was collected on participants; the heaviest was 280 lbs., so none would in fact have required either of the two highest capacity levels (so the mild utility decline may be attributable to wanting to avoid excess capacity or even not wishing to appear as if it were needed). Natural cubic spline functions Ψ, shown in Figure 4.5, were fit to these β values for each characteristic and for price. Based on discussions with a major scale manufacturer, (exogenous) values for cost c_V = \$3 per unit and for initial investment c_I = \$1 million

[3] Estimation for the conjoint model was based on maximum likelihood using standard gradient search methods; all starting values converged to identical optima. At the optimum, the log-likelihood, LL = -10983. This model can be compared to a series of nested alternatives: to a seven-parameter model which sets equal levels within characteristics, but allows the characteristics themselves to vary (LL = -12066); to a one-parameter model which estimates only the "no choice" option's relative attractiveness (LL = -12716); and to a 'zero parameter' model which assigns equal probability to all choices (LL = -12753). Each can be very strongly rejected against the preceding one.

were assumed, as well as for market size *S*, which was set to 5 million, the approximate yearly market for dial scales in the United States. Using this last figure and the estimated splines, demand *q* was computed using Eq.(4.7)-(4.8).

Table 4.2: Part-worth coefficient beta values

Weight Capacity		**Platform Area**		**Size of Number**	
200 lbs.	-0.534	100 in.^{2}	0.015	0.75 in.	-0.744
250 lbs.	0.129	110 in.^{2}	-0.098	1.00 in.	-0.198
300 lbs.	0.228	120 in.^{2}	0.049	1.25 in.	0.235
350 lbs.	0.104	130 in.^{2}	0.047	1.50 in.	0.291
400 lbs.	0.052	140 in.^{2}	-0.033	1.75 in.	0.396
Platform Aspect Ratio		**Interval Mark Gap**		**Price**	
0.75	-0.058	2/32 in.	-0.366	\$10	0.719
0.88	0.253	3/32 in.	-0.164	\$15	0.482
1.00	0.278	4/32 in.	0.215	\$20	0.054
1.14	-0.025	5/32 in.	0.194	\$25	-0.368
1.33	-0.467	6/32 in.	0.100	\$30	-0.908

4.4.2 Engineering Design Subproblem

Reverse engineering was used to create the engineering design submodel. Three scales of different construction were purchased and disassembled; this allowed a determination of the relevant functional components and their dependencies and interrelations. These are shown in Figure 4.7, with the resultant design variables for the engineering design submodel.

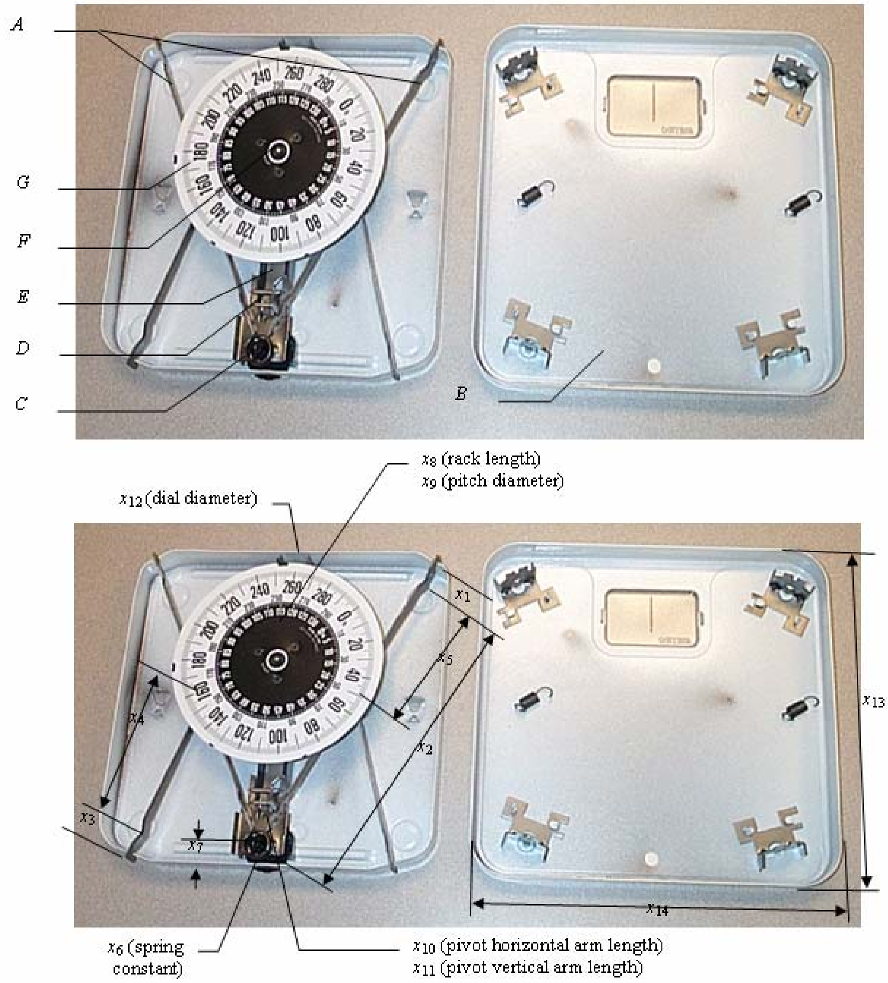


Figure 4.7: Disassembled scale showing components and design variables

Analysis of the three different scales indicated they operated on essentially identical principles. In Figure 4.7, levers *A* create mechanical advantage and translate the force of the user's weight from the cover *B* to coil spring *C* which is displaced proportionally to the applied force; a pivot lever *D* transfers the vertical motion of the spring to the horizontal motion of gear rack *E*, after which pinion gear *F* translates the rack's linear motion to rotation of the dial *G*. Although this basic topology is common to the three scales examined, dimensions vary; for example, the ratio of dial-turn per applied force depends on the dimensions of the levers, the rack and pinion, and the spring properties. Because the topology is common, it is possible to represent a parametric space of design alternatives using a set of design variables. Figure 4.7 shows the set of fourteen design variables chosen for this study, all of which are real-valued, positive, and continuous in nature. Other dimensions were considered to be fixed parameters **y** with values based on the observed scales, as shown in Table 4.3.

Table 4.3: Engineering design model parameters

Name	Description	Value	Units
y_1	Gap between base and cover	0.30	in
y_2	Minimum distance between spring and base	0.50	in
y_3	Internal thickness of scale	1.90	in
y_4	Minimum pinion pitch diameter	0.25	in
y_5	Length of window	3.00	in
y_6	Width of window	2.00	in
y_7	Distance between top of cover and window	1.13	in
y_8	Number of lbs measured per tick mark	1.00	lbs
y_9	Horizontal distance between spring and pivot	1.10	in
y_{10}	Length of tick mark + gap to number	0.31	in
y_{11}	Number of lbs that number length spans	16.00	lbs
y_{12}	Aspect ratio of number (length/width)	1.29	-
y_{13}	Min. allowable dist. of lever at base to centerline	4.00	in

Eight mathematical constraint functions **g**(**x**) were developed based on geometric and mechanical relationships to ensure that the design variable vector **x** represents a meaningful, feasible design. First, the dial (G) diameter x_{12} must be small enough to fit inside the base widthwise, where the base width is measured as the cover width x_{14} minus the gap y_1 between the cover and the base on both sides:

$$x_{12} \le x_{14} - 2y_1 \; . \tag{4.9}$$

The dial G must also fit lengthwise inside the scale base (x_{13}-2y_1) with sufficient room for the spring plate (x_7+y_9):

$$x_{12} \le x_{13} - 2y_1 - x_7 - y_9 \,. \tag{4.10}$$

The length of the short levers (x_3+x_4) must be small enough to fit inside the base lengthwise (x_{13}-2y_1):

$$\left(x_3 + x_4\right) \le x_{13} - 2y_1 \,. \tag{4.11}$$

The position along the long lever of the short lever joint x_5 must be within the bounds of the long lever length x_2:

$$x_5 \le x_2 \,. \tag{4.12}$$

In the fully extended position, the end of the rack E (x_7+y_9+x_{11}+x_8) must fit inside the scale body lengthwise (x_{13}-2y_1):

$$x_7 + y_9 + x_{11} + x_8 \le x_{13} - 2y_1 \,. \tag{4.13}$$

However, the length x_8 of the rack E must be long enough to span the space between the pivot lever D and the pinion F:

$$x_8 \geq (x_{13} - 2y_1) - \left(\frac{x_{12}}{2} + y_7\right) - x_7 - y_9 - x_{10} \,. \tag{4.14}$$

The two long levers connect to the top edge of the base rather than the side. Therefore the lever length (x_1+x_2) is limited by the width dimension of the scale body (x_{14}-2y_1). Using the Pythagorean Theorem:

$$(x_1 + x_2)^2 \leq (x_{13} - 2y_1 - x_7)^2 + \left(\frac{x_{14} - 2y_1}{2}\right)^2 . \tag{4.15}$$

However, for stability the long levers must be long enough (x_1+x_2) to attach to the top edge of the base at a minimum distance y_{13} from the centerline. Again, using the Pythagorean Theorem:

$$\left(x_1 + x_2\right)^2 \geq \left(x_{13} - 2y_1 - x_7\right)^2 + y_{13}^2 \,. \tag{4.16}$$

In addition, simple bounds are provided to ensure that all variables are positive. Given that all x are positive, any real-valued vector $\mathbf{x}$ that satisfies Eq.(4.10)-(4.16) represents a valid, feasible design.

Next, the response functions $\mathbf{r}(\mathbf{x})$ that calculate product characteristics $\mathbf{z}$ in terms of the design variables $\mathbf{x}$ are defined. Assuming the scale is made up of rigid bodies (except for the spring) and using standard static force and moment balancing (Hibbeler, 1993), the weight capacity z_1 can be derived as a function of the position of the cover force on the long (x_1) and short (x_3) levers, the length of the long (x_1+x_2) and short (x_3+x_4) levers, the position of the joint x_5, the dimensions of the pivot (x_{10} and x_{11}), the pitch diameter of the pinion x_9 and the spring constant x_6:

$$z_1 = \frac{4\pi x_6 x_9 x_{10}\left(x_1 + x_2\right)\left(x_3 + x_4\right)}{x_{11}\left(x_1\left(x_3 + x_4\right) + x_3\left(x_1 + x_5\right)\right)}. \tag{4.17}$$

The aspect ratio is the length of the cover divided by its width:

$$z_2 = \frac{x_{13}}{x_{14}}. \tag{4.18}$$

The area of the scale cover is its length times its width:

$$z_3 = x_{13}x_{14}. \tag{4.19}$$

The arc length of the gap between 1-lb interval tick marks is proportional to the dial diameter x_{12} and inversely proportional to the weight capacity z_1 (see Eq.(4.17)):

$$z_4 = \pi\frac{x_{12}}{z_1}. \tag{4.20}$$

Finally, the number length, a measure of overall printed number size, is calculated in terms of the dial diameter x_{12} and weight capacity z_1 using trigonometry based on the fixed span of numbers along the tick marks y_{10} (the printed number is assumed to span a fixed number of tick marks), the positioning of the numbers on the dial y_{11}, and the aspect ratio (length / width) of the rectangular space allocated for the number y_{12}:

$$z_5 = \frac{\left(2\tan\left(\frac{\pi y_{11}}{z_1}\right)\right)\left(\frac{x_{12}}{2} - y_{10}\right)}{\left(1 + \frac{2}{y_{12}}\tan\left(\frac{\pi y_{11}}{z_1}\right)\right)}. \tag{4.21}$$

Equations (4.17)-(4.21) form the vector function $\mathbf{r}(\mathbf{x})$, which maps design variables $\mathbf{x}$ onto product characteristics $\mathbf{z}$ so that product characteristics can be calculated for any design. The design variables $\mathbf{x}$ are constrained to feasible values by the constraint functions $\mathbf{g}(\mathbf{x})$; therefore, the resultant product characteristic combinations $\mathbf{z} = \mathbf{r}(\mathbf{x})$ are also restricted to feasible combinations.

4.5 RESULTS

The engineering design and marketing subproblems were solved iteratively until convergence using the Matlab function *fmincon*, based on the sequential quadratic programming method (Papalambros and Wilde, 2000), to solve each subproblem. This gradient-based search algorithm generates local optima, and global optima can only be found through multi-start. The reported solution represents the best local optimizer found over several starting points based on the dimensions of scales used for reverse engineering[4]. At the solution, shown in Table 4.4 the optimal scale design is bounded by

---

[4] The marketing and engineering design subproblems were each solved using the Matlab 6.5.1 function *fmincon*, based on SQP (Papalambros and Wilde, 2000), with default parameter settings. Convergence of the ATC subproblem coordination was strictly defined as occurring when each subproblem is unable to improve its objective function value over the optimal solution from the previous iteration. The number of ATC coordination iterations required to converge varies depending on the starting point and weighting coefficients used. Using the starting point generated by the disjoint case with weighting coefficients of 10^5, the system converged in 1815 ATC iterations, and the resulting inconsistency between marketing targets and engineering design characteristics are less than 0.3% for all characteristics. Use of smaller weighting coefficients yields faster convergence but greater inconsistency between marketing targets and engineering design characteristics, as discussed in Chapter 3. For example, weighting coefficients of 10^4 yield convergence in only 31 ATC iterations with inconsistencies less than 10% for all characteristics.

active engineering constraints that ensure the dial, the spring plate, and the levers are not too large to fit inside the scale. The optimal scale characteristics are within the range of scales found in online e-commerce, and none of the variable bounds are active except for x_7, which is unique because it is explicitly bound by the specified parameter value of y_2, rather than an arbitrary bound.

Table 4.4: Optimal scale design

	Variable	Description	Value	Lower Bound	Upper Bound
Marketing Variables	z_1	Weight capacity	254 lbs.	200 lbs.	400 lbs.
	z_2	Aspect ratio	0.997	0.75	1.33
	z_3	Platform area	133 in^2	100 in^2	140 in^2
	z_4	Tick mark gap	0.116 in.	1/16 in.	3/16 in.
	z_5	Number size	1.33 in.	0.75 in.	1.75
	p	Price	$26.41	$10.00	$30.00
Engineering Variables	x_1	Length from base to force on long lever	2.30 in.	0.125 in.	36 in.
	x_2	Length from force to spring on long lever	8.87 in.	0.125 in.	36 in.
	x_3	Length from base to force on short lever	1.34 in.	0.125 in.	24 in.
	x_4	Length from force to joint on short lever	1.75 in.	0.125 in.	24 in.
	x_5	Length from force to joint on long lever	0.41 in.	0.125 in.	36 in.
	x_6	Spring constant	95.7 lb./in.	1.00 lb./in.	200 lb./in.
	x_7	Distance from base edge to spring	0.50 in.	0.50 in.	12 in.
	x_8	Length of rack	7.44 in.	1.00 in.	36 in,
	x_9	Pitch diameter of pinion	0.25 in.	0.25 in.	24 in.
	x_{10}	Length of pivot's horizontal arm	0.50 in.	0.50 in.	1.9 in.
	x_{11}	Length of pivot's vertical arm	1.90 in.	0.50 in.	1.9 in.
	x_{12}	Dial diameter	9.34 in.	1.00 in.	36 in.
	x_{13}	Cover length	11.54 in.	1.00 in.	36 in.
	x_{14}	Cover width	11.57 in.	1.00 in.	36 in.

In the engineering model several product characteristics are functions of the ratios of some of the design variables. For example, an increase in lever length can be traded off for a changed spring constant, force placement, pinion gear pitch diameter, or pivot lever dimensions to yield an equivalent weight capacity. This means that two different designs with appropriate design variable ratios may exhibit the same product characteristics, and also that an infinite number of design solutions are equivalent from a marketing perspective. One such design is reported in Table 4.4. Additional models representing cost structures in terms of design variables or part commonality among product variants in a product line could be used to select a single design among the set of otherwise equivalent designs; however, this possibility was not explored here.

4.5.1 Comparison of ATC with Disjoint Decision-Making

One might question whether the joint method proposed here has a substantial impact on product design and, ultimately, resulting profit. The role of ATC in avoiding infeasible products has been emphasized; however, this example demonstrates the impact that ATC can have on profitability. Let us examine a case of disjoint decision-making by marketing and engineering design, similar to the methodology proposed by Cooper *et al*. (2003), where (1) marketing defines desired product characteristics, (2) engineering designs a feasible product to meet the requested characteristics as closely as possible, and (3) marketing prices the actualized product.

In the first step, marketing chooses the optimal price and product characteristic combination conditional on the monopolist / single-product framework and known consumer preference data (arrived at using conjoint, a discrete choice model, and the profit function). This step is referred to as Analytical Target Setting (Cooper *et al*., 2003). Based on the optimal price and characteristics at this stage, expected price, market share and profit are \$28.04, 64.3% and \$79.5M, respectively as shown in Table 4.5. There is no guarantee that a feasible product can be designed that exhibits the desired target characteristics. So, engineering designs a *feasible* product that meets the product characteristics requested by marketing as closely as possible. At this point the product design is considered fixed, but price is an easily changed variable, so it can be reconsidered based on the characteristics of the achieved design (a simple form of Cooper *et al.*'s "Reduced ATS" problem). The resulting price, share and profit in this scenario are \$25.54, 54.8% and \$60.8M as shown in Table 4.5. This involves a sizeable decrease in price and share, and a truly enormous drop in profit from what marketing had originally planned using target product characteristics. Consumers often desire combinations of product characteristic values that are difficult or impossible to produce together, so it is important to examine the *realizable* share and profit levels. In a scenario such as this, marketing may accuse engineering design of failing to deliver, while engineering design may blame marketing for requesting a product that could not be built, causing substantial unnecessary compromises in the final product design. With contingent, sequential decision-making each side would be in the right from its own perspective, but the final decisions would be inferior.

For comparison, let these two groups use ATC as a tool for communication, considering both the tradeoffs among desirable product characteristics and the feasibility of obtaining these characteristics. In this case, an entirely different product is designed, with price, market share and profit of \$26.41, 59.0% and \$68.0M, as shown in Table 4.5. Although the price is not much higher than in the disjoint case above, share and profit are significantly improved. This difference in profitability is non-trivial, approximately \$7,200,000, a 12% increase over the 'best feasible' design offered by engineering design based on 'optimal' marketing target specifications alone. Thus, ATC, using the same submodels, converges to a jointly optimal solution offering far better market prospects.

Table 4.5: Comparison of disjoint solution vs. ATC-coordinated solution

			Disjoint		Joint
	Description	**Unit**	**Initial Marketing Plan**	**Final Product Design**	**ATC**
z_1	Weight capacity	lbs	283	222	254
z_2	Aspect ratio	--	0.946	1.041	0.997
z_3	Platform area	in.^{2}	124.2	127.8	133.4
z_4	Tick mark gap	in.	0.136	0.1322	0.116
z_5	Number size	in.	1.75	1.478	1.33
p	Price	$	$28.04	$25.54	$26.41
P_j	Market Share	%	64.3%	54.8%	59.0%
Π	Profit	$	$79.5 M	$60.8 M	$68.0 M

4.6 CONCLUSIONS

From the perspective of the producer, marketing and engineering design ideally work together to achieve a common goal: creating the product with greatest value for the firm. As detailed here and in earlier cited research, goals, language and modes of operation in the two disciplines tend to insulate each from the other. The practical upshot is that each tends to solve problems relative to constraints 'exogenously' set by the other. The proposed methodology, based on analytical target cascading, allows the disciplines to remain largely independent yet to link their product subproblems formally, using time-tested models from both fields.

It is instructive to consider what this joint methodology offers each of the constituent communities. For the marketing community, ATC goes beyond merely facilitating communication and cutting down time-consuming iterations; it helps whenever marketers confront even moderately complex products and/or production processes in which some combinations of desired characteristics are technologically impractical or even physically impossible. This 'feasible set' of products is seldom one which can be easily described in the product characteristic space, and is ordinarily a function of the technical decisions of the product. In short, the method allows marketers to dispense, at least initially, with questions of "what can be made?" and focus instead on what they do best: discerning what consumers most value.

The method offers distinct benefits for the engineering design community, helping to contextualize design decisions within the larger framework of the firm and how it satisfies its customer base. Instead of resolving engineering tradeoffs – for example, among competing performance objectives, as in multiobjective optimization – purely in terms of technological or physical possibility, it allows such decisions to be tied directly into the firm's overall objective, that of producing a successful and profitable product. The proposed methods would allow, for example, sensitivity analysis, where small design changes could be mapped to their eventual profit implications. Such an analysis would be unthinkable without a conjoined system of consumer needs and resultant demand, as provided by the marketing sub-model and linked through ATC.

This ATC coordination methodology can be readily extended to far greater complexity using known methods. For example, the consumer response model was made as simple as possible, based on a homogenous-coefficient logit model. Well-known hierarchical Bayes methods could be substituted to allow inference for heterogeneous populations, and probit models with full error covariance could help account for potential IIA problems (Kahneman and Tversky 1979), albeit at great loss of tractability. In turn, models allowing heterogeneous preferences (and thus demand) would allow one to design product *lines*. This extension is explored in detail in Chapter 5. It would even be possible to improve extant conjoint methods by allowing them to generate only *feasible* tasks: those which maximize utility measurement accuracy within the range of technologically possible product configurations.

On the engineering side, a great emphasis was placed on an overarching engineering design sub-model, which was based on a single product topology appropriate for rectangular dial-readout scales. Product variety could be enhanced by incorporating multiple product topologies, with the potential for automatic topology generation (as in, for example, Campbell *et al*., 1998). A deliberately simple cost model was chosen for illustration; however, more detailed cost models can be integrated to the engineering decision-model. By doing so, it would be possible to have another sort of feedback, wherein the marketing sub-model sets target production cost and the engineering design sub-model designs feasible products that meet cost targets. This extension is explored in detail in Chapter 6. Product lines or families can be accommodated on the engineering side as well, enabling study of component- and process-sharing effects on the production cost structure (Fellini *et al*. 2003) or the use of flexible and reconfigurable manufacturing equipment (Koren *et. al.*, 1999).

In closing, there are several points to stress for *both* communities. The first involves the viewpoint, common in marketing, that design constraints can generally be overcome by allotting appropriate funds. In some cases they cannot. Marketing methods must learn to take note not only of costly designs, but also of utterly infeasible ones, a concept foundational in the ATC formulation presented here. This can only improve predictive accuracy, and simultaneously reduce data requirements for the dominant models used in new product forecasting. In parallel, the engineering design community must accept that price and consumer preferences are aspects of design just as real as those determined by physics. Second, determining which product characteristic combinations are infeasible can be difficult even when producing only a single product as simple as the scales considered in this chapter. Even if infeasible combinations are eliminated in conjoint questions, optimal solutions may still be infeasible; this is particularly important for continuous variable formulations. The ATC approach allows marketing and engineering design each to formulate its own sub-model, using methods most familiar to each, and to link them afterward, so that an optimal joint decision can be reached. Finally, designs reached using ATC necessarily converge on joint optimality and, as such, guarantee better profitability – or any other chosen metric – than the suboptimal solutions achieved by solving the engineering design and marketing design problems sequentially. Given its relative ease of implementation, this last benefit will likely prove a deciding factor in the willingness of firms to adopt ATC processes for complex design projects

4.7 REFERENCES

[4.1] Andrews, R.L., A. Ainslie and I.S. Currim (2002) "An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity," *Journal of Marketing Research*, v39 p479-487.

[4.2] Andrews, R.L., A. Ansari and I.S. Currim (2002) "Hierarchical Bayes versus finite mixture conjoint analysis models: a comparison of fit, prediction, and partworth recovery," *Journal of Marketing Research*, v39 n1 p87-98.

[4.3] Balling, R.J. and Sobieszczanski-Sobieski (1994) "Optimization of coupled systems: a critical overview of approaches," *Proceedings of the 5th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization,* AIAA-94-4330-CP, Panama City, p753-773.

[4.4] Braun, R. (1996) *Collaborative Optimization: An Architecture for Large-Scale Distributed Design*, PhD Dissertation, Stanford University.

[4.5] Brown, Shona L. and Kathleen M. Eisenhardt (1995) "Product development: past research, present findings, and future directions," *Academy of Management Review*, v20 (April) p343-378.

[4.6] Campbell, M., K. Kotovsky and J. Cagan (1998) "Agent-based synthesis of electro-mechanical design configurations," *Proceedings of the 1998 ASME Design Engineering Technical Conferences*, DTM-5673, September.

[4.7] Cattin, Philippe and Dick R. Wittink (1982) "Commercial use of conjoint analysis: a survey," *Journal of Marketing*, v46 (Summer) p44-53.

[4.8] Choudhary, R., A. Malkawi and P.Y. Papalambros (2003) "A hierarchical design optimization framework for building performance analysis" *Proceedings of the Eight International Building Performance Simulation Association (IBPSA) Conference*, Eindhoven, Netherlands, August 11-14.

[4.9] Cooper, A.B., P. Georgiopoulos, H.M. Kim and P.Y. Papalambros (2003) "Analytical target setting: an enterprise context in optimal product design," *Proceedings of the 2003 ASME Design Automation Conference*, DAC-48734, Chicago, Illinois, USA, September 2-6.

[4.10] Diaz, A. (1987) “Interactive solution to multiobjective optimization problems,” *International Journal for Numerical Methods in Engineering*, v24 p1865-1877.

[4.11] Fellini, R., M. Kokkolaras and P.Y. Papalambros (2003) “A rigorous framework for making commonality and modularity decisions in optimal design of product families,” *Proceedings of the International Conference on Engineering Design*, Stockholm, Sweden, August 19-21.

[4.12] Garcia, R. and R. Calantone (2002) “A critical look at technological innovation typology and innovativeness terminology: a literature review,” *Journal Of Product Innovation Management*, v19 n2 p110-132.

[4.13] Georgiopoulos, P., M. Jonsson and P.Y. Papalambros (2004) “Linking optimal design decisions to the theory of the firm: the case of resource allocation,” *ASME Journal of Mechanical Design*, (forthcoming).

[4.14] Georgiopoulos, P. (2003) *Enterprise-wide Product Design: Linking Optimal Design Decisions to the Theory of the Firm*, D. Eng. Thesis, University of Michigan, Ann Arbor.

[4.15] Green, Paul E. and V. Srinivasan (1990) “Conjoint analysis in marketing: new developments with implications for research and practice,” *Journal of Marketing*, v54 n4 p3-19.

[4.16] Griffin, Abbie (1997) “The effect of project and process characteristics on product development cycle time,” *Journal of Marketing Research*, v34 (February) p24-35.

[4.17] Griffin, Abbie and John R. Hauser (1992) “Patterns of communication among marketing, engineering and manufacturing – a comparison between two new product teams,” *Management Science*, v38 (March) p360-373.

[4.18] Gu, X., J.E. Renaud, L.M. Ashe, S.M. Batill, A.S. Budhiraja, and L.J.Krajewski (2002) “Decision-based collaborative optimization,” *Journal of Mechanical Design*, v124 p1-13.

[4.19] Gupta A. K., S. P. Raj and D. Wilemon (1986) “A model for studying research-and-development - marketing interface in the product innovation process,” *Journal of Marketing*, v50 (April) p7-17.

[4.20] Gupta, S.K. and A.K. Samuel (2001) “Integrating market research with the product development process: a step towards design for profit,” *Proceedings of the ASME Design Engineering Technical Conferences*, Pittsburgh, PA. September.

[4.21] Haaijer, M. E. (1999) “Modeling conjoint choice experiments with the probit model,” Labyrint Publications, <http://www.ub.rug.nl/eldoc/dis/eco/m.e.haaijer/>.

[4.22] Halman, Johannes I. M., Adrain P. Hofer and Wim van Wuuren (2003) “Platform-driven development of product families: linking theory with practice,” *Journal Of Product Innovation Management*, v20 n2 p149-162.

[4.23] Hauser, John R. (2001) “Metrics Thermostat,” *Journal of Product Innovation Management*, v18 n3 p134-153.

[4.24] Hazelrigg, G.A. (1988) “A framework for decision-based engineering design” *ASME Journal of Mechanical Design*, v120 p653-658.

[4.25] Henard, David H. and David M. Szymanski (2001) “Why some new products are more successful than others,” *Journal of Marketing Research*, v38 (August) p362-375.

[4.26] Hibbeler, R.C. (1993) *Statics and Mechanics of Materials*, Macmillan Publishing, New York.

[4.27] Hultink, Erik Jan, Susan Hart, Henry S.J. Robben and Abbie Griffin (2000) “Launch decisions and new product success: an empirical comparison of consumer and industrial products,” *Journal of Product Innovation Management*, v17 n1 p5-23.

[4.28] Kahn, K.B. (2002) “An exploratory investigation of new product forecasting practices,” *Journal of Product Innovation Management*, v19 n2 p133-143.

[4.29] Kahneman, Daniel and Amos Tversky (1979) “Prospect theory: an analysis of decision under risk,” *Econometrica*, v47 n2 p263-91.

[4.30] Kalyanam, K. and T.S. Shively (1998) “Estimating irregular pricing effects: a stochastic spline approach,” *Journal of Marketing Research*, v35 p16-29.

[4.31] Kaul, Anil and Vithala Rao (1995) “Research for product positioning and design decisions: an integrative review,” *International Journal of Research in Marketing*, v12 p293-320.

[4.32] Kim, H. M. (2001) *Target Cascading in Optimal System Design*, Ph.D. Dissertation, Dept. of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan.

[4.33] Kim, H. M., M. Kokkolaras, L. Louca, G. Delagrammatikas, N. Michelena, Z. Filipi, P.Y. Papalambros, and D. Assanis (2002) “Target cascading in vehicle redesign: a class VI truck study,” *International Journal of Vehicle Design*, v29 n3 p199-225.

[4.34] Kim, H.M., N.F. Michelena, P.Y. Papalambros, and T. Jiang (2003a) “Target cascading in optimal system design,” *ASME Journal of Mechanical Design*, v125 p481-489.

[4.35] Kim, H.M., D.G. Rideout, P.Y. Papalambros, and J.L. Stein (2003b) "Analytical target cascading in automotive design," *ASME Journal of Mechanical Design*, v125 p481-489.

[4.36] Kim, J.G., U. Menzefricke and F.M. Feinberg (2003c) "Capturing heterogeneous nonlinear utilities: a bayesian spline approach," working paper, University of Michigan Business School.

[4.37] Kokkolaras, M., R. Fellini, H.M. Kim, N. Michelena, and P. Papalambros (2002) "Extension of the target cascading formulation to the design of product families," *Journal of Structural and Multidisciplinary Optimization*, v14 n4 p293-301.

[4.38] Koren, Y., U. Heisel, F. Jovane, T. Moriwaki, G. Pritschow, G. Ulsoy and H. Van Brussel (1999) "Reconfigurable manufacturing systems," a keynote paper, *Annals of the CIRP*, v48 n2 p527-540.

[4.39] Krishnan, V. and K. Ulrich (2001) "Product development decisions: a review of the literature," *Management Science*, v47 n1 p1-21.

[4.40] Kuhfeld, Warren F. (2003) "Conjoint analysis," SAS Institute Publications, available at *http://support.sas.com/techsup/tnote/tnote_stat.html*.

[4.41] Leenders, Mark A.A.M. and Berend Wierenga (2002) "The effectiveness of different mechanisms for integrating marketing and R&D," *Journal Of Product Innovation Management*, v19 n4 p305-317.

[4.42] Li, H. and S. Azarm (2000) "Product design selection under uncertainty and with competitive advantage," *ASME Journal of Mechanical Design*, v122 p411-418.

[4.43] Lilien, Kotler and Moorthy (1992) *Marketing Models*, Chapter 10, Englewood Cliffs: Prentice-Hall.

[4.44] Louviere J. J. and G. Woodworth (1983) "Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data," *Journal of Marketing Research*, v20 p350-367.

[4.45] Louviere, J.J. (1988) *Analyzing decision making: Metric conjoint analysis*, Newbury Park, CA: Sage Publications, Inc.

[4.46] Louviere, J., D. Hensher, and J. Swait (2000) *Stated Choice Methods: Analysis and Application*, Cambridge University Press, UK.

[4.47] McAlister, Leigh (1982), "A dynamic attribute satiation model of variety-seeking behavior," *Journal of Consumer Research*, v9 (September) p141-150.

[4.48] Michalek, J. and P.Y. Papalambros (2005a) "An efficient weighting update method to achieve acceptable consistency deviation in analytical target cascading," *ASME Journal of Mechanical Design*, (forthcoming).

[4.49] Michalek, J. and P.Y. Papalambros (2005b) "Technical brief: weights, norms, and notation in analytical target cascading," *ASME Journal of Mechanical Design*, (forthcoming).

[4.50] Michalek, J.J, F.M. Feinberg and P.Y. Papalambros (2004), "An optimal marketing and engineering design model for product development using analytical target cascading," *Proceedings of the Tools and Methods of Competitive Engineering Conference*, Lausanne, Switzerland. April.

[4.51] Michalek, J.J., F.M. Feinberg and P.Y. Papalambros (2005), "Linking marketing and engineering product design decisions via analytical target cascading," *Journal of Product Innovation Management*, v22 p42-62.

[4.52] Michalek, J., P.Y. Papalambros and S.J. Skerlos (2005) "A study of fuel efficiency and emission policy impact on optimal vehicle design decisions," *ASME Journal of Mechanical Design*, (forthcoming)

[4.53] Michelena, N., H. Park, and P. Papalambros (2003) "Convergence properties of analytical target cascading," *AIAA Journal*, v41 n5 p897-905.

[4.54] Montoya-Weiss, Mitzi M. and Roger J. Calantone (1994), "Determinants of new product performance: a review and meta-analysis," *Journal of Product Innovation Management*, v11 (November) p397-417.

[4.55] Papalambros and Wilde (2000) *Principles of Optimal Design: Modeling and Computation*, 2nd ed., Cambridge Univ. Press, New York.

[4.56] Papalambros, P.Y. (2001) "Analytical target cascading in product development," keynote lecture, *Proceedings of the 3rd ASMO UK / ISSMO Conference on Engineering Design Optimization*, Harrogate, North Yorkshire, England, July 9-10.

[4.57] Pullman, M.E., W.L. Moore, D.G. Wardell (2002) "A comparison of quality function deployment and conjoint analysis in new product design," *Journal Of Product Innovation Management*, v19 n5 p354-364.

[4.58] Sethi, Rajesh, Daniel C. Smith and C. Whan Park (2001) "The effect of cross-functional teams on creativity and the innovativeness of new products," *Journal of Marketing Research*, February, p73-85.

[4.59] Song, X. Michael and Mark E. Parry (1992) "The R&D-marketing interface in Japanese high-technology firms," *Journal of Product Innovation Management*, v9 (June) p91-112.

[4.60] Souder, William E. (1988) "Managing relations between R&D and marketing in new product development projects," *Journal of Product Innovation Management*, v5 (March) p6-19.

[4.61] Srinivasan V, W.S. Lovejoy and D. Beach (1997) “Integrated product design for marketability and manufacturing,” *Journal of Marketing Research*, v23 p154-154.

[4.62] Train, Kenneth E. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press.

[4.63] Urban, Glen L. and John R. Hauser (1993) *Design and Marketing of New Products*, Englewood Cliffs, N.J.: Prentice-Hall.

[4.64] Wassenaar, Henk and Wei Chen (2003) “An approach to decision-based design with discrete choice analysis for demand modeling,” *ASME Journal of Mechanical Design*, v125 p490-497.

[4.65] Wittink Dick. R., Marco Vriens, and Wim (1994) “Commercial use of conjoint analysis in Europe: results and critical reflections,” *International Journal of Research in Marketing*, v11 p41-52.

[4.66] Wittink, Dick R. and Philippe Cattin (1989) “Commercial use of conjoint analysis: an update,” *Journal of Marketing*, p91-96.

CHAPTER 5

HETEROGENEITY OF PREFERENCES AND THE DESIGN OF PRODUCT LINES

The previous chapter introduced a methodology for coordinating product development decisions between marketing and engineering design using analytical target cascading. The scope was limited to design of a single product, and the preference model used to predict demand for the product was homogeneous: i.e., differences in individual preference are explained only as random draws of the error term in a single demand model with a single set of parameters. In other words, all heterogeneity of preferences across the population is modeled in the unobserved error term, and the model does not attempt to describe the structure of heterogeneity over the population. Under the properties of this model, including the IIA property (Louviere *et al*., 2000), a single product emerges as the optimum, and design of a line of products is not meaningful. In this chapter, a more complex model of consumer preferences is developed, accounting for heterogeneity of preferences across the population, and the model is used to optimize a product line. The decomposition offered by ATC is particularly important in this instance because the design of each individual product is treated as a separate subproblem. If all products were to be designed at once, the dimensionality and complexity of the search space would dramatically increase with the number of products. ATC offers a method for coordinating these individual design problems in such a way that the combination of the products in the marketplace performs optimally. The material in this chapter is based on an article by Michalek *et al*. (2005).

5.1 BACKGROUND

It is a truism bordering on cliché that producers sell not merely isolated products, but product lines. As such, each producer competes against both itself (cannibalization) and rivals, and it must determine how to do so effectively. The academic community has responded with an array of techniques to select sets of products, which, considered jointly, optimize a given objective function, typically total profit for the line.

Marketing approaches to what has become known as the *product line selection problem* (Green and Krieger 1985; McBride and Zufryden 1988; Dobson and Kalish 1988) or *product line design problem* (Kohli and Sukumar 1990) have been many and varied. Certain commonalities, however, can be noted across them: elicitation of consumer preferences; an econometric model for mapping preferences onto product characteristics; an explicit demand or objective function; and, linking these, an optimization method. The space of product characteristics is typically modeled with discrete variables, and searching the space is often prohibitive (certain m are known to be NP-hard), so the lion's share of prior research has been devoted to efficient, scalable optimization algorithms (Chen and Hausman 2000; Steiner and Hruschka 2002). Each of the components of this body of work is sufficiently well-developed to consider it a mature methodology, deployable when market and product characteristics are well understood.

In the engineering community, design decision-making models typically involve maximizing a set of performance objectives to best meet a set of exogenously determined product characteristic targets. Recently, researchers in this community have begun to study optimal product line design in the context of models of producer objectives (Li and Azarm, 2002); however, the bulk of such efforts in engineering focus on the study of product commonality and platforms (Fellini, 2003; Simpson, 2004).

One might question why firms do not simply integrate marketing decision making with engineering design and production from the outset. The conclusion of the extant literature on the topic is that there are non-trivial impediments to such integration (e.g., Griffin 1997, Wind and Mahajan 1997). Hauser (2001) highlights the difficulty of coordinating resources and agents across functional areas, each with idiosyncratic notions of "success," suggesting a set of metrics for that purpose. The goal here is less one of measurement than pragmatism: embedding marketing and engineering product line design models into a deployable formal framework that admits of real consumer choice data and that can be reliably optimized.

In the real world of product development, marketers and engineers face the continual challenge of having to conform to one another's constraints. Marketers, including those in the academic community, predicate their methods and models on an understanding of what might be termed the "product characteristic space": the characteristics, and levels of those characteristics, which can be manipulated to attract consumers and generate profit. In contrast, design engineers work in the "design space," selecting values for detailed, technical design variables in order to achieve desired target product characteristics. If engineering is unable to design and produce a feasible product that is sufficiently close to the targets identified in the marketing analysis, additional rounds of measurement, design and intercommunication must be undertaken. It has been amply documented in the product development literature that this back-and-forth passing of product characteristic targets between those who study preferences for products and those who engineer and build them is costly, time-consuming and, most importantly, can lead to sub-optimal final designs (Griffin and Hauser, 1992; Krishnan and Ulrich, 2001). How this might be rectified in formal product line design models has remained an open question, and the extension presented in this chapter addresses this question.

Part and parcel of creating a formal system for product line optimization is to allow marketing and engineering design models to 'speak to' one another in a well-defined, efficient, and convergent manner. Their domains and terminology differ, even when restricted to questions of product design (Krishnan and Ulrich, 2001). A first step, therefore, in fashioning an overarching model for product line optimization is standardizing terminology and defining proper coordination of overall objectives so that each discipline can bring to bear its own proven modeling techniques. One historical obstacle to achieving this objective is differences in control variables: Marketing makes use of primary research methods, for example, questionnaires and focus groups, to infer what consumers want; that is, consumer preferences are taken as primitives. A typical approach in the literature is to view each product as a "bundle of attributes" (McAlister 1982) and use measured consumer preferences to locate "sweet spots" in the space of consumer-defined product characteristics. Engineers, however, do *not* optimize with respect to product characteristics or attributes themselves, but instead manipulate "lower-level" design variables to achieve goals for these characteristics or attributes. For example, what may be conveyed to a consumer as "sturdy" or "durable" (perceptual attributes) must be translated into the set of technical specifications for physical characteristics (loading conditions, allowable deflections, yield strength) that engineers will meet through manipulation of design variables (metal thickness, spring tension, etc.), subject to inviolable physical and geometric constraints. This interaction is largely absent from the marketing literature on product line optimization, and it is among the main contributions of the methodology introduced in this chapter.

By extending the material introduced in Chapter 4, this chapter introduces a novel approach to the product line optimization problem, one which formally links marketing and engineering product design submodels in order to ensure feasibility of the product line while pursuing producer objectives with respect to customer preferences. This approach calls on proven strengths of models in both disciplines best suited to intercommunication and joint optimization. From marketing, hierarchical Bayesian (HB) methods and choice-based conjoint analysis are employed to build data-driven discrete choice models of demand using a general representation of consumer preference heterogeneity and capitalizing on Monte Carlo techniques. From engineering design, engineering performance design optimization models are utilized and the ATC methodology is extended for the multi-product case. ATC coordination of existing marketing methods, such as conjoint analysis, discrete choice analysis and demand forecasting, with engineering performance design optimization models allows optimal product line designs to emerge seamlessly and without tedious, human-mediated iterations between marketing and engineering design submodels. As in Chapter 4, joint optima are produced that are superior to those arising from independent marketing and engineering design optimization.

This chapter proceeds by outlining key concepts and reviewing the relevant methods and literature from both marketing and engineering product design. ATC is then adapted to link models from the two disciplines for product line design formally, and several alternative models of demand are constructed for use within the ATC methodology. Finally, it is demonstrated how a full product line might be devised for the dial-readout bathroom scale example.

5.1.1 Marketing Product Line Design Literature

The product line design literature in marketing is largely concerned with setting product characteristic and price levels to maximize profit or some alternative metric like sales or market share. As discussed at length subsequently, this optimization is bedeviled both by combinatorial complexity and by the simple need to avoid product characteristic combinations that engineering design cannot physically produce.

Product line optimization entails a number of intrinsic tradeoffs. Rutenberg (1971) put the main one succinctly, as balancing "…the disutility of refusing to provide each segment of customers with an item fitting its exact requirements versus economies of scale achieved in producing and inventorying each item." From this premise, he explored the optimal 'depth' of a product line (the number of unique items to produce) through dynamic programming techniques. A main contribution of the ensuing train of work on product line optimization is the integration of formal methods of nonlinear and discrete optimization with value/utility measurement techniques such as conjoint analysis. Given a set of either individual-, segment- or market-level preference functions – and there are important differences among these, which are explored subsequently – research in marketing has focused on efficiently optimizing some objective conditional on them. Among the earliest work in this vein was that of Zufryden (1977, 1982), who proposed an integer programming formulation for optimal design of both single and multiple products in a line. Green and Krieger (1985) subsequently showed how to extend existing single product design optimization heuristics to entire product lines. They were also among the first to acknowledge the prohibitive nature of any sort of global optimization and suggest heuristic alternatives that perform well in practice (Green and Krieger 1987b, 1992).

The general product line optimization problem was rendered into what might be called its current form by Dobson and Kalish (1988) and by McBride and Zufryden (1988). Dobson and Kalish consider a known set of related or substitute products and ask how their characteristics can be optimally set by a monopolist to maximize some criterion of interest, ordinarily line profit. Importantly, they show how appropriate data can be obtained from consumers, how the model can be estimated, and what sort of heuristics can be applied to solve the line optimization problem for real applications. They did, however, need to invoke a number of assumptions as a nod to technology as it existed at the time. The most important of these is that the market is either composed of homogeneous customer segments or that estimation be done at the individual level, for which there is rarely the luxury of sufficient per-respondent data. McBride and Zufryden (1988) treat the problem in a rigorous mathematical fashion, extracting individual consumer measurements from conjoint analysis, and selecting appropriate product characteristic levels using integer programming. They show that their proposed system can solve realistic product line problems using what are, by today's standards, modest computational resources.

A great deal of subsequent research sought computationally efficient algorithms to pinpoint the 'best' product (or set of products) from all available candidates. Kohli and Krishnamurti (1987) and Kohli and Sukumar (1990) provide an extensive analysis of dynamic programming heuristics for single-product solutions using hybrid conjoint designs relative to a number of objective functions. In looking back to these proposed methods, one must remember that computational efficiency is an evolving concept. The Bayesian econometric and ATC approach advocated here, for example, were well beyond feasibility even a decade ago. Dobson and Kalish (1993) incorporate conjoint data into a mathematical programming formulation with differential costs across various product profiles. Moreover, they not only show formal equivalence to the well-known (NP-complete) problem of plant location, but that their generalized version of the product line design problem can be addressed using the greedy interchange heuristic. Realizing that practical progress on the line design problem depended on robust, scalable heuristic methods to search the high-dimensional discrete space of possible solutions, Nair, Thakur and Wen (1995) showed that so-called beam search heuristics were arguably superior to any which had been previously analyzed, and could find near-optimal multi-product solutions even for line design problems with many product characteristics and levels. Chen and Hausman (2000) extend prior approaches to *choice-based* conjoint, highlighting certain special properties which allow the product line problem to be solved efficiently. Unfortunately, these properties fail to hold unless consumer preferences are presumed homogeneous, so their approach cannot be adapted for much practical work on choice modeling. Among the most recent methods is that of Steiner and Hruschka (2002) who adopt a genetic algorithm formulation (Balakrishnan and Jacob 1996) to locate near-optimal product line designs efficiently. They find GAs to be practical even for conjoint designs with many product characteristics and levels, identifying the globally optimal solution in the vast majority of cases and never erring by more than 4%.

5.1.2 Conjoint Analysis in Product Line Design

Among the most durable of marketing methodologies is conjoint analysis, the dominant approach to measuring 'consumer value systems' and thereby projecting the success of new product concepts. Since its introduction (e.g., Green and Rao 1971), the method has grown in sophistication along with computational resources and econometric methods (Louviere 1988; Green and Srinivasan, 1990; Green and Krieger 1996; Kuhfeld 2003). Indeed, conjoint analysis has been used successfully in the design of thousands of products across the world (Cattin and Wittink 1982, Wittink and Cattin 1989, Wittink *et al*. 1994). Much recent work focuses on the accurate representation of consumer heterogeneity, with a growing consensus in favor of HB formulations (Lenk *et al*., 1996; Huber and Train, 2001; Allenby and Rossi 2003), which are adopted here. Although not conceived purely as a method for product design optimization, conjoint analysis has nonetheless consistently provided the raw material upon which such frameworks have been built. Kaul and Rao (1995) expertly survey this literature.

By computational necessity, early research relied on the most tractable forms of conjoint analysis, based on either rankings or ratings, and only subsequently were extensions made to choice-based data (e.g., Louviere and Woodworth 1983). In addressing this problem, Chen and Hausman (2000) underscore difficulties hindering prior research, notably, that "one practically needs to build the entire choice simulator into the (optimization) program." They also point out that choice-based conjoint analysis typically presumes "the aggregated probabilistic representation of customer choice, which… diminishes most of the model's capability of accessing market heterogeneity at the ‘individual’ level and potentially leads to counterintuitive results.” While this is true if researchers rely on a homogeneous representation of consumer preference, one can appeal to a number of alternative heterogeneity formulations, as surveyed by Huber and Train (2001) and Andrews, Ansari and Currim (2002).

5.1.3 Engineering Design Literature

By contrast to the marketing literature, the bulk of the engineering design literature relevant to product line design focuses on studying product commonality and product platforms (Fellini, 2003; Simpson, 2004; Kokkolaras *et al.,* 2002), which is outside the scope of this dissertation. Within the DBD literature, introduced in Chapter 2 and Chapter 4, Li and Azarm (2000) proposed a two-step approach that they later extended to address product line design (2002) by first generating a set of designs that are Pareto-optimal in performance and then selecting a product line from that set based on a “first choice” model of demand using stochastic algorithms. This sequential approach can be effective for product line design cases where preferences for product characteristics are strictly monotonic across the consumer population (e.g., fuel economy, reliability, etc.) because in these cases the same Pareto set is relevant to all consumers, and individuals differ only in their preferred tradeoff among the competing performance objectives. In the scale example no such universal Pareto set exists because different consumers have different ideal points for size, shape, and other valued product characteristics.

Additionally, as discussed in Chapter 2 and Chapter 4, a host of MDO methods exist for decomposing complex problems into subproblems and coordinating their solutions. The ATC methodology is used here because of its capacity to accommodate arbitrarily large hierarchies through coordinating targets and linking variables and because it has proven convergence properties.

5.1.4 Proposed Methodology

Like Chen and Hausman (2000), a number of assumptions are invoked in this model to focus on product line optimization issues; specifically: that the total potential market size can be (exogenously) determined; that customers purchase either zero or one product from the line; that a customer's purchase decision is not overtly influenced by those of other customers; and that production can be scaled up or down as needed to suit demand. This formulation is well-suited to stable, durable goods categories, which comprise the majority of line optimization decisions. It is arguably less appropriate for rapidly-developing product classes (e.g., high-tech products), or those where aggregate demand is erratic. Unlike Chen and Hausman (2000) or other prior research in the area, only mild parametric assumptions are made about how to represent consumer preferences. Indeed, one of the main attractions of this approach is a fully Bayesian account of a most general form of preference heterogeneity and a method for line optimization conditional upon it. And, as in Chapter 4, a framework is provided for dealing with technologically 'unachievable' product combinations, treating them as an integral part of the modeling formulation.

Contingent on customer preference information, Steiner and Hruschka (2002) classify profit optimization methods into "one step" and "two step" types. The former optimizes directly over product lines as described by part-worth values and an objective (profit) function; the latter, realizing the combinatorial difficulties of such a search, first engage in some form of reduction of the full set of potential products, and search this reduced set for the best item(s). Generally speaking, culling the full set of potential products into a set of 'promising candidates' is non-trivial (Green and Krieger 1987a, 1987b, 1989). In this regard, the work here follows Kohli and Sukumar (1990), searching the entire space of possibilities without prior restrictions or pre-processing. However, the discrete nature imposed by the form of conjoint data (characteristics discretized into levels) is relaxed through natural cubic spline interpolation at the individual level to account for preferences at intermediate characteristic values. While the resulting 'profit surface' can therefore be highly non-linear and complex, it is locally smooth and continuous, allowing the use of efficient nonlinear programming algorithms and ensuring that ATC will converge, typically in a reasonable amount of time.

The proposed product line design methodology entails three stages: First, consumers choose among products in a choice-based conjoint setting, given a set of presented product characteristics; second, respondent choices are used as input data to estimate heterogeneous demand models, and the resulting preference coefficients are interpolated using splines; and third, ATC is used to coordinate optimization over the space of feasible product designs that yield optimal product characteristics. The first two stages are viewed as preprocessing for the ATC model.

The chapter proceeds by defining the ATC methodology for coordinating a product line in Section 5.2, conditional on a model to predict demand; describing three alternative discrete choice demand model specifications in Section 5.3; and finally, demonstrating the methodology with the application to dial-readout scales in Section 5.4 and discussing results in Section 5.5.

5.2 ATC COORDINATION OF MARKETING AND ENGINEERING MODELS

In this section the framework from Chapter 4 is extended so that a single "marketing subproblem" that plans the product line is coordinated with a set of "engineering design subproblems," one for each product in the line. An informal depiction of this process appears in Figure 5.1: The marketing subproblem involves determining price and (target) product characteristics for the full product line to maximize a known objective function, which can be profit or some other measure of interest to the firm, without deviating too much from the feasible designs achieved by engineering, while each "engineering design subproblem" requires determining a *feasible* design – one conforming to known constraints – that hews as closely as possible to product characteristic targets set in the marketing subproblem.

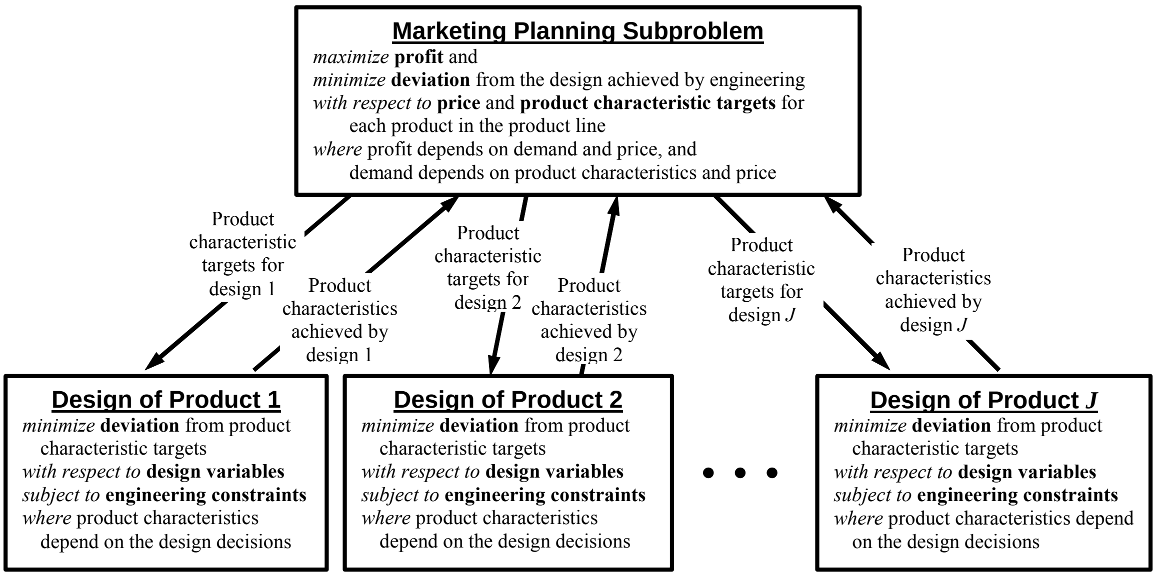


Figure 5.1: ATC statement of product planning and engineering design coordination

The marketing product planning subproblem differs from that in typical marketing applications in one critical way: Supposedly "optimal" product characteristics and price must be associated with a product that *can be achieved* by a feasible engineering design. The engineering design subproblems differ from those in typical engineering design applications in that no arbitrary weights need be assigned to specify multiobjective optimization tradeoffs in each of the engineering design subproblems: Instead, they are resolved by iteratively optimizing to meet product characteristic targets set by marketing. It is the interplay of minimizing deviations between target and achieved characteristics while pursuing a management goal (profit maximization) that ATC formalizes through iteration. Iteration between groups of real people – as opposed to mathematical subproblems – entails substantial fiscal and opportunity costs, both costs which ATC helps avoid.

The chief organizational benefit of ATC is that it separates models by discipline: Marketers can build submodels based on, say, conjoint analysis and new product demand forecasting, engineers can formulate models for product design and production, and other functional groups can focus on what they know how to do well. No functional area need become an expert in modeling the others, since ATC coordinates models with well-defined interfaces.

The following sections lay out the engineering design and marketing design subproblems explicitly.

5.2.1 ATC Marketing Subproblem

The marketing subproblem objective is to maximize profit Π with respect to the price p_j and the vector of real-valued product characteristic targets $\mathbf{z}^{\mathrm{M}}{}_j$ for each product j in the product line $j = \{1, 2, ..., J\}$ while minimizing deviation from the characteristics of the feasible design achieved by engineering $\mathbf{z}^{\mathrm{E}}{}_j$. Many profit formulations are suitable for the marketing objective function, and firms can specify arbitrarily sophisticated functions based on their experience, internal accounting and historical demand forecasting. As before, an especially simple profit (Π) formulation is adopted, revenue minus cost, so that

$$\Pi = \sum_{j=1}^{J} \left(\left(p_j - c_j^{\mathrm{V}} \right) q_j - c_j^{\mathrm{I}} \right), \tag{5.1}$$

where p_j is the (retail) price of product j, $c^{\mathrm{V}}{}_j$ is the unit variable cost of product j, $c^{\mathrm{I}}{}_j$ is the investment cost for product j, which represents all costs of setting up a manufacturing line for product j, and q_j is quantity of product j sold (demand), which is a function of the product characteristics $\mathbf{z}^{\mathrm{M}}{}_{j'}$ and price $p_{j'}$ of all products $j' = \{1, 2, ..., J\}$. It is assumed that product commonalities enabling investment cost sharing and improving economies of scale do not exist, so each new product requires new manufacturing investment. Further elaborations on this basic model are easily accommodated; for example, discounted future earnings, salvage values for finite horizons, and loss or risk functions can all be readily grafted onto Eq.(4.2) using known techniques. In general, c^{V}_{j} and c^{I}_{j} can be considered functions of market conditions or engineering decisions, although in this example they are taken as constants. Given a demand model to calculate q_j as a function of the vector of (target) product characteristics $\mathbf{z}^{M}_{j}$ and price p_j for all products j in the line, which will be developed in the next section, the profit function is fully defined; however, the objective function also involves minimizing deviation from the characteristics of the feasible design achieved by engineering $\mathbf{z}^{E}_{j}$, which are held constant while solving the marketing subproblem. As before, this deviation for each product j is represented in ATC using the square of the l_2 norm of the weighted deviation vector,

$$\left\| \mathbf{w}_j \circ \left(\mathbf{z}_j^{\mathrm{M}} - \mathbf{z}_j^{\mathrm{E}} \right) \right\|_2^2, \tag{5.2}$$

where $\| \|_2^2$ denotes the square of the l_2 norm, $\mathbf{w}_j$ is the weighting coefficient vector for product j, and the $\circ$ operation indicates term-by-term multiplication, so that $[a_1\ a_2\ ...\ a_n] \circ [b_1\ b_2\ ...\ b_n] \equiv [a_1b_1\ a_2b_2\ ...\ a_nb_n]$. For this formulation, the weighting coefficients $\mathbf{w}_j$ need only be chosen sufficiently large so that the deviation between the product characteristic targets set by marketing $\mathbf{z}^{M}_{j}$ and those achieved by engineering $\mathbf{z}^{E}_{j}$ at the solution is acceptably small. Finally, the marketing subproblem, conditional on a model for demand, is written as:

$$\begin{aligned} &\underset{\mathbf{z}_j^{\mathrm{M}}, p_j; \forall j \in \{1,2,\ldots,J\}}{\text{maximize}} \sum_{j=1}^{J} \left(\left(\left(p_j - c_j^{\mathrm{V}} \right) q_j - c_j^{\mathrm{I}} \right) - \left\| \mathbf{w}_j \circ \left(\mathbf{z}_j^{\mathrm{M}} - \mathbf{z}_j^{\mathrm{E}} \right) \right\|_2^2 \right). \\ &\text{where } q_j = q_j \left(\mathbf{z}_{j'}^{\mathrm{M}}, p_{j'}; \forall j' \right) \end{aligned} \tag{5.3}$$

In Section 5.3, conjoint analysis, discrete choice modeling and Bayesian Markov Chain Monte Carlo (MCMC) methods are used to develop the functional relationship between demand q and the marketing decision variables $\mathbf{z}^{\mathrm{M}}$ and p.

5.2.2 ATC Engineering Design Subproblems

In each engineering design subproblem j, a vector of design variables $\mathbf{x}_j$ is manipulated to achieve product characteristics $\mathbf{z}^{\mathrm{E}}_j$, where $\mathbf{z}^{\mathrm{E}}_j$ is expressed as a function of $\mathbf{x}_j$ using a vector valued response function $\mathbf{z}^{\mathrm{E}}_j = \mathbf{r}(\mathbf{x}_j)$. The design variable vector $\mathbf{x}_j$ is restricted to feasible values by a set of constraint functions $\mathbf{g}(\mathbf{x}_j) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x}_j) = \mathbf{0}$, and so vales for product characteristics $\mathbf{z}^{\mathrm{E}}_j = \mathbf{r}(\mathbf{x}_j)$ are implicitly restricted to values that can be achieved by a feasible design. Papalambros and Wilde (2000) present general procedures for defining $\mathbf{x}$, $\mathbf{r}(\mathbf{x})$, $\mathbf{g}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$ for engineering design problems, although details are necessarily case-specific. The objective of each engineering design subproblem is to minimize deviation between the engineering design product characteristics $\mathbf{z}^{\mathrm{E}}_j$ and the marketing targets $\mathbf{z}^{\mathrm{M}}_j$, which are held constant in the engineering design subproblems, and where deviation is measured as the square of the l_2 norm of the weighted deviation vector. The engineering optimization problem for product j can then be written as

$$\begin{aligned} &\underset{\mathbf{x}_j}{\text{minimize}} \ \left\| \mathbf{w}_j \circ \left(\mathbf{z}_j^{\mathrm{M}} - \mathbf{z}_j^{\mathrm{E}} \right) \right\|_2^2 \\ &\text{subject to } \mathbf{g}\left(\mathbf{x}_j\right) \leq \mathbf{0}, \mathbf{h}\left(\mathbf{x}_j\right) = \mathbf{0}, \\ &\text{where } \mathbf{z}_j^{\mathrm{E}} = \mathbf{r}\left(\mathbf{x}_j\right) \end{aligned} \tag{5.4}$$

5.2.3 Complete ATC Formulation

Figure 5.2 conveys a mathematical description of the concepts in Figure 5.1, showing the flow of the ATC-based product line optimization model for a single producer, where the number of products in the line J is determined through a parametric study: i.e., during optimization the number of products J in the product line $j = \{1, 2, \ldots, J\}$ is held fixed, separate optimization solutions are found for each value of $J = \{1, 2, ...\}$, and the value of J that produces the solution with the highest profit is selected. In the marketing product planning subproblem (Eq.(5.3)) vectors of characteristic targets $\mathbf{z}^{\mathrm{M}}_{j}$ and price p_j are chosen for each product j in order to maximize profit Π (i.e., revenue minus cost), as defined in Eq.(4.2). This is achieved while minimizing the deviation between the product characteristic targets set by marketing $\mathbf{z}^{\mathrm{M}}_{j}$ and those achieved by each engineering design $\mathbf{z}^{\mathrm{E}}_{j}$, using weighting coefficients $\mathbf{w}_j$ to specify the tradeoff between the objectives. The demand model, which will be developed in the next section, is left as a generic function in Eq.(5.3) and Figure 5.2. In each engineering design subproblem (Eq.(4.1)), design variables $\mathbf{x}_j$ are chosen to minimize the deviation between characteristics achieved by the design $\mathbf{z}^{\mathrm{E}}_{j}$ and targets set by marketing $\mathbf{z}^{\mathrm{M}}_{j}$ subject to engineering design constraints $\mathbf{g}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$.

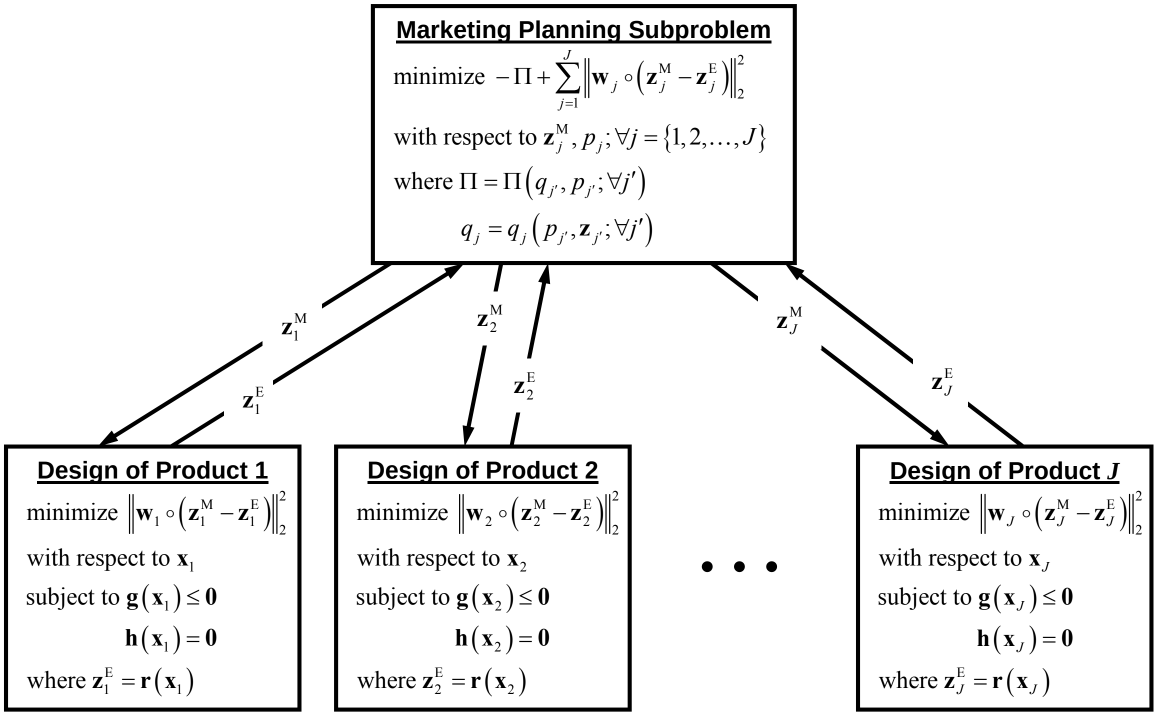


Figure 5.2: Mathematical ATC formulation of product planning and engineering design coordination

5.3 MODELS OF PRODUCT DEMAND

This section outlines the models used to predict demand for a product as a function of the product characteristics and prices of all of the products in the marketing subproblem of the ATC formulation. Three different discrete choice models of demand, which differ in their assumptions about the form of the stochastic error terms and the representation of preference heterogeneity, are presented, and each model is fit to data obtained through the online choice-based conjoint study.

5.3.1 Conjoint Choice Data Collection

In the choice-based conjoint survey, as described earlier, the respondent is presented with a series of questions or "choice sets" $t = \{1, 2, ..., T\}$. In each choice set t, the respondent is presented a set of product alternatives $j \in \mathcal{J}_t$ including the option to not choose any of the product alternatives: the "no choice" option. Each choice set contains hypothetical products with each characteristic and price set to one level among the sets of discrete levels shown in Table 5.1.

Table 5.1: Product characteristic and price levels

	Description	Metric	Levels					Units
z_1	Weight capacity	Weight causing a 360° dial turn	200	250	300	350	400	lbs
z_2	Aspect ratio	Platform length divided by width	6/8	7/8	8/8	8/7	8/6	--
z_3	Platform area	Platform length times width	100	110	120	130	140	in^2
z_4	Tick mark gap	Distance between 1 lb tick marks	2/32	3/32	4/32	5/32	6/32	in.
z_5	Number size	Length of readout number	0.75	1.00	1.25	1.50	1.75	in.
p	Price	US Dollars	10	15	20	25	30	$

The characteristic and price levels of the product alternatives in each choice set are chosen to be balanced and to collect data efficiently (as per Kuhfeld 2003, who provides a comprehensive overview). The resulting data are the observed choices that each respondent makes in each choice set: Φ_{ijt}, where $\Phi_{ijt} = 1$ if respondent i chooses alternative j in choice set t, and $\Phi_{ijt} = 0$ otherwise. These $\{\Phi_{ijt}\}$ are then used to estimate the parameters of the demand model for the marketing subproblem, as illustrated in Figure 5.3.

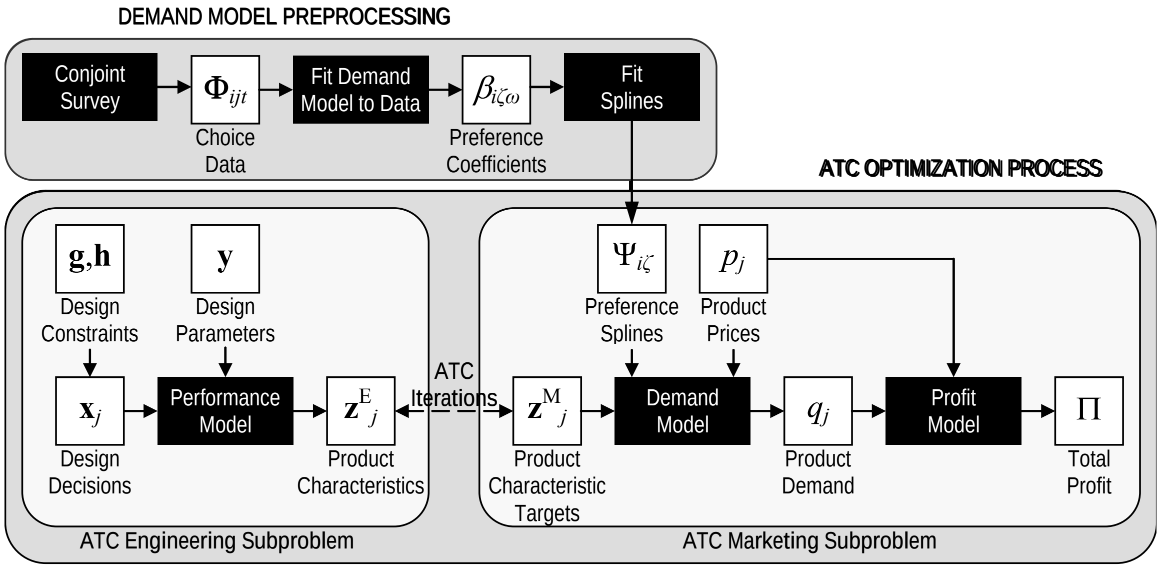


Figure 5.3: Diagram of the modeling process

5.3.2 Discrete Choice and Random Utility Models

As before, a utility framework is adopted (Train, 2003) to derive total line demand. Many of the previous product line optimization formulations (e.g., Dobson and Kalish, 1988; McBride and Zufryden, 1988; Li and Azarm, 2002) adopt a "first-choice" model of demand, where utility is written as a deterministic function of the product and consumer characteristics, and the product with the highest utility is presumed to be chosen deterministically. Here a random utility model is used instead where the utility of each product to each consumer is assumed to depend on the product's characteristics, the consumer's idiosyncratic preferences for those characteristics (which are estimated from choice data), and a random error component, for which a specification must be given explicitly and from which choice probabilities formally arise. Demand is then written in terms of the probability of each individual choosing each alternative. The use of a random utility model avoids the discontinuities intrinsic to the 'first choice' model, and is important for generalizing the product design space to a continuous space so that efficient gradient-based nonlinear programming optimization algorithms can be called upon.

Throughout, subscripts are used to designate a set of representative individual consumers (or groups thereof) $i = \{1, 2, ..., I\}$ and a set of product alternatives $j = \{1, 2, ..., J\}$. Although the set of product alternatives can be consumer- or choice-occasion-specific (for example, in the conjoint choice sets), for clarity the additional subscripting these would require is avoided. Individuals i are assumed to derive from each product j some utility value u_{ij} that is composed of a deterministic component v_{ij}, which is a function of the observable aspects of the choice scenario, and an unobservable random error component ε_{ij}, so that $u_{ij} = v_{ij} + \varepsilon_{ij}$. It is assumed that each individual will choose the alternative that gives rise to the highest utility, i.e., alternative j is chosen by individual i if $u_{ij} > u_{ij'}$ for all $j' \neq j$. The probability, P_{ij}, that alternative j is chosen by individual i on a particular occasion can therefore be computed as:

$$P_{ij} = \Pr\left[v_{ij} + \varepsilon_{ij} \geq \left\{ v_{ij'} + \varepsilon_{ij'} \right\}_{\forall j' \neq j} \right]. \tag{5.5}$$

The value of P_{ij} depends on the assumed error (ε_{ij}) distribution. Two such distributions are used in essentially all choice modeling work: normal and double exponential, resulting in the logit and probit models, respectively. It is well-known (Amemiya 1985) that very large samples are required to distinguish results produced by the logit and probit specifications. Each offers advantages: The normal distribution is expedient for Bayesian computation, due to conjugacy properties; the double exponential distribution allows a closed-form expression for P_{ij}, which is precise and computationally efficient for optimization.

The deterministic component of utility v is taken as a function of observables such as price and product characteristics, or even consumer covariates, etc. Here only price and product characteristics are considered, elements under joint marketing and engineering control. As a practical matter, some rule is required to map characteristics into v. Recent econometric work explores non-parametric representations for individual-level (Kalyanam and Shively 1998; Kim *et al*. 2004b) or latent utility functions (Andrews, Ainslie and Currim 2002). In order to accommodate *post hoc* iterative optimization, each characteristic ζ of the real-valued characteristic vector $\mathbf{z}^{M}j$ is divided into a set of discrete levels $\omega = \{1, 2, \ldots \Omega_{\zeta}\}$ and a linear mapping of discrete product characteristic levels (conjoint part-worths) is adopted, accounting for possible non-linearities with respect to the underlying continuous variables using spline interpolation for intermediate values, which will be discussed later. The deterministic utility v_{ij} that individual i derives from product j is therefore written as

$$v_{ij} = \sum_{\zeta=0}^{Z} \sum_{\omega=1}^{\Omega_{\zeta}} \beta_{i\zeta\omega} \delta_{j\zeta\omega} \,, \tag{5.6}$$

where the binary dummy $\delta_{j\zeta\omega} = 1$ indicates alternative j possesses characteristic ζ at level ω, and $\beta_{i\zeta\omega}$ is the part-worth coefficient of characteristic ζ at level ω for individual i. To clarify notation, in $\delta_{j\zeta\omega}$ the elements of the product characteristic vector $\mathbf{z}^{M}_{j}$ (which does not include price) are enumerated $\zeta = \{1, 2, \ldots Z\}$, and price p is included in $\delta_{j\zeta\omega}$ and labeled as element $\zeta = 0$. Note that each product characteristic ζ is either intrinsically discrete or is discretized into Ω_{ζ} levels, $\omega = \{1, 2, \ldots \Omega_{\zeta}\}$. The use of discrete levels is crucial: It does not presume linearity with respect to the underlying continuous variables (e.g., a \$1 price decrease cannot be presumed to have the same effect for a \$10 product and a \$100 one, and preferences for values of product characteristics cannot, in general, be assumed monotonic).

In using choice models, one must account for the real possibility that none of the presented alternatives is deemed acceptable; this is accomplished through an ‘outside good' or 'no choice option.' This outside good is included throughout the empirical application, and it plays the vital role of ensuring that the total demand for a set of undesirable or overly expensive products will be low. In the absence of an outside good, consumers are forced to choose from the inside goods, and demand depends only on differences in offered product characteristics (cf. Berry, 1994). Haaijer, Kamakura and Wedel (2001) survey the useful properties of such a "no choice" option as well as how to best accommodate it in conjoint designs. Throughout, the no choice option is indexed as alternative 0, with error ε_{i0} and attraction value v_{i0} for individual i, where v is normalized to zero ($v_{i0} = 0$; $\forall i$) for identification purposes.

For purposes of comparison, three demand models are presented that arise from different specifications in each of the following elements: (1) heterogeneity in the part-worth coefficients β over the population, and (2) the distributional form of the stochastic error components of utility ε_{ij}; details on estimation of the part-worth coefficients are presented in each case.

5.3.3 A Simple Homogenous Heterogeneity Specification

In the simple homogeneous case, as used in Chapter 4, all individuals are presumed to belong to the same segment (i.e., they are all presumed to have the same preference coefficients$\beta_{\zeta\omega}$, and therefore the same deterministic component of utility v_j for a given product), so the consumer segment index i drops out. If, the stochastic error terms ε_{ij} are taken to be independent and identically distributed across products and to follow the "extreme value" (or "double exponential") distribution (i.e., $\Pr(\varepsilon_{ij} < \alpha) = -\exp(-\exp(-\alpha))$), then the choice probability P_{jt} in Eq.(5.5) for product j in choice set t reduces to the well-known logistic (logit) expression:

$$P_{jt} = \frac{e^{v_j}}{\sum_{j' \in \mathcal{J}_t} e^{v_{j'}}}, \tag{5.7}$$

where v_j is defined in Eq.(5.6). In this and the next section, the logit model is used as it allows closed-form expressions for choice probabilities; the probit case, discussed in detail in a subsequent section, is analogous in every way other than the error specification, and neither is to be preferred on *a priori* grounds. Given observed consumer choice data Φ, the parameters $\beta_{\zeta\omega}$ can be estimated using maximum likelihood techniques that are available in many statistical packages: Values for the part-worth coefficients $\beta_{\zeta\omega}$ are found in order to maximize the probability that the model in Eq.(5.6)-(5.7) predicts the observed choice data. Specifically, if Φ_{ijt} defines the observed choices such that $\Phi_{ijt} = 1$ if individual i selects alternative j from choice set $\mathcal{J}_t$ on choice occasion t, and $\Phi_{ijt} = 0$ otherwise, then the probability that (a random draw from) the model will predict the same choices observed in the data is given by

$$\prod_{t=1}^{T} \prod_{i=1}^{I_S} \prod_{j \in \mathcal{J}_t} \Phi_{ijt} P_{jt} \ . \tag{5.8}$$

When choosing the β coefficients to maximize this quantity, it is standard practice to maximize the log of the likelihood (which, due to monotonicity, retains the same maximum) in order to improve search-based numerical properties of the formulation, so that

$$\beta_{\zeta\omega} = \underset{\beta_{\zeta\omega}}{\arg\max} \left(\sum_{t=1}^{T} \sum_{i=1}^{I_S} \sum_{j \in \mathcal{J}_t} \log\left(\Phi_{ijt} P_{jt}\right) \right), \tag{5.9}$$

where P_{jt} is defined by Eq.(5.7), and v_j is defined by Eq.(5.6). Equation (5.9) yields an infinity of equivalent solutions for the β coefficients because, for example, adding a constant to each v_j leaves their relative values unchanged, and so does not affect choice probabilities in the logit model (location invariance). To therefore 'identify' the model, a linear combination of the β coefficients for each characteristic is set to zero. Specifically, we select from the infinity of equivalent coefficient vectors the one where the mean value $\sum_{\omega=1}^{\Omega_\zeta} \beta_{\zeta\omega} / \Omega_\zeta$ is the same for all ζ. By adding this constraint, the model has $1+\sum_{\zeta=1}^{Z}\left(\Omega_\zeta - 1\right)$ degrees of freedom per individual, and the solution is uniquely identified.

Such models have two important shortcomings. The first limitation is the assumption of independently and identically distributed error specifications across alternatives, producing unrealistic substitution patterns: Often referred to as the independence from irrelevant alternatives (IIA) property, the assumption of uncorrelated error terms in the model implies that the model will underestimate the degree to which two highly similar objects compete with one another. Secondly, the model above assumes homogenous preferences for the consumers where differences among consumer choices are explained only through random draws of the error component, which is unrealistic in most cases (Leeflang *et al*., 2000) since consumers typically differ in behavior and preferences. Failure to correctly model this heterogeneity leads to biased parameter estimates, which can adversely affect profitability of the resulting solution, particularly when designing product line. In the following sections two alternative models are presented that have more realistic patterns for heterogeneity.

5.3.4 A Discrete Mixture Heterogeneity Specification

The limitations of the homogeneity restriction can be addressed by presuming that there exist a fixed number of consumer segments B, where each segment b = 1,2,...,B

contains s_b fraction of the population, and individuals in segment b have identical preferences $\beta_{b\zeta\omega}$, but preferences differ between segments. This is the well-known discrete mixture or *latent class model* (Kamakura and Russell 1989). In this model, a certain individual belongs to segment b with probability s_b, and the deterministic component of utility v_{bj} derived from product j by a member of segment b is given by Eq.(5.6) using the homogeneous within-segment preferences $\beta_{b\zeta\omega}$. If an independent and identically distributed (i.i.d.) extreme value error distribution is again specified, the unconditional choice probability P_{jt} of choosing product j from choice set t is now the weighted sum:

$$P_{jt} = \sum_{b=1}^{B} \left(s_b \frac{e^{v_{bj}}}{\sum_{j' \in \mathcal{J}_t} e^{v_{bj'}}} \right), \tag{5.10}$$

where v_{bj} is defined in Eq.(5.6). As with the previous aggregated model, the discrete mixture model is straightforward to estimate using the maximum likelihood technique with respect to the β and s_b terms. As before, the probit case differs only in error distribution. Model selection (to choose among different values of B) can be achieved through standard measures as BIC, which is used here (Kadane and Lazar, 2004).

Latent classes b can be viewed as market segments, a concept with which managers are comfortable, and which are appealing if there are advantages in targeting these segments separately, for example, via production or advertising (Leeflang *et al.*, 2000; Wedel and Kamakura, 2000). The discrete mixture model also helps address the IIA criticism. It is well-established that such problems are exacerbated, particularly in choice-based conjoint (CBC) models, when heterogeneity is not adequately accounted for. In summarizing their vast empirical experience with CBC models, Sawtooth (Orme, 1998) concluded "IIA is much more problematic with aggregate logit and much less a problem with methods that recognize respondent heterogeneity, such as Latent Class and ICE (Individual Choice Estimation)." Nevertheless, it is often argued that a discrete representation of heterogeneity is too restrictive, in particular when true preferences are continuously distributed. In such a case, a finite mixture model leads to an artificial partitioning into homogeneous segments (Leeflang *et al*., 2000). If a continuous mixing distribution is assumed instead (e.g., a random coefficient specification with a *single* normal mixing component), results are known to be sensitive to the (subjectively) specified distribution of the random components, and an overt misspecification can lead to biased results (Nevo, 2000, or Kim, Menzefricke and Feinberg 2004a). In the following section a model is introduced that combines both discrete and continuous heterogeneity to approximate a wide variety of heterogeneity 'shapes.'

5.3.5 A Heterogeneous Normal Mixture Specification

The heterogeneous probit model presented here includes several past conjoint choice models as special cases (Keane 1992, McCulloch and Rossi 1994, Haaijer 1999), as well as both model types described above. A specification is adopted which allows for *correlated* part-worth values within each mixing component, unlike in the previous two specifications. (Note that these correlations are between model β coefficients, not unobserved ε error terms, which is a separate issue.) We can test explicitly whether these correlations are warranted by restricting them to zero and noting whether and how fit measure results suffer. Formally, individual aggregate part-worth coefficient vectors $\boldsymbol{\beta}_i$, containing the elements $\beta_{i\zeta\omega}$ for individual i, are assumed to be drawn from a mixture of multivariate normal distributions. Although the individual-level part-worth parameters $\beta_{i\zeta\omega}$ for the previous models could be obtained by maximum likelihood (e.g., Louviere *et al*., 2000), such an approach is impractical for the 'continuous' heterogeneity representations in this section. Instead, hierarchical Bayes methods are used to estimate model parameters via MCMC techniques (Andrews, Ainslie, and Currim, 2002). Allenby and Rossi (2003) persuasively argue for the superiority of HB-based methods over alternative approaches. The DIC statistic (Spiegelhalter *et al*., 2002) is then used to compare candidate models (i.e., to choose the best number and form of the mixture components), as detailed later.

If error terms ε_{ij} are normally distributed and are independent across individuals, alternatives and choice sets, Eq.(5.5) gives rise to a probit model, which offers certain attractive conjugacy properties in a Bayesian setting, but does not yield probabilities in closed form[5]. Furthermore, it is assumed that $\boldsymbol{\beta}_i$ is drawn from a mixture of B multivariate normal distributions, so that

$$\boldsymbol{\beta} \sim \sum_{b=1}^{B} s_b N\left(\boldsymbol{\theta}_b, \boldsymbol{\Lambda}_b\right), \tag{5.11}$$

where s_b is the fraction of the market in 'segment' b. Here $\boldsymbol{\theta}_b$ is the vector of means for $\boldsymbol{\beta}_i$, and $\boldsymbol{\Lambda}_b$ is a full-variance covariance matrix (against which diagonality can be checked). This allows for a very flexible representation, with both discrete and continuous heterogeneity (Lenk and DeSarbo 2000).

This specification is still not free of the IIA criticism, due to i.i.d. errors. While it is important not to downplay potential IIA problems, they are mitigated by a reasonably large number of 'well mixed' conjoint choice sets, as is the case for the scale application, and, above all, appropriate heterogeneity correction. In this case, because the multivariate distribution describing $\boldsymbol{\beta}_i$ has a non-diagonal covariance matrix even when there is a

---

[5] As discussed above, here normal error terms are used and the consequent probit model for P_j. It is in any case typically possible to post-process MCMC draws to convert between estimates obtained through various error specifications, using auxiliary Metropolis steps.

single mixing component, product alternative part-worth values are not independent across consumers, reducing the impact of IIA (Nevo, 2000; or for an alternative account, see Haaijer *et al*. 1998).

Although Bayesian estimation methods are more forgiving in this regard than classical counterparts (Rossi and Allenby 2003), care still must be taken to ensure appropriate model identification; as before, the model specification is completed by setting the "no choice" value (Haaijer *et al*. 2001) to zero separately for each individual. In this way, the estimated individual-level coefficients are easily interpreted as (1) deviations from the "no choice" value, and (2) as being the same, on average, across product characteristics. Note, then, that the *range* of the coefficients within-characteristic corresponds to how strongly that characteristic influences choice for a particular individual. Relatively unimportant characteristics will be 'flat', and this flatness will vary by individual.

Estimation of this particular form of probit model is facilitated through data augmentation. The latent variables u_{ij} enable straightforward application of MCMC techniques when $\boldsymbol{\beta}_i$ comes from a single normal distribution. Conditionally on these latent values, the formulation is simply a Bayesian linear model (Lenk *et al*. 1996, Spiegelhalter *et al*. 1996, McCulloch and Rossi, 1994). Incorporation of the mixture of normals instead of a single normal, however, does not complicate matters when Dirichlet distribution is used to define a conjugate prior for s_b, because the full conditionals are also ordered Dirichlet (Lenk and DeSarbo 2000). Given conjugate priors, the full conditional distributions can be readily derived for all model parameters (e.g., Rossi and McCulloch 1994, Lenk and DeSarbo 2000). Specifically, the following MCMC scheme is used:

Draw sequentially from $u_{ijt}\,|\,\Xi,\ \boldsymbol{\beta}_i\,|\,\Xi,\ \boldsymbol{\theta}_b\,|\,\Xi,\ \boldsymbol{\Lambda}_b\,|\,\Xi,\ \varphi_i\,|\,\Xi,\ s_b\,|\,\Xi,$

where Ξ is notational shorthand for all previously sampled quantities, i.e.,$\{u_{ijt}$, $\boldsymbol{\beta}_i$, $\boldsymbol{\theta}_b$, Λ_b φ_i, $s_b\}$, plus the observed choice data, $\{\Phi_{ijt}\}$. Under this scheme, u_{ijt} are drawn from a truncated normal distribution; $\boldsymbol{\beta}_i$'s are drawn from subject-specific multivariate normal distributions; $\boldsymbol{\theta}_b$ is drawn form a multivariate normal distribution; $\boldsymbol{\Lambda}_b$ is drawn from an Inverse Wishart distribution; and segment probabilities s_b are drawn from an ordered Dirichlet distribution. The variable φ_i is used only within the MCMC chain to draw the group membership of individual i at each iteration, conditional on the other parameters, so that $\varphi_I \in \{1, 2, ..., B\}$.

The Gibbs sampler draws sequentially from these full conditionals and, after numerous iterations, a sample from the true posterior density is obtained (Diaconis and Saloff-Coste 1998). Estimates from a classical mixture of probit models are used as starting values, and the Gibbs sampler is iterated until a stationary posterior is obtained. To mitigate autocorrelation in the Gibbs sequence, only every 10^{th} draw is retained for data thinning, after a burn-in of 50,000 iterations. Convergence for the MCMC routine was examined through iteration plots and by re-running the program with different starting values, including those obtained by estimating a homogeneous version of the model. In all cases, posterior means were stable across start values, and histograms of posterior marginals revealed no systematic differences.

In order to choose among the different number of segments B in the mixture representation for $\boldsymbol{\beta}_i$, the Deviance Information Criterion (DIC) statistic proposed by Spiegelhalter *et al*. (1998) is used. This method is particularly suited to complex hierarchical (Bayesian) models in which the number of parameters is "not clearly defined" (Spiegelhalter *et al*. 2002). Classical methods like the Akaike's information criteria (AIC) or Bayesian information criteria (BIC), which are appropriate for inference using aggregate models and discrete mixtures (as in the previous two sections), rely on a synthetic measure of complexity, typically a (linear or multiplicative) function of the number of parameters in the model and the total quantity of observations. By contrast, the

DIC statistic determines the ‘effective number of parameters’ entailed by the model specification itself, and so eschews calculations based on the number of parameters or observations, likelihoods, and particular penalty functions. Spiegelhalter *et al*. (1998) illustrate the use of DIC in a wide range of applications and find it to perform well across them, particularly so for hierarchical models. Furthermore, DIC is readily computed from standard MCMC output.

Formally, DIC is defined as

$$\mathrm{DIC} = 2\overline{D(\Xi)} + D(\overline{\Xi}), \tag{5.12}$$

where Ξ is an aggregation vector containing all model parameters, D is the deviance function

$$D(\Xi) = -2\log\left(\Pr\left(\Phi \middle| \Xi\right)\right) + 2\log\left(f\left(\Phi\right)\right), \tag{5.13}$$

the overbar indicates posterior mean, and f is a scaling function related only to the data and not to model parameters. After the Gibbs sampler has converged and a sample of both Ξ and $D(\Xi)$ is obtained, $\overline{D(\Xi)}$ and $D(\overline{\Xi})$ are readily computed as sample means across the MCMC draws. Spiegelhalter *et al*. (2002) provide additional detail.

Computation of DIC for a mixture representation of the part-worth coefficients requires, unlike in the Gibbs sampler itself, the direct calculation of P_{ij}, as would the LL for the probit model itself [6]. This probability must be computed at every iteration of the

[6] See also §7.3 of Spiegelhalter *et al*. (1998), for a longitudinal binary-choice probit example with individual-specific random intercepts.

Gibbs sequence (after the burn-in), for each individual and each choice set (though only for the chosen alternative). In the scale application there are four alternatives in each conjoint choice set (including the no-choice option), so this entails reducing a four-dimensional integral (e.g., Haaijer, 1999), which can require long running times for high accuracy (Genz, 2004). If the errors are independent, it is possible to reduce this to a one-dimensional integral by integrating across the error distribution for the item in question, so that

$$
\begin{aligned}
P_{ijt} &= \Pr\left[v_{ij} + \varepsilon_{ijt} > \left\{ v_{ij'} + \varepsilon_{ij't} \right\}_{\forall j' \in \{\mathcal{J}_t - j\}} \right] \\
&= \int_{\varepsilon_{ijt}=-\infty}^{\infty} \left[\prod_{j' \in \{\mathcal{J}_t - j\}} F_\varepsilon\left(v_{ij} - v_{ij'} + \varepsilon_{ijt} \right) \right] dF_\varepsilon\left(\varepsilon_{ijt} \right)
\end{aligned}
\tag{5.14}
$$

where F_ε is the normal cumulative distribution function. There are many specialized methods for carrying out this computation, including numerical integration, both full and quasi Monte Carlo algorithms, and the highly efficient GHK simulator (Hajivassiliou *et al*. 1996; or, see Haaijer 1999). To ensure accuracy for this important quantity – model comparison hinges upon it – it can be computed by a variety of methods: GHK; adaptive Simpson's and Radau quadrature; Niederreiter sequences; Genz's specialized algorithm; and GAUSS' (Aptech, 2000) built-in procedure. In the scale application no differences were found among them to at least five significant digits, suggesting a high degree of confidence in the numerical results.

The model put forth thus far yields several sets of quantities, and it is helpful to take stock of them in non-technical terms. For each survey respondent, the model offers a set of draws from (the posterior distribution of) that respondent's coefficients $\beta_{i\zeta\omega}$. One could then use this information for inferences about that particular individual, or the specific set of individuals used to calibrate the model. Taking a Bayesian perspective and focusing instead on parameters of the hierarchical model alone, $\{s_b, \theta_b, \Lambda_b\}$, which describe the mixture distribution, these parameters can be thought of as *generating* the individual-level $\boldsymbol{\beta}_i$ values, and so are the appropriate quantities for inference and prediction. It is important to realize that these are not known with certainty (i.e., s_b, $\boldsymbol{\theta}_b$, and $\boldsymbol{\Lambda}_b$ are themselves random variables), and they possess a jointly-specified posterior. It is this posterior over which optimization occurs. To do this requires integration over the *posterior choice surface* arising from the MCMC chain for these parameters.

In the empirical application, optimization over this posterior surface is achieved through Monte Carlo integration, as follows: When the chain has stabilized, *new* values of $\boldsymbol{\beta}_i$ are generated as the chain continues to run, and they are thinned to reduce serial correlations; specifically, 10000 values are generated and every 10^{th} is retained. It is this resulting set of 1000 $\boldsymbol{\beta}_i$ draws which are used to represent the population (the posterior surface) throughout the optimization. Note that higher fidelity to this surface can be achieved simply by generating additional $\boldsymbol{\beta}_i$ values.

5.3.6 Spline Interpolation Conditional on Generated Coefficients

The part-worth coefficient values $\beta_{i\zeta\omega}$ represent individual-level preferences for each characteristic at each discrete level used in the conjoint study, and as such convey no direct information about intermediate values of the product characteristics and price. Optimizing over presumably continuously-valued product characteristics (including price) requires interpolation to such intermediate values. *Natural cubic splines* in particular have a number of desirable properties for interpolating vis-à-vis optimization, particularly so near endpoints (De Boor, 2001). Although a greater number of within-characteristic levels would enable higher-order splines, previous research has found even quadratic splines to suffice in a variety of Bayesian choice model applications (Kim *et al*. 2004b). Through interpolation the deterministic component of utility can be written as a function of continuous-valued product characteristics $\mathbf{z}_j$ and price p_j using natural cubic spline functions $\Psi_{i\zeta}$ fit through the discrete levels $\omega = \{1, 2, ..., \Omega_\zeta\}$ of the discrete part-worth coefficients $\beta_{i\zeta\omega}$ for each individual i and each product characteristic (plus price) ζ.

Natural cubic splines are fit through the set of discrete points by solving a system of equations corresponding to the coefficients of a set of cubic polynomials, each defined over the range between a pair of adjacent discrete levels, such that each polynomial passes through the two discrete points at its range boundaries, each polynomial matches the first and second derivative of adjacent polynomials at the discrete levels, and the second derivative of the extreme endpoints are zero. The resulting spline curve $\Psi_{i\zeta}$ for each individual i and each characteristic ζ passes through the discrete $\beta_{i\zeta\omega}$ values at each discrete level ω (each node), and the curve will possess a continuous derivative function, which is important for gradient-based optimization. Indexing characteristics as $\zeta = 1,\ldots,$ Z and price as $\zeta = 0$, the interpolated value of the observable component of utility is

$$\hat{v}_{ij} = \Psi_{i0}\left(\beta_{i0\omega}, p_j\right) + \sum_{\zeta=1}^{Z} \Psi_{i\zeta}\left(\beta_{i\zeta\omega}, \left\langle \mathbf{z}_j^{\mathrm{M}} \right\rangle_\zeta\right), \tag{5.15}$$

where $\langle \mathbf{z}^{\mathrm{M}}_j \rangle_\zeta$ indicates the ζ^{th} element of the vector $\mathbf{z}^{\mathrm{M}}_j$. These interpolated $\hat{v}_{ij}$ give rise, through the random utility specification, to expected individual choice probabilities P_{ij}, which are summed across individuals to generate total expected market share. It is important to realize that the ATC methodology requires only that demand be rendered as a function of price and product characteristics and does *not* hinge on a particular functional specification for this relationship. The demand surface emerging from this entire procedure – Bayesian estimation of the mixture-of-normals heterogeneity model, generation of candidate coefficients for Monte Carlo integration, and natural cubic spline interpolation – is complex, and may possess local minima.

5.3.7 Assessing Demand

In summary, three random utility discrete choice model specifications, with varying levels of capacity to represent heterogeneity, have been presented: a simple homogeneous model, a heterogeneous discrete mixture, and a heterogeneous (continuous) normal mixture. Data used to assess consumer preference are collected using a choice-based conjoint exercise designed to provide the greatest efficiency (lowest parameter variance) relative to the number of choice tasks (Kuhfeld, 2003). The part-worth coefficients of each model are then fit to the choice data using classical maximum likelihood techniques or, in the case of the (HB) heterogeneous mixture of normal distributions, MCMC methods. The part-worth coefficients, which express preferences for discrete levels of the product characteristics and price, are then used as inputs to fit natural cubic splines (one for each individual for each product characteristic) to project preferences for intermediate values. This set of splines is the final result of the demand modeling, and it is used directly to calculate demand in the marketing subproblem of the ATC formulation. This process is summarized in Figure 5.3.

Given that all products are available to all consumers at the time of purchase, calculating demand for product j is simply a matter of multiplying the (unconditional) probability that a randomly selected individual will choose product j by the market potential S. The market potential is assumed to be exogenously determined through any of a number of pre-market forecasting techniques (see Lilien *et al*. 1992), and it is taken as a given. To illustrate, under simple logit-based demand, demand for product j is then given by

$$q_j = S\sum_i \left(\frac{s_i e^{\hat{v}_{ij}}}{\sum_{j'=0}^{J} e^{\hat{v}_{ij'}}} \right), \tag{5.16}$$

where $\hat{v}_{ij}$ is calculated in Eq.(4.7) using the spline functions derived from the part-worth values of the appropriate discrete choice model specification. For the simple homogeneous case, there is only one "individual" i, and $s_i = 1$. For the discrete mixture case, i indexes the latent classes (segments), so that $i = \{1, 2, ..., B\}$, and si is the size of segment i. For the mixture-of-normals heterogeneity case, i refers to the draws taken from the MCMC chain, so that $i = \{1, 2, ..., I_D\}$, where I_D is the number of draws, and $s_i = 1/I_D$. For a particular case, Eq.(5.16) can be used to specify demand in Figure 5.2 for ATC product line optimization coordination.[7]

5.4 APPLICATION

The proof of the proposed methodology lies in its ability to design a real product line. Here, the methodology is applied to the design of a line of dial-readout bathroom scales. While the methodology can be used for durables of any type, so long as they are in a fairly stable market, scales possess several features which make them attractive for purposes of illustration, as mentioned previously. From a marketing perspective, consumers are highly familiar with scales through personal interaction, have a good idea of which product features appeal to them, and the number of these features is not prohibitively large. Moreover, mechanical bathroom scales are neither highly differentiated nor a commodity product, with prices in a relatively restricted and well-known range. From an engineering design perspective, dial-readout scales are straightforward enough that one can readily specify a succinct list of design variables and a set of physical and geometrical constraints they need obey. ATC has been applied to

[7] The probit model entails a normal error specification, so that choice probabilities must be calculated using numerical methods, rather than the standard logit formula of Eq.(5.16). In this applications, several such methods were used – quasi Monte Carlo (Niederreiter sequences), quadrature (Radau) and a logit-based approximation (as per Amemiya, 1985), which allow direct application of Eq.(5.16) – finding them to agree to at least five significant digits.

design vastly more complex products (for example, Kim *et al*., 2002); in fact, one of its chief virtues is that it can, through hierarchical decomposition, accommodate products consisting of numerous interacting subsystems. That is, the method illustrated here seamlessly 'scales' to far more complex products, so long as their designs are appropriately codified. Next, the marketing planning and engineering design subproblems are defined, after which they are formally linked via ATC.

5.4.1 Marketing Subproblem

The same product characteristic discretization and conjoint survey data described in Chapter 4 are used here to fit each of the three demand models. Exogenous parameters were set as follows: c^{V} = \$3 per unit, c^{I} = \$3 million for initial investment, and market size S = 5 million, the approximate yearly US dial-readout scale market. Being completely exogenous, these values are easily altered, and in fact the entire demand specification can be. Finally, the marketing subproblem in Figure 5.2 is formulated as in Eq.(5.3), with the demand model specified in Eq.(5.6) and Eq.(5.16) and using functions constructed through the part-worth coefficients β obtained from the conjoint choice data Φ.

5.4.2 Engineering Design Subproblems

The same engineering design variables, constraint functions, and response functions from Chapter 4 are used here, and a copy of the engineering model is constructed for each product j, formulated as in Eq.(4.1), where the response functions $\mathbf{r}$ and constraint functions $\mathbf{g}$ are given in Chapter 4, and no equality constraints $\mathbf{h}$ exist in this example. This completes the engineering design model. While the specifics of such a model necessarily differ across various product types, the method by which the model was developed is general.

5.5 RESULTS

There are two main components to the approach advocated here: econometric, for the extraction of decomposed individual-level preferences and generation of the preference splines, and optimization-based, for the determination of the best number of products and their positionings conditional on the preference splines. These are each examined in turn.

5.5.1 Demand Model Results

Table 5.2 lists DIC results for the normal mixture model and classical likelihoods for the discrete mixture and homogeneous models. The continuous heterogeneity models appear up through the two-segment solutions, which were chosen by DIC, while the latent class models appear up through 7 segments, at which point BIC topped out.

Table 5.2: Comparison of heterogeneity specifications: discrete latent class vs. HB random parameters

Discrete Heterogeneity (Classical Latent Class)			
Segments	**df**	**Classical LL**	**BIC**
1	25	-10983	22194
2	51	-10239	20944
3	77	-9784	20271
4	103	-9537	20014
5	129	-9336	19850
6	155	-9187	19788
7	**181**	**-9059**	**19770**
8	207	-8948	19785
Continuous Heterogeneity (HB Random Parameters)			
Segments	**Cov(Beta)**	**Classical LL***	**DIC**
1	Diagonal	-3813	12432
2	Diagonal	-3713	12073
3	**Diagonal**	**-3656**	**11961**
4	Diagonal	-3638	12029
1	Full	-4051	11742
2	**Full**	**-4016**	**11674**
3	Full	-4017	11745

It is plainly apparent that: (1) continuous heterogeneity (normal mixure) alone is superior[8] to discrete heterogeneity (latent class) alone, up through even a fairly large number of segments (as per Allenby and Rossi 2003); (2) a correlated (random) coefficients specification for the normal mixture model is superior to an uncorrelated one; and (3) more than one segment in the normal mixture model is supported. In short, the most general specification fares best, and each of its attributes – correlated coefficients, and both discrete and continuous heterogeneity – is useful in presenting an accurate representation of consumer preferences. In the following sections, this 'full' model is referred to primarily, and others are called upon only peripherally for comparison purposes.

As described at length earlier, the hyperparameters in the hierarchical set-up were estimated using the Bayesian MCMC chain and used to generate a large number (1000) of part-worth coefficient vectors ($\boldsymbol{\beta}_i$), which serve to stochastically integrate over the posterior choice surface. Posterior means of these resulting $\boldsymbol{\beta}_i$ values are listed in Table 5.3, and the resulting splines are shown graphically in Figure 5.4, along with mean values for the discrete mixture and homogeneous cases; recall that for identification purposes these values are scaled so that the sum in each set of characteristics is zero, making for easier visual comparison. There are several things to note here. First, the listed mean beta values are generated from a mixture distribution, and as such do not correspond to any model parameters directly, but to weighted averages of them. Second, a comparison of the values used in the optimization (1000 generated values from the posterior mixing distribution) to the averaged posterior Betas for our actual (n = 184) respondents indicates very close agreement; the greatest deviation is less than 0.06, small relative to

[8] The classical homogeneous, latent class and Bayesian continuous heterogeneity representations are informally compared here using classical LL values (calculated at the posterior mode for the Bayesian models). These differences are dramatically in favor of the continuous heterogeneity specifications, far more so than can be attributed to posterior uncertainty (as captured by DIC).

the overall range of Beta values. Finally, in each of the six attribute spline graphs, the heterogeneous model is most 'arched' or highly sloped, suggesting the presence of some consumers with relatively strong preference differentials across attribute levels (it should be emphasized that part-worth values have a nonlinear mapping onto choice probabilities, and hence demand, so an "averaged part-worth" is only a rough guide to comparing across heterogeneity specifications).

Table 5.3: Average part-worth coefficient beta values with each characteristic zero-mean scaled

		homo-geneous	discrete mixture	mixture of normals
z_1	200 lbs.	-0.5295	-0.6944	-1.1888
	250 lbs.	0.1331	0.1449	0.2519
	300 lbs.	0.2321	0.2782	0.4539
	350 lbs.	0.1081	0.1505	0.2783
	400 lbs.	0.0562	0.1210	0.2047
z_2	6/8	-0.0543	-0.1009	-0.1406
	7/8	0.2570	0.3157	0.4748
	8/8	0.2816	0.3974	0.5007
	8/7	-0.0211	0.0056	0.0362
	8/6	-0.4632	-0.6181	-0.8712
z_3	100 in.^{2}	0.0192	0.0531	-0.0431
	110 in.^{2}	-0.0940	-0.0198	-0.0480
	120 in.^{2}	0.0533	0.0566	0.1186
	130 in.^{2}	0.0508	0.0034	0.0998
	140 in.^{2}	-0.0293	-0.0937	-0.1273
z_4	2/32 in.	-0.3622	-0.7065	-0.6915
	3/32 in.	-0.1595	-0.1240	-0.1967
	4/32 in.	0.2191	0.3332	0.3806
	5/32 in.	0.1983	0.2580	0.3126
	6/32 in.	0.1044	0.2389	0.1950
z_5	0.75 in.	-0.7402	-0.8462	-1.4291
	1.00 in.	-0.1941	-0.2500	-0.2604
	1.25 in.	0.2394	0.2792	0.4555
	1.50 in.	0.2947	0.3805	0.6200
	1.75 in.	0.4002	0.4366	0.6140
p	\$10	0.7227	0.9682	1.5863
	\$20	0.4862	0.6802	1.1556
	\$30	0.0583	0.2023	0.4637
	\$40	-0.3636	-0.4697	-0.5490
	\$50	-0.9036	-1.3815	-2.6567
no choice		0.0247	-0.0387	0.8366

A number of trends are apparent across these mean estimated coefficient values. Unsurprisingly, price appears to exert the strongest influence, and is decisively downward-sloping (this is true of the posterior means for each of the n = 184 original participants). One might have expected similarly monotonic preferences for number size and weight capacity, but this is only true for the former; apparently, too high a capacity was viewed as 'suboptimal' by the respondents, on average. Note that these beta values reflect pure consumer preference, and *not* any sort of constraint resulting from infeasible designs, which can only arise from the engineering design submodel. Preferences for the other three variables (platform area, aspect ratio (i.e., shape) and interval mark gap) all have interior maxima.

5.5.2 Product Line Optimization Results

Conditional on the generated splines arising from the HB conjoint estimates (using the mixture-of-normals-based demand model), the engineering design and marketing subproblems are solved iteratively until convergence. Optimization was carried out in MATLAB, with each subproblem solved using the sequential quadratic programming method (Papalambros and Wilde 2000), an efficient gradient-based algorithm. The ATC system is solved for a fixed product line size J, and a parametric study is performed to determine the value of J that produces the most profitable overall product line (i.e., solutions are found for cases $J = \{1, 2, ...\}$, and the solution that produces the highest profit is chosen). As is typical, local optima are generated; global optima can only arise using multi-start. Figure 5.5 shows the best resulting profit across several local minima (found using ten runs with random starting points) for each case $J = \{1, 2, ..., 7\}$. It is clear that a product line with four products is most profitable, and Table 5.4 details this solution.

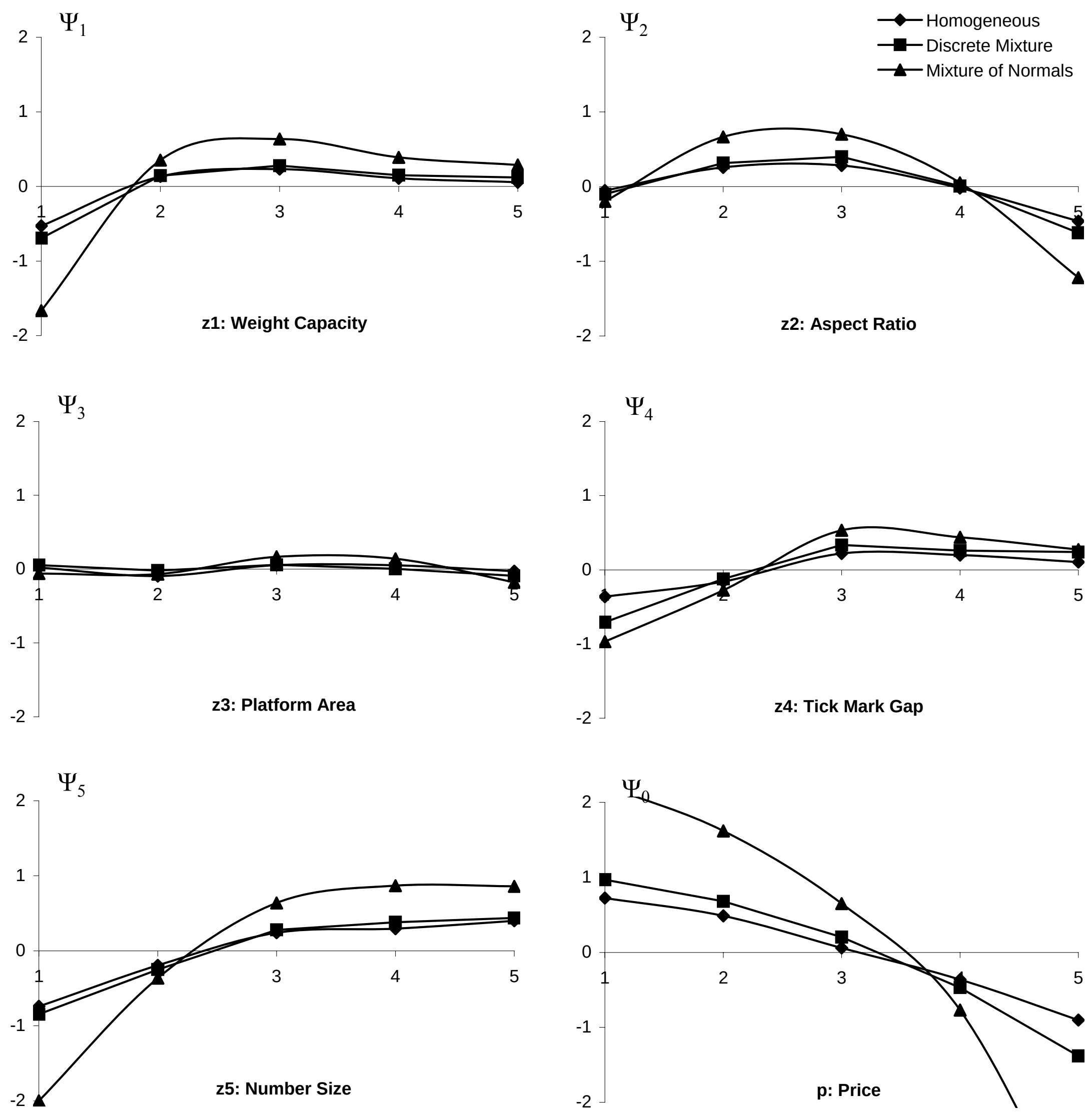


Figure 5.4: Plots of the average splines for each product characteristic and price under the three demand models

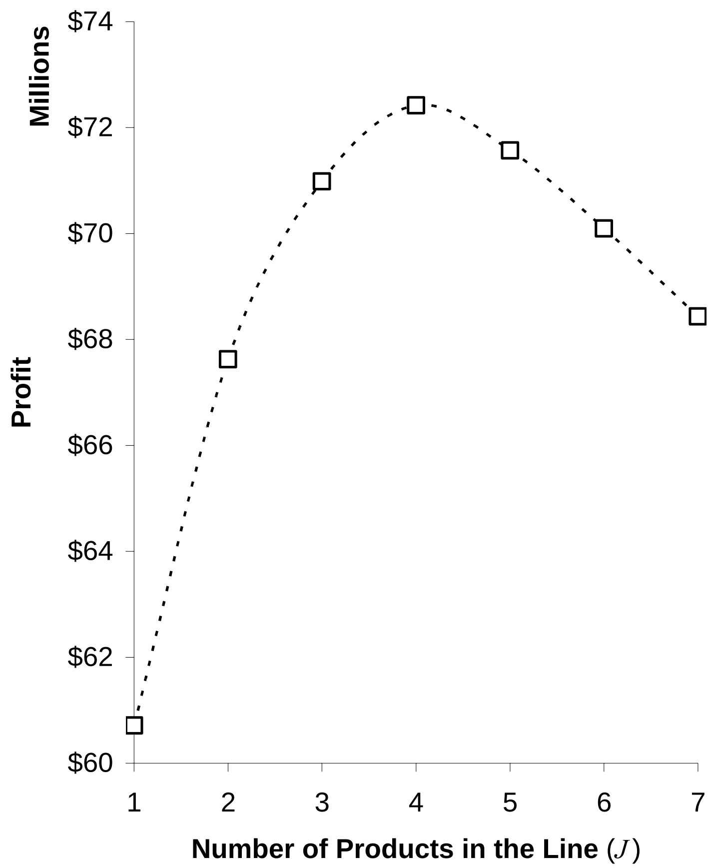


Figure 5.5: Resulting profit as a function of the number of products in the line

Note that several of the resulting scale designs are bounded by active engineering design constraints; this is necessary to ensure that the scale is physically tenable, for example, that the dial, spring plate and levers can be accommodated in the case. Note as well that all the scales in the line lie well within the range available through online retailers. Looking across the table, and considering primarily marketing attributes, one might term the resulting products "large square scale" (27.4% of the market), "large-number portrait scale" (21.0%), "small, low capacity landscape scale" (18.6%) and "high-priced, middle-of-the-road" scale (11.3%).

Table 5.4: Optimal scale designs for the heterogeneous market

				Product Line (Four Products) 1	2	3	4
	Π	Profit (Millions)	$	$72,421,000			
		Market share	%	27.4%	21.0%	18.6%	11.3%
Mkt. Decision Variables	z_1	Weight capacity	lbs.	292	262	200	255
	z_2	Aspect ratio	- - -	0.980	1.156	0.921	0.986
	z_3	Platform area	in.^{2}	140	122	105	135
	z_4	Tick mark gap	in.	0.103	0.116	0.121	0.116
	z_5	Number size	in.	1.221	1.351	1.293	1.331
	p	Price	$	$22.89	$24.53	$23.84	$30.00
Engineering Design Variables	x_1	Lever dimension	in.	8.53	8.62	6.42	7.28
	x_2	Lever dimension	in.	2.82	3.17	3.21	4.20
	x_3	Lever dimension	in.	1.95	9.72	24.00	15.75
	x_4	Lever dimension	in.	0.13	0.65	4.32	1.95
	x_5	Lever dimension	in.	2.82	1.92	3.18	2.47
	x_6	Spring constant	lb./in.	200.0	200.0	2.7	48.7
	x_7	Spring location	in.	0.50	0.50	0.50	0.50
	x_8	Rack length	in..	7.07	4.40	3.41	3.89
	x_9	Pinion size	in.	0.25	0.29	12.02	1.91
	x_{10}	Pivot dimension	in.	0.50	0.80	1.43	0.54
	x_{11}	Pinion dimension	in.	0.64	1.43	1.90	1.78
	x_{12}	Dial diameter	in.	9.52	9.67	7.66	9.35
	x_{13}	Cover length	in.	11.72	11.87	9.86	11.55
	x_{14}	Cover width	in.	11.95	10.27	10.70	11.71
Engineering Constraint Functions	g_1	Dial fits in base	in.	-1.83	0.00	-2.44	-1.75
	g_2	Dial fits in base	in.	0.00	0.00	0.00	0.00
	g_3	Lever fits in base	in.	-8.17	-8.70	-1.75	-6.54
	g_4	Lever joint location	in.	0.00	-1.25	-0.02	-1.73
	g_5	Rack fits in base	in.	-1.81	-3.85	-2.35	-3.69
	g_6	Sufficient rack length	in.	-3.93	-1.49	-2.13	-0.88
	g_7	Lever fits in base	in.	-16.17	-0.38	-9.50	-8.41
	g_8	Sufficient lever length	in.	-0.03	-6.99	0.00	-6.43

It is important to emphasize again that the design space does *not* map one-to-one with characteristics communicated to consumers. This comes about because the engineering design model specifies some product characteristics as functions of interactions between design variables. To take one example, weight capacity (z_1) can be achieved by adjusting the relative values of the pivot lever dimensions (x_{10}, x_{11}), lever lengths (x_1+x_2, x_3+x_4), spring constant (x_6), force placement (x_1, x_3), joint position (x_5) and pinion gear pitch diameter (x_9). The resulting 'iso-weight-capacity' surface is highly non-linear, but the key point is that an infinite number of design solutions can appear equivalent in terms of what can be conveyed about them to consumers; that is, multiple designs may exhibit identical product characteristics. A manager could enact any number of criteria *post hoc* to choose from among such a continuum, or detailed cost data and preferences for commonality could drive selection of a single engineering design among the set of possibilities, although such strategies are not pursued here since the primary intention is to focus on ensuring existence of a feasible design that attains the preferred product characteristic targets.

In the following sections ATC product line coordination results are compared to an alternative disjoint approach, and the effects of heterogeneity representations on product line solutions are explored.

5.5.3 Effectiveness of ATC Coordination

A major contribution of the methodology presented here is to provide rigorous coordination between marketing and engineering models in order to find a joint solution that is optimal under consideration of both customer preferences and engineering feasibility. As in Chapter 4, to demonstrate the importance of this coordination, the coordinated ATC solution was compared to the solution obtained through a disjoint sequential approach, which has previously been referred to as Analytical Target Setting (Cooper *et al*. 2003). In the disjoint scenario, marketing sets price and product characteristic targets based on consumer preference data without engineering feasibility information (marketing subproblem) and passes these targets to engineering design teams. The engineering teams then design feasible products that meet the targets as closely as possible (engineering subproblems) without further iteration, leaving price as the only marketing variable that can be changed post-design. In this disjoint scenario, marketing produces a plan for a line of four scales with a predicted market share of 83.4% and resulting profit of $81.2 million. There is no reason to believe these products will be feasible, as they are based solely on consumer preferences, which often involve contradictory characteristics. That is, when one looks beyond consumer preferences and considers only the space of realizable products, product lines and their profit potentials can change substantially. In the disjoint case, these (unachievable) targets are passed to engineering teams who each design a feasible product to achieve product characteristics as close as possible to the targets requested by marketing without further iteration. The resulting products differ significantly from the initial plan and therefore have characteristics that are less preferred by consumers, resulting in an actual market share of 70.5% and profit of $67.9 million: 16% less than marketing's original (unachievable) prediction. If, instead, the ATC process is used to iteratively coordinate marketing profit and preferences with engineering performance and feasibility, the resulting joint solution is a line of four different products, resulting in 78.2% market share and $72.4 million profit. In this case, coordination resulted in a feasible product line with 6% higher profitability than that resulting from disjoint decision-making. These results are summarized in Table 5.5.

Table 5.5: ATC coordination vs. disjoint decision-making

	Resultant Line Market Share	Resultant Profit
Best line targets predicted by marketing (infeasible)	83.4%	$81.2
Feasible engineering designs closest to marketing targets	70.5%	$67.9
ATC coordination to produce a joint solution	78.2%	$72.4

It is worth noting that in this case the original marketing plan and the ATC coordinated solution both contain four products in the line, but, in general, these two cases need not yield lines of the same size, and in many cases marketing will have a tendency to overestimate the optimal line size *J* because of an overoptimistic assumption about how much "coverage" of the product characteristic space is possible. Also, in this disjoint scenario marketing "leads" by developing the original plan and engineering design "follows" by attempting to meet product characteristic targets. The reverse situation, where engineering "leads," is possible when all consumers have monotonic preferences for product characteristics (e.g.: fuel economy and performance for a vehicle) by first having engineering design a set of Pareto optimal products and then allowing marketing to pick a line from that set of products (see Li and Azarm, 2002). However, in this example, preferences for characteristics are non-monotonic, so no such common Pareto set exists, and without preference information, engineering has no well-defined optimization objective.

5.5.4 Heterogeneity Representation

The methodology presented here has several independent components. As such, it is reasonable to ask whether one or another might be made less sophisticated or sidestepped entirely, with little change in substantive outcome. Specifically, what might be the profit implications of doing so? While it is true that the methodology is modular – ATC can be applied, for example, with a different demand formulation or another form of heterogeneity correction – it is valuable to ascertain whether the modeling choices made within that modular framework are warranted, and how strongly. Although it is not the main focus of this research, it is instructive to consider whether similar optimal line designs can be obtained through simpler forms of heterogeneity modeling. Using the simple homogeneous model of demand is pointless for generating product lines because the IIA property leads to product lines with duplicate products, so the discrete mixture model, or latent class model, is compared with the normal mixture model. Since the discrete mixture model is natively supported in many statistical packages, it might prove convenient for line optimization. Though fit statistics (Table 5.2) alone argue that the discrete mixture model is dramatically inferior to the normal mixture specification, this does not necessarily mean that, *conditional* on the resulting estimates, the resulting optimal line will be dramatically inferior.

Table 5.6 lists a comparison between the resulting profitability (evaluated *post hoc* with the normal mixture model) of the optimal solutions found using the discrete and continuous mixture approaches for optimization. Not only do these cases result in different product line solutions (a line of six products under the discrete mixture model vs. a line of four products under the normal mixture model), but the solution deriving from the discrete mixture model suffers a decrement in profit of 18.4%[9]. Furthermore, the discrete mixture specification results in a profit surface with more numerous and more pronounced local minima, including local minima with duplicate products such as the solution found in Table 5.4, acting to impede the optimization process. Thus, even a relatively sophisticated heterogeneity representation can offer very different, and potentially sub-optimal, product line results.

Table 5.6: Optimal single-product and product line solutions under each demand specification

		Single Product Solutions			Product Line Solutions									
					Discrete Mixture						Normal Mixture			
		Homo-geneous	Discrete Mixture	Normal Mixture	1	2	3	4	5	6	1	2	3	4
Π Profit (Millions)*	$	$54.1	$58.3	$60.7	$59.1						$72.4			
Market share*	%	48.8%	57.8%	65.0%	25.1%	8.7%	8.7%	8.7%	6.9%	4.9%	27.4%	21.0%	18.6%	11.3%
z_1 Weight capacity	lbs.	255	254	256	238	257	257	257	253	248	292	262	200	255
z_2 Aspect ratio	- - -	0.996	1.047	1.002	1.045	1.041	1.039	1.039	1.062	1.051	0.980	1.156	0.921	0.986
z_3 Platform area	in^2	134	127	130	100	131	131	131	123	114	140	122	105	135
z_4 Tick mark gap	in.	0.116	0.117	0.115	0.106	0.116	0.116	0.116	0.114	0.111	0.103	0.116	0.121	0.116
z_5 Number size	in.	1.334	1.339	1.315	1.193	1.341	1.337	1.337	1.316	1.268	1.221	1.351	1.293	1.331
p Price	$	$26.41	$24.21	$22.61	$23.96	$30.00	$30.00	$30.00	$30.00	$29.37	$22.89	$24.53	$23.84	$30.00

[9] The profit of the line generated under the discrete mixture model was evaluated for profitability *post hoc* using the normal mixture model , which has superior properties.

While it may be somewhat expected that more restricted representations of heterogeneity can lead to suboptimal product lines, it is common practice to presume that a homogeneous model is sufficient for the design of a single product. The results for this case study suggest that this assumption must be made with care. Table 5.6 lists the solutions when one performs optimization for a single product under the three demand model scenarios. In this case the more restrictive models do a fairly good job predicting the optimal product characteristics, although this need not always be true; however, the price that can be fetched for the product is significantly overestimated by the more restrictive models, resulting in a loss of 7% market share using the discrete mixture model and a loss of 14% market share using a simple homogeneous model, relative to the normal mixture model. These results suggest that the issue of how heterogeneity specification affects contingent *optimization* results is worthy of further study on its own.

5.6 CONCLUSIONS

Product-producing firms work to plan and design lines of products that best suit their market and profitability goals. Different functional entities within the firm can interpret this imperative idiosyncratically: suiting consumer needs for marketers, maximizing performance at minimum cost within technological constraints for engineers. Considered independently, these goals often lead to conflict, both in practice and with respect to optimization models in each discipline, and disjoint sequential “throw it over the wall” approaches to resolving the conflict can lead to suboptimal decision-making.

In this chapter, the ATC methodology was extended to iteratively coordinate models from each discipline and arrive at joint optima for a marketplace-viable, technologically-feasible product line. The decomposition and coordination approach offered by ATC allows such joint solutions to be achieved while maintaining disciplinary modeling focus and relieving the need for a single modeler to become an expert in all areas. For example, marketers need not explicitly enumerate the set of attainable product characteristic combinations or constrain survey design to these combinations, difficult goals when at least some product characteristics are defined on a continuous space. Likewise, engineering designers need not develop artificial representations to describe tradeoff preferences among competing performance objectives or deal directly with differences in preference points across the population of users. Instead, the methodology presented offers a rigorous definition of the interaction between the two modeling scopes, and modelers in each discipline can focus on issues relevant to their expertise. The iteration of these decision-models then acts to reduce the need for more costly human iteration.

The ATC-based analysis presented here offers a number of concrete conclusions for the product line studied. More research is warranted to determine the extent to which these carry over to other types of products. Nevertheless, several ‘main points’ emerged with clarity:

1. Consideration of engineering feasibility during marketing product planning can lead to significantly different designs with different characteristics than those obtained through a pure marketing analysis. ATC allows only *feasible* product lines by iteratively coordinating marketing product characteristic targets with actual product characteristics achieved by a feasible engineering design, and this coordination can significantly improve the profitability of the resulting product line.
2. In product design optimization, the form of heterogeneity matters: The normal mixture model HB approach was far superior at representing the underlying preferences leading to the observed choices than the homogenous and discrete mixture representations, and use of a more restrictive model can lead to different solutions with significant reductions in profitability.

3. Multiple products are more profitable and better able to satisfy preferences in a heterogeneous market than single products, unless fixed costs are prohibitive. Even if a firm is designing only a single product, when that product is optimized for a heterogeneous market it can differ from one optimized for a presumed-homogeneous market calibrated on the same consumer preference data.
4. The results do not advocate four different consumer products – there are clearly far more in the market – but four different product *designs*: Other elements (e.g., color, packaging) can be altered on the fly independently from the engineering design. This is an important observation to keep in mind when using the proposed methodology.

A chief strength of the ATC formulation is its intrinsic modularity: Each component can be readily modified without having to mathematically re-tool the entire system. Incorporating other forms of heterogeneity, for example, will affect only the demand model, and changing the product topology model or including multiple product topologies is also straightforward in that one must adjust only the engineering design sub-model. ATC can also accommodate more complex hierarchies to organize decision-making in marketing or engineering, which is particularly useful for scaling to large problems such as automobile production, where numerous systems and subsystems interact. While relatively simple cost and profit models are used in the case study, with several important qualities specified exogenously, individual models can be altered with any number of well-known refinements. For example, Chapter 6 introduces detailed manufacturing planning models and coordinates them with cost and production volume targets set by marketing.

A number of relevant constructs could serve to provide other useful extensions to the model presented here, and several of these are discussed below. First, the account provided here is 'normative,' in that certain economic, technological and perceptual forces observed in practice are not modeled directly. A number of authors have pointed out forces that can affect the normative account of line optimization arising from conjoint analysis and various operations research-based search methods. Moorthy (1984), based on consumer self-selection and third-degree price discrimination, shows how various consumer segments might be profitably aggregated even in the absence of economies of scale. Kekre and Srinivasan (1990) suggest that, whereas manufacturing seeks to limit product line depth to ensure a smooth production flow, marketing seeks to satisfy as many consumer segments as practicable. Villas-Boas (1998) underscores the role of customer communication in optimal line design, showing that certain costs, in particular advertising, can lead to firms' offering fewer product variants in their lines than they otherwise should. Using a related framework, Villas-Boas (2003) demonstrates that, because retailers, not manufacturers, control the channel and ultimate product targeting, manufacturers may need to increase line length to help appropriately segment the market. In our approach, as in that of nearly all prior research in the area, retailers play a limited direct role, and various communication costs are considered known and stable.

Second, several other studies have incorporated elements that have been assigned a peripheral role here. Perhaps the most prominent of these is likely competitive reaction. A pioneering study in this regard is Kadiyali, Vilcassim and Chintagunta (1996), who adopted a "new empirical industrial organization" (NEIO) approach to determine optimal policies for competitive product line pricing. A related formulation is that of explicit competitive reaction models using game theory. For single brand firms, Choi and DeSarbo (1993) and Green and Krieger (1997) have illustrated how to derive competitive strategies using conjoint analysis via Nash equilibrium. Chapter 7 discusses a method for using game theory to predict the engineering design decisions of competing automotive firms with respect to measured customer preferences under several government regulation scenarios.

Third, models of market expansion as a function of the product line's depth or composition could enrich model predictions over the simple 'outside good' methodology used here. Bayus and Putsis (1999), for example, outline strategic and cost-based consequences of various complex market factors on product line composition. Shugan and Desiraju (2001) consider the effects of perpetual new variants within and deletions from a product line, and how retailers should react in the many product classes where this is the norm. They find that it is often optimal to decrease prices on an entire line, for example, when costs fall disproportionately for higher-quality products. In this chapter, as is traditional in product line optimization models, product lines are taken to be stable across the planning horizon.

Finally, consideration of how the structure and depth of a product line impacts the equity of its brands or the line itself could provide further insight (Randall, Ulrich and Reibstein 1998). Recent work in this area by Draganska and Jain (2004) examines product line length as arising from just this sort of endogenous system. Here, the conjoint data is used to describe the market "as is." It would be worthwhile to extend the proposed modeling framework to allow for the likely market restructuring arising from just this sort of endogenous introduction.

In closing, several maxims are relevant for both the marketing and engineering design/production communities. First, although engineers are keenly aware at every step of their work of real, inviolable constraints, marketers tend to work to find desirable product characteristic targets for exploring new markets. A tacit belief is that most, if not all, design constraints can be vanquished by ingenuity or sufficient capital. While this is sometimes true, often it is not. ATC encodes non-negotiable technological infeasibilities directly into its conceptual foundations. As such, marketers using their own models within an overarching ATC formulation can gain a "gut feel" for what will work, and what will not, in terms of actual, deliverable products, to supplement their intuitive understanding of the consumer marketplace. The flip side is that engineers can come to terms with 'consumer space,' every bit as real as the geometry and physics underlying their own models, and resolve tradeoffs among competing performance goals through coordination with marketing. Second, while it may appear simple to specify which product characteristic combinations cannot co-exist, in practice it is often impractical: This 'infeasible hull' can snake through the product characteristic space in ways difficult to visualize or translate into meaningful consumer terms. ATC frees marketers from considering such issues when collecting consumer preferences; iterative coordination readily avoids infeasible product line configurations *post hoc*. Third, heterogeneity matters: it must be accounted for in sufficient generality. And finally, the question arises of whether ATC, this newcomer to the enterprise modeler's toolbox, can be trusted alongside mainstays like conjoint analysis and discrete choice modeling. Although the best verdict is always that of posterity, ATC is proven, for a broad class of problems, to converge to joint optimality across its various subsystems. As such, it can literally guarantee better profitability – or sales, or consumer surplus, or indeed any quantity of managerial interest – than one could arrive at by optimizing the engineering design and marketing submodels independently. With its scalability, relatively low computational overhead, and ability to key into a wide variety of extant modeling techniques, ATC offers strong reasons for adoption as a cross-disciplinary platform for the design of complex products and product lines.

5.7 REFERENCES

[5.1] Allenby, Greg M. and Peter E. Rossi (2003) "Perspectives based on 10 years of HB in marketing research," *Sawtooth Software Conference Proceedings*, (forthcoming).

[5.2] Amemiya, T. (1985) *Advanced Econometrics*, Harvard University Press, Cambridge.

[5.3] Andrews, R.L., A. Ainslie and I.S. Currim (2002) "An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity," *Journal of Marketing Research*, v39 p479-487.

[5.4] Andrews, R.L., A. Ansari and I.S. Currim (2002) "Hierarchical Bayes versus finite mixture conjoint analysis models: a comparison of fit, prediction, and partworth recovery," *Journal of Marketing Research*, v39 n1 p87-98.

[5.5] Aptech (2000) "GAUSS language reference," Aptech Systems, Inc., Maple Valley.

[5.6] Balakrishnan, P. V. and Jacob V. S. (1996) "Genetic algorithms for product design," *Management Science*, v42 n1 p1105-1117.

[5.7] Balling, R.J. and Sobieszczanski-Sobieski (1994) "Optimization of coupled systems: a critical overview of approaches," *Proceedings of the 5th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization,* AIAA-94-4330-CP, Panama City, p753-773.

[5.8] Bayus, B., and W. Putsis (1999) "Product proliferation: an empirical analysis of product line determinants and market outcomes," *Marketing Science*, v18 n2 p137-53

[5.9] Berry, Steven (1994) "Estimating discrete choice models of product differentiation," *RAND Journal of* Economics, v25 p242–262.

[5.10] Bohl, Patrick (1997) "Conjoint literature database CLD," University of Mainz, Germany, at *http://www.uni-mainz.de/~bohlp/cld.html*.

[5.11] Braun, R. (1996) *Collaborative optimization: An Architecture for Large-Scale Distributed Design*, PhD Dissertation, Stanford University.

[5.12] Brown, Shona L. and Kathleen M. Eisenhardt (1995) "Product development: past research, present findings, and future directions," *Academy of Management Review*, v20 (April) p343-378.

[5.13] Cattin, Philippe and Dick R. Wittink (1982) "Commercial use of conjoint analysis: a survey," *Journal of Marketing*, v46 (Summer) p44-53.

[5.14] Chen, K. D. and Hausman W. H. (2000) "Technical note: mathematical properties of the optimal product line selection problem using choice-based conjoint analysis," *Management Science*, v46 n2 p327-332.

[5.15] Choi, S. C. and DeSarbo, W. S. (1993) "Game theoretic derivations of competitive strategies in conjoint analysis," *Marketing Letters*, v4 p337-348.

[5.16] Choudhary, R. (2004) *A Hierarchical Optimization Framework for Simulation-Based Architectural Design*, PhD Dissertation, University of Michigan.

[5.17] Cooper, A.B., P. Georgiopoulos, H.M. Kim and P.Y. Papalambros (2003) "Analytical target setting: an enterprise context in optimal product design," *Proceedings of the 2003 ASME Design Automation Conference*, DAC-48734, Chicago, Illinois, USA, September 2-6.

[5.18] De Boor (2001) *A Practical Guide to Splines* (revised ed.), Springer, New York.

[5.19] Diaconis, P. and L. Saloff-Coste (1998) "What do we know about the Metropolis algorithm?" *Journal of Computer and System Sciences*, v57 n1 (August) p20-36.

[5.20] Dobson, G. and Kalish, S. (1988) "Positioning and pricing a product line," *Marketing Science*, v7 n2 p107-125.

[5.21] Dobson, G. and Kalish, S. (1993) "Heuristics for pricing and positioning a product-line using conjoint and cost data," *Management Science*, v39 n2 p160-175.

[5.22] Draganska, M. and Jain, D. (2004) "Product line length as a competitive tool," working paper, Stanford Graduate School of Business.

[5.23] Fellini, R. (2003) *A Model-Based Methodology for Product Family Design*, Ph.D. Dissertation, Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan, USA

[5.24] Genz, Alan (2004) "Numerical computation of rectangular bivariate and trivariate normal and t probabilities," *Statistics and Computing*, v14 p151-160.

[5.25] Green, P. E. and Krieger, A. M. (1985) "Models and heuristics for product line selection," *Marketing Science*, v4 p1-19.

[5.26] Green, P. E. and Krieger, A. M. (1987a). "A consumer-based approach to designing product line extensions," *Journal of Product Innovation Management*, v4 p21-32.

[5.27] Green, P. E. and Krieger, A. M. (1987b) "A simple heuristic for selecting 'good' products in conjoint analysis," *Applications of Management Science*, v5 p131-153.

[5.28] Green, P. E. and Krieger, A. M. (1989) "Recent contributions to optimal product positioning and buyer segmentation," *European Journal of Operational Research*, v41 p127-141.

[5.29] Green, P. E. and Krieger, A. M. (1992) "An application of a product positioning model to pharmaceutical products," *Marketing Science*, v11 p117-132.

[5.30] Green, P. E. and Krieger, A. M. (1996) "Individualized hybrid models for conjoint analysis," *Management Science*, v42 n6 p850-867.

[5.31] Green, P. E. and Krieger, A. M. (1997) "Using conjoint analysis to view competitive interaction through the customer's eyes." In: Day, G. S. and Reibstein, D. J. (eds.). *Wharton on dynamic competitive strategy*, Wiley & Sons, New York, p343-368.

[5.32] Green, P. E. and Rao, V. R. (1971) "Conjoint measurement for quantifying judgmental data," *Journal of Marketing Research*, v8 p355-363.

[5.33] Green, Paul E. and V. Srinivasan (1990) "Conjoint analysis in marketing: new developments with implications for research and practice," *Journal of Marketing*, v54 n4 p3-19.

[5.34] Green, Paul E., J. Wind and V. R. Rao (1998) "Conjoint analysis: methods and applications," in *Handbook of Technology Management*, R. Duff (ed.). Boca Raton, FL: CRC Press.

[5.35] Griffin, Abbie (1997) "The effect of project and process characteristics on product development cycle time," *Journal of Marketing Research*, v34 (February) p24-35.

[5.36] Griffin, Abbie and John R. Hauser (1992) "Patterns of communication among marketing, engineering and manufacturing – a comparison between two new product teams," *Management Science*, v38 (March) p360-373.

[5.37] Gu, X., J.E. Renaud, L.M. Ashe, S.M. Batill, A.S. Budhiraja, and L.J.Krajewski (2002) "Decision-based collaborative optimization," *Journal of Mechanical Design*, v124 p1-13.

[5.38] Gupta, S. and A. Samuel (2001) "Integrating market research with the product development process: a step towards design for profit," *Proceedings of the 2001 ASME DETC*, DFM-21202.

[5.39] Haaijer, R. (1999) "Modeling conjoint choice experiments with the probit model," Labyrint Publications, at *http://www.ub.rug.nl/eldoc/dis/eco/m.e.haaijer/*.

[5.40] Haaijer, R., Kamakura, W. and Wedel, M. (2001) "The 'no-choice' alternative in conjoint choice experiments," *International Journal of Marketing Research*, v43 n1 p93-106.

[5.41] Haaijer, R., M. Wedel, M. Vriens and T.J. Wansbeek (1998) "Utility covariances and context effects in conjoint MNP models," *Marketing Science*, v17 p236-252.

[5.42] Hajivassiliou, V., McFadden D., and Ruud P. (1996) "Simulation of multivariate normal rectangle probabilities and their derivatives: theoretical and computational results," *Journal of Econometrics*, v72 p85-134 .

[5.43] Hauser, John R. (2001) "Metrics thermostat," *Journal of Product Innovation Management*, v18 n3 p134-153.

[5.44] Hazelrigg, G.A. (1988) "A framework for decision-based engineering design," *ASME Journal of Mechanical Design*, v120 p653-658.

[5.45] Hibbeler, R.C. (1993) *Statics and Mechanics of Materials*, Macmillan Publishing, New York.

[5.46] Huber, J., and K. Train (2001) “On the similarity of classical and Bayesian estimates of individual mean partworths,” *Marketing Letters*, v12 n3 p257-267.

[5.47] Kadane J.B. and N.A. Lazar (2004) “Methods and criteria for model selection,” *Journal of the American Statistical Association*, v99 p279-290.

[5.48] Kadiyali, Vrinda, Naufel Vilcassim, and Pradeep Chintagunta (1996) ‘‘Empirical analysis of competitive product line pricing decisions: lead, follow or move together?’’ *Journal of Business*, v9 p459-87.

[5.49] Kalyanam, K. and T.S. Shively (1998) “Estimating irregular pricing effects: a stochastic spline approach,” *Journal of Marketing Research*, v35 p16-29.

[5.50] Kamakura, W.A. and G.J. Russell (1989) “A probabilistic choice model for market segmentation and elasticity structure,” *Journal of Marketing Research*, v26 p379-390.

[5.51] Kaul, Anil and Vithala Rao (1995) “Research for product positioning and design decisions: an integrative review,” *International Journal of Research in Marketing*, v12 p293-320.

[5.52] Keane, M.P. (1992) “A note on identification in the multinomial probit model,” *Journal of Business and Economic Statistics*, v10 p193-200.

[5.53] Kekre, S. and K. Srinivasan (1990) “Broader product line: a necessity to achieve success?” *Management Science*, v36 n10 p1216-1231.

[5.54] Kim, H. M. (2001) *Target Cascading in Optimal System Design*, Ph.D. Dissertation, Dept. of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan.

[5.55] Kim, H. M., M. Kokkolaras, L. Louca, G. Delagrammatikas, N. Michelena, Z. Filipi, P.Y. Papalambros, and D. Assanis (2002) “Target cascading in vehicle redesign: a class VI truck study,” *International Journal of Vehicle Design*, v29 n3 p199-225.

[5.56] Kim, H.M., D.G. Rideout, P.Y. Papalambros, and J.L. Stein, (2003b) “Analytical target cascading in automotive design,” *ASME Journal of Mechanical Design*, v125 p481-489.

[5.57] Kim, H.M., N.F. Michelena, P.Y. Papalambros, and T. Jiang (2003a), “Target cascading in optimal system design,” *ASME Journal of Mechanical Design*, v125, p481-489.

[5.58] Kim, J.G., U. Menzefricke and F.M. Feinberg (2004a) “Assessing heterogeneity in discrete choice models using a Dirichlet process prior,” *Rev. Marketing Science*, v2 n1.

[5.59] Kim, J.G., U. Menzefricke and F.M. Feinberg (2004b) "Capturing heterogeneous nonlinear utilities: a Bayesian spline approach," working paper, MIT Sloan School of Management.

[5.60] Kohli, R., Krishnamurti, R. (1987) "A heuristic approach to product design," *Management Science*, v 33 n12 p1523-1533.

[5.61] Kohli, R., Sukumar, R. (1990) "Heuristics for product-line design using conjoint analysis," *Management Science*, v 36 n12 p1464-1478.

[5.62] Kokkolaras, M., R. Fellini, H.M. Kim, N. Michelena, and P. Papalambros (2002) "Extension of the target cascading formulation to the design of product families," *Journal of Structural and Multidisciplinary Optimization*, v14 n4 p293-301.

[5.63] Krishnan, V. and K. Ulrich (2001) "Product development decisions: a review of the literature," *Management Science*, v47 n1 p1-21.

[5.64] Kuhfeld, Warren F. (2003) "Conjoint analysis," SAS Institute Publications, available at *http://support.sas.com/techsup/tnote/tnote_stat.html*.

[5.65] Leeflang, P.S.H, D.R. Wittink, M. Wedel, and P.A. Naert (2000) *Building Models for Marketing Decisions*, Dordrecht/Boston: Kluwer Academic Publishers.

[5.66] Leenders, Mark A.A.M. and Berend Wierenga (2002) "The effectiveness of different mechanisms for integrating marketing and R&D," *Journal Of Product Innovation Management*, v19 n4 p305-317.

[5.67] Lenk, P. and DeSarbo W.S. (2000), "Bayesian inference for finite mixtures of generalized linear models with random effects," *Psychometrika*, v65 n1 p93-119.

[5.68] Lenk, P., DeSarbo W.S., Green P.E. and Young M.R. (1996) "Hierarchical Bayes conjoint analysis: recovery of partworth heterogeneity from reduced experimental designs, *Marketing Science*, v15 n2 p173-191.

[5.69] Li, H. and S. Azarm (2000) "Product design selection under uncertainty and with competitive advantage" *ASME Journal of Mechanical Design*, v122 p411-418.

[5.70] Li, H. and S. Azarm (2002) "An approach for product line design selection under uncertainty and competition," *ASME Journal of Mechanical Design*, v124 p385-392.

[5.71] Lilien, Kotler and Moorthy (1992) *Marketing Models*, Chapter 10, Englewood Cliffs: Prentice-Hall.

[5.72] Louviere J. J. and G. Woodworth (1983) "Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data," *Journal of Marketing Research*, v20 p350-367.

[5.73] Louviere, J. J., D. Hensher, and J. Swait (2000) *Stated Choice Methods: Analysis and Application*, Cambridge University Press, UK.

[5.74] Louviere, J. J. (1988) *Analyzing Decision Making: Metric Conjoint Analysis*, Newbury Park, CA: Sage Publications, Inc.

[5.75] Marton, M. and F. Mistree (1998) "An implementation of expected utility theory in decision based design," *Proceedings of the 1998 ASME DETC*, DTM-5670, p1-11.

[5.76] McAlister, Leigh (1982), "A dynamic attribute satiation model of variety-seeking behavior," *Journal of Consumer Research*, v9 (September) p141-150.

[5.77] McBride, R. D., Zufryden, F. S. (1988) "An integer programming approach to the optimal product line selection problem," *Marketing Science*, v7 n2 p126-140.

[5.78] McCulloch, R.E. and Rossi P.E. (1994) "Exact likelihood analysis of the multinomial probit model," *Journal of Econometrics*, v64 p207-240.

[5.79] Michalek, J.J., F.M. Feinberg, F. Adgüzel, P. Ebbes, and P.Y. Papalambros (2005) "Realizable product line design optimization: coordinating marketing and engineering models via analytical target cascading," in review, *Marketing Science*.

[5.80] Michelena, N., H. Park, and P. Papalambros (2003) "Convergence properties of analytical target cascading," *AIAA Journal*, v41 n5 p897-905.

[5.81] Moorthy, Sridhar (1984) "Market segmentation, self-selection, and product line design," *Marketing Science*, Fall.

[5.82] Nair, S. K., L. S. Thakur, K. Wen, (1995) "Near optimal solutions for product line design and selection: beam search heuristics," *Management Science*, v41 n5 p767–785.

[5.83] Nevo, A. (2000) "Mergers with differentiated products: the case of the ready-to-eat cereal industry," *RAND Journal of Economics*, v31 p395-421.

[5.84] Orme, Bryan K. (1998) "The benefits of accounting for respondent heterogeneity in choice modeling," Sawtooth Software, Inc., at *http://www.sawtoothsoftware.com/download/techpap/benefits.pdf*

[5.85] Papalambros and Wilde (2000) *Principles of Optimal Design: Modeling and Computation*, 2nd ed., Cambridge Univ. Press, New York.

[5.86] Rossi, Peter E. and Greg M. Allenby (2003) "Bayesian statistics and marketing," *Marketing Science*, v22 p304-328.

[5.87] Randall, T., K. Ulrich, and D. Reibstein (1998) “Brand equity and vertical product line extent,” *Marketing Science*, v17 n4 p356-370.

[5.88] Rutenberg, D.P. (1971) “Design commonality to reduce multi-item inventory: optimal depth of a product line,” *Operations Research*, v19 p491-509.

[5.89] Shugan, S. M. and Desiraju, R. (2001), “Retail product-line pricing strategy when costs and products change,” *Journal of Retailing*, v77 n1.

[5.90] Simpson, T. (2004) “Product platform design and customization: status and promise,” *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, v18 n1.

[5.91] Spiegelhalter, D.J., Best N.G., and Carlin B.P. (1998) “Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models,” technical report, Division of Biostatistics, University of Minnesota.

[5.92] Spiegelhalter, D.J., Best N.G., Carlin B.P. and Linde van der A. (2002) “Bayesian measures of model complexity and fit,” *Journal of Royal Statistical Society*, Series B, v64 p583-640.

[5.93] Spiegelhalter, D.J., Richardson S., Gilks W.R. (1996) *Markov Chain Monte Carlo in Practice*, Chapman & Hall, New York.

[5.94] Steiner, W. and Hruschka, H. (2002) “A probabilistic one-step approach to the optimal product line design problem using conjoint and cost data,” *Review of Marketing Science*, Working Paper 441.

[5.95] Train, Kenneth E. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press.

[5.96] Villas-Boas (1998) “Product line design for a distribution channel,” *Marketing Science*, v17 n2 p156-169.

[5.97] Villas-Boas (2003) “Communication strategies and product line design,” *Marketing Science*, (forthcoming).

[5.98] Wassenaar and Chen (2003) “An approach to decision-based design,” *ASME Journal of Mechanical Design*, v125 p490-497.

[5.99] Wedel, M. and Kamakura, W. A. (2000) *Market Segmentation: Conceptual and Methodological Foundations*, 2nd Edition, Boston: Kluwer.

[5.100] Wedel, M., Kamakura, W., Arora, N., Bemmaor, A., Chiang, J., Elrod, T., Johnson, R., Lenk, P., Neslin, S., and Poulsen, C. S. (1999) “Discrete and continuous representations of unobserved heterogeneity in choice modeling,” *Marketing Letters*, v10 p219-232.

[5.101] Wind, Jerry and Vijay Mahajan (1997) "Issues and opportunities in new product development: an introduction to the special issue," *Journal of Marketing Research*, v34 p1-12.

[5.102] Wittink, Dick R. and Philippe Cattin (1989) "Commercial use of conjoint analysis: an update," *Journal of Marketing*, p91-96.

[5.103] Wittink, Dick. R., Marco Vriens, and Wim (1994) "Commercial use of conjoint analysis in Europe: results and critical reflections," *International Journal of Research in Marketing*, v11 p41-52.

[5.104] Zufryden, F. S. (1977) "A conjoint-measurement-based approach for optimal new product design and product positioning," A. D. Shocker, ed. *Analytical Approaches to Product and Market Planning*, Marketing Science Institute, Cambridge, MA, 100–114.

[5.105] Zufryden, F. S. (1982) "Product-line optimization by integer programming," Annual ORSA/TIMS Meeting, San Diego, CA.

CHAPTER 6

MANUFACTURING INVESTMENT AND ALLOCATION IN PRODUCT LINE DESIGN DECISION-MAKING

In Chapter 4 a method was introduced for coordinating marketing and engineering design decisions for product development to achieve joint solutions. In Chapter 5 this method was extended to design product lines using a heterogeneous model of demand. However, in both chapters the modeling of product cost was quite simple. In many cases, tradeoffs exist between improving the manufacturability of a product to reduce costs and improving the design of the product to achieve desirable product characteristics. Studies in Design for Manufacturability (DFM) encourage designers to consider manufacturing issues when making decisions, but generally these methods do not quantify the tradeoff between the expected additional revenue and cost generated by making a design change desirable in the market but costly to manufacturing. In this chapter, the product line model is further extended to include machine investment decisions along with allocation of production to the purchased machines. The material in this chapter is based on a working paper by Michalek, Ceryan, and Papalambros (2005).

6.1 METHODOLOGY

The methodology for including manufacturing decisions in the ATC hierarchy is built on top of the model presented in Chapter 5, referencing Sriraman, Imfeld and Swisher (2002) for manufacturing model development. In this extension, the marketing planning subproblem is augmented to include the setting of overall cost targets and production volume targets for each design to be achieved in a fixed time period.

Production volume targets are achieved by each engineering design subproblem by allocating design of the product's components to available machines, while ensuring that each component can only be made on machines capable of manufacturing the component, given the design. Life-cycle and dynamic manufacturing issues are not considered. Instead, it is assumed that a set of candidate machine types are available for purchase, and a single manufacturing investment subproblem is added to the ATC hierarchy to manage decisions of how many machines of each machine type will be purchased to match the cost targets set by marketing and simultaneously provide sufficient machine capacity for manufacturing the components designed in each engineering design subproblem. Because the production volume achievable for each product depends on the amount of machine time available for that product, linking variables are included to coordinate machine time requests and allocations between the engineering design subproblems and the manufacturing subproblem. A diagram of this process is provided in Figure 6.1, where linking variables are shown as connections between the two subproblems sharing those variables; while in actual implementation they are coordinated at the parent element, as described in Chapter 3.

6.1.1 Marketing Planning Subproblem

The marketing planning subproblem is formulated as in Chapter 5, using a heterogeneous mixture model of demand obtained through hierarchical Bayesian procedures using choice-based conjoint response data. However, in Chapter 5, decision variables of the marketing subproblem included only the price p_j of each product and the target product characteristic vectors $\mathbf{z}^{\mathrm{M}}{}_{j}$ passed down to each engineering design subproblem. It was assumed that production volume V_j for each product was equal to demand q, and the details of how this production volume would be achieved were outside the scope of the model. Similarly, cost was modeled as an exogenously determined investment cost c^{I} paid per design to set up infrastructure for production plus a variable cost c^{V} paid per product.

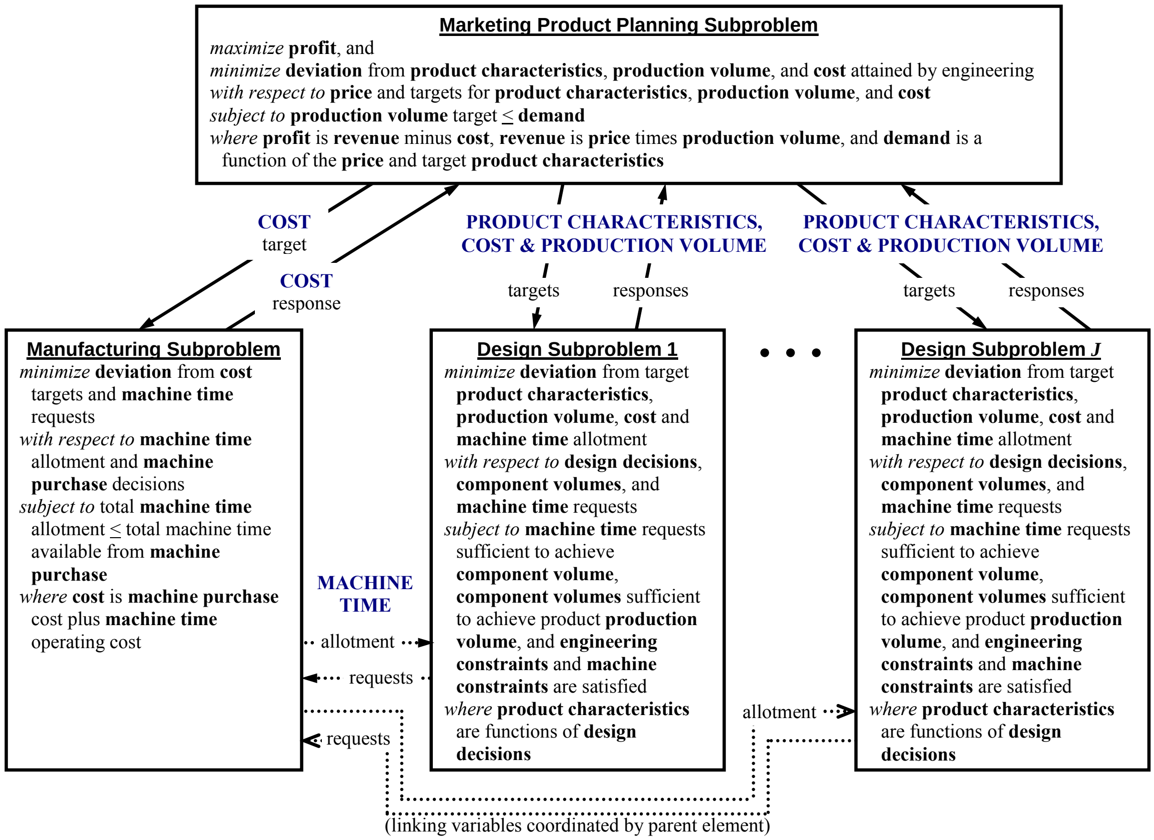


Figure 6.1: Description of ATC coordination of marketing, engineering design, and manufacturing decisions

In this chapter a more comprehensive model of manufacturing cost and capacity is required; therefore the marketing subproblem is altered so that the target production volume V^{M}_j, for a fixed time period T, is set in the marketing subproblem to be achieved by each engineering design subproblem V^{E}_j during that period. Likewise, target unit material cost c^{M}_j is set in the marketing subproblem to be achieved by each engineering design subproblem c^{E}_j, and a target for total investment and operating cost C^{M} is set in the marketing subproblem to be achieved by the manufacturing subproblem C^{P}. In this way the details of the cost and capacity calculations are handled outside of the marketing subproblem, but they are coordinated with marketing targets for these values.

Finally, if the production volume V^{M}_j of each product j is less than or equal to demand, the resulting profit is calculated as

$$\Pi = \left(\sum_j V_j^{\mathrm{M}} \left(p_j - c_j^{\mathrm{M}} \right) \right) - C^{\mathrm{M}}, \tag{6.1}$$

where p_j is the price of product j. If production volume were to be greater than demand, profit would be calculated differently, but here a constraint is added to ensure that $V^{\mathrm{M}}_j \leq q_j$. As before, it is still true that $V^{\mathrm{M}}_j = q_j$ at the solution, so it is not necessary to allow V^{M}_j to deviate from q_j. However, if $V^{\mathrm{M}}_j = q_j$ for intermediate iterations, this results in the marketing subproblem working to make product characteristics undesirable in order to match the lower production volumes achieved by engineering at intermediate iterations. This can increase computational time and also result in driving the marketing subproblem into an undesirable area of the design space where it may settle to a local minimum of lower global quality. Allowing $V^{\mathrm{M}}_j \leq q_j$ speeds up computation time by allowing each product to attract more than V^{E}_j individuals, reducing the need for marketing to make the product unattractive in order to match small production volume responses at intermediate iterations and mitigating opportunity for the algorithm to fall into a local minimum far from the starting point.

Finally, the marketing subproblem includes a coordinating liking variable T^{M}_{jm} to coordinate machine time allocation between the manufacturing subproblem T^{P}_{jm} and each engineering design subproblem T^{E}_{jm}. More detail about this variable will be provided in later sections as the marketing subproblem serves only as a parent coordination element with respect to T^{P}_{jm} and T^{E}_{jm}, as described in detail in Chapter 3.

The objective function of the marketing subproblem is then to maximize profit and minimize deviation between targets and responses of the unit material cost $c^{\mathrm{M}}{}_j$, investment and operating cost C^{M}, product characteristics $\mathbf{z}^{\mathrm{M}}{}_j$, production volume $V^{\mathrm{M}}{}_j$, and machine time $T^{\mathrm{M}}{}_{jm}$ variables for all products j and machines m. The full formulation of the marketing planning subproblem and its relationship with the other subproblems is shown in Figure 6.2, with vector norm weighting coefficients removed for clarity.

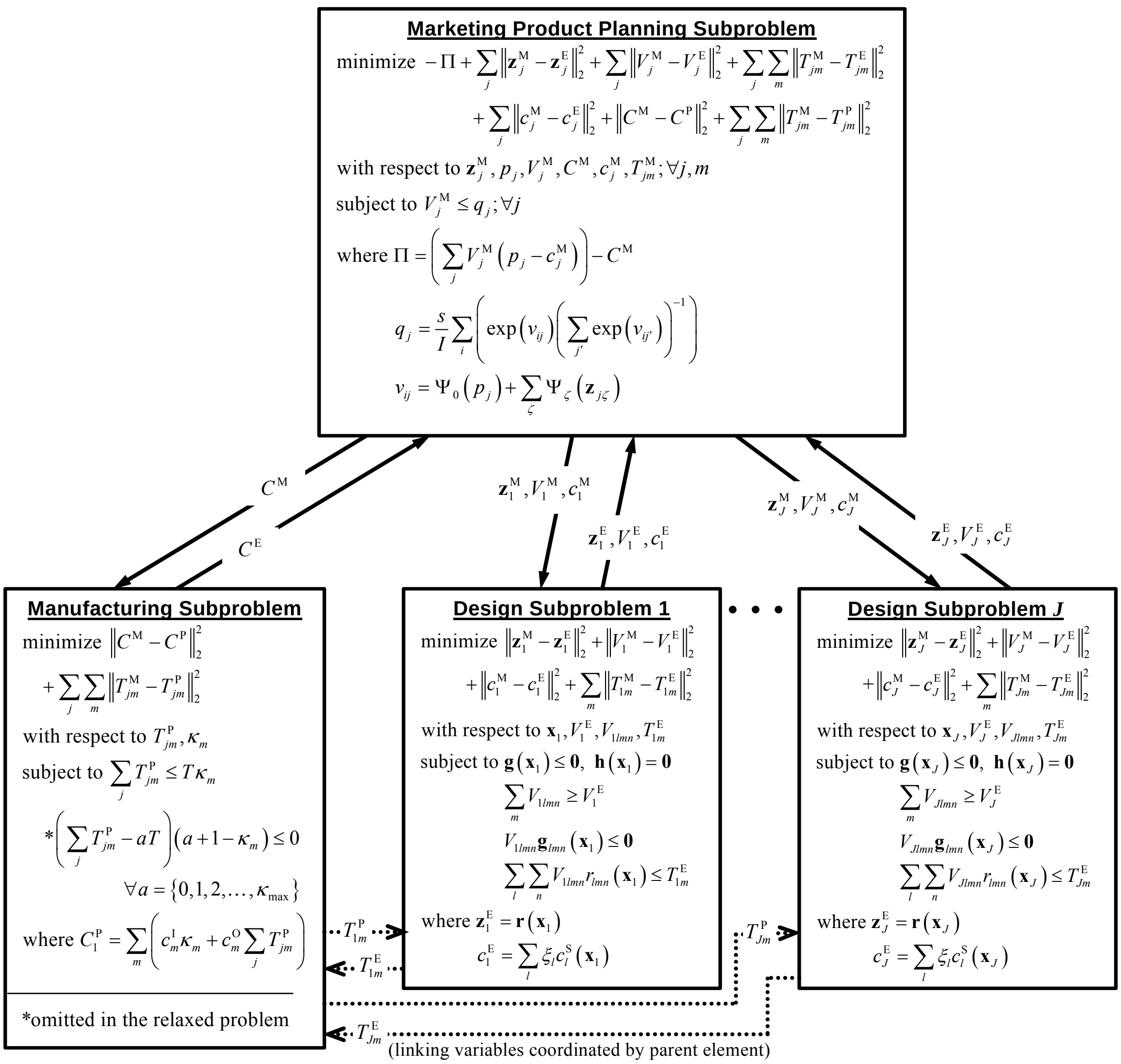


Figure 6.2: Mathematical description of ATC coordination of marketing, engineering design, and manufacturing decisions

6.1.2 Manufacturing Investment Subproblem

It is assumed that a fixed number of machine types $m = \{1, 2, \ldots, M\}$ is available from which to choose, and the firm must decide how many machines of each machine type to purchase κ_m. The possibility of leasing equipment is not considered. The manufacturing subproblem is tasked with dividing up the purchased machine time among products in the line by setting decision variables T^{P}_{jm}, indicating the amount of time on machine m that is allocated to product j. Only the allocation of machine time is considered here; Issues such as configuration (Spicer *et al.*, 2002), sequencing (Kurnaz *et al.*, 2005), line balancing (Son *et al.*, 2000), and multiple facilities (Benjaafar *et al.*, 2004) are not considered. If the parameter T represents the amount of machine time available per machine in a fixed period (i.e., the number of working hours over the period), then $\kappa_m T$ is the total time available from κ_m machines. Therefore, T^{P}_{jm} is constrained such that

$$\sum_j T^{\mathrm{P}}_{jm} \leq \kappa_m T \,. \tag{6.2}$$

In practice, each κ_m must be a nonnegative integer (0, 1, 2, …) because it is not possible to pay for a fraction of a machine at a fraction of the cost to receive a fraction of the capacity. However, the formulation is designed so that this requirement can be relaxed, permitting purchase of fractional numbers of machines. The solution to this relaxed problem will provide an upper bound on the amount of profit achievable by the more realistic situation where κ_m is restricted to integers. One way to restrict κ_m to integers is to do so explicitly, resulting in a mixed integer nonlinear programming problem (MINLP). However, the ATC formulation presented in Chapter 3 currently is not proven to converge for discrete formulations, and further research is necessary to extend the applicability of ATC to these problems. Alternatively, it is possible to solve such a problem under the current ATC formulation using branch-and-bound; however, the branching and bounding steps must be performed on *the entire ATC hierarchy*, not just the subproblem containing integer restrictions. To avoid the computational complexity of this strategy, it is possible to restrict the κ_m terms to integer values while working entirely in a continuous space. For a particular value of a, the following constraint

$$\left(\sum_j T_{jm}^{\mathrm{P}} - aT\right)\left(a+1-\kappa_m\right) \le 0 \quad \forall a \in \{0,1,2,\ldots\}, \forall m \tag{6.3}$$

ensures that when fewer than (a+1) machines are purchased (i.e., when $\kappa_m < a+1$), the total machine time allocated must not be greater than aT, the time provided by a machines. A set of these constraints for all a = {0, 1, 2, …} enforced together ensures that at least a machines must be purchased in order to use a machines worth of time for all values of a. In implementation, values of a need only be considered up to the maximum number of machines that may be purchased. While this set of constraints enables operation in a continuous domain and results in integer solutions for κ_m, it does not completely solve the problem. This set of constraints creates a "stair step" shaped feasible region, and because of the shape of the objective function, there are many cases where the shape of the feasible region creates several local minima: each at an integer value. Therefore, while the formulation allows operation in a continuous domain, solving for the optimum integer value of κ_m requires global search.

The strategy used here is to solve the relaxed problem (without the constraints in Eq.(6.3)) to obtain an upper bound on the profit achievable by the more restrictive problem. Next, starting from the optimum of the relaxed problem, penalty functions representing Eq.(6.3) are added to the objective function with a penalty coefficient parameter that increases over time until the solution is forced out of the infeasible region. This procedure results in a local minimum that is nearby the solution to the relaxed problem. The solution is not guaranteed to be the global solution; however, if it is within an acceptable deviation from the solution of the relaxed problem, it may be considered an acceptable and useful local solution.

Additionally, the cost of purchasing κ_m machines of type m is given by $\kappa_m c^{\mathrm{I}}_m$, where c^{I}_m is the investment cost per machine of type m. The cost to operate the machines of type m is equal to

$$\sum_j c^{\mathrm{O}}_m T^{\mathrm{P}}_{jm} , \tag{6.4}$$

where c^{O}_m is the cost per unit time to operate machine m (labor cost plus machine use cost). The total production cost C^{P} is composed of investment and operating cost, so that

$$C^{\mathrm{P}} = \sum_m \left(c^{\mathrm{I}}_m \kappa_m + \sum_j c^{\mathrm{O}}_m T^{\mathrm{P}}_{jm} \right). \tag{6.5}$$

Finally, the manufacturing subproblem objective is to minimize deviation from the cost targets C^{M} passed from the marketing subproblem and minimize deviation from the machine time allocation linking variables T^{M}_{jm} with values requested by each engineering design subproblem, and coordinated by the parent marketing subproblem. The full formulation of the manufacturing subproblem is provided in Figure 6.2.

6.1.3 Engineering Design Subproblems

Each engineering design subproblem is formulated similarly to those in Chapter 5, except that in addition to matching target product characteristics passed from marketing, each engineering design subproblem must attempt to match production volume targets, material cost targets, and machine allocation time linking variables. In addition, the individual components $l = \{1, 2, \ldots, L\}$ that make up each product j are considered, since each component is manufactured separately, where the parameter ξ_l defines the number of units of component l contained in each product. Each component l must undergo several manufacturing operations $n = \{1, 2, \ldots, N_l\}$; for example, a single component may require shearing, drawing, and bending operations. Production of the components $l = \{1, 2, \ldots, L\}$ that make up the product j must be allocated to machines $m = \{1, 2, \ldots, M\}$ in such a way that each component design meets the capability requirements of each machine on which it is made (Huang *et al*., 2003), and the total time requests made for each machine do not exceed the amount of time allocated. It is assumed that none of the designs in the product line share components. This is a limitation since it is common to design product families that share specific components among different product designs in a line to save costs (Kota *et al.,* 2000; Thonemann and Margaret, 2000; Lee, 2001; Qureshi, 2001; Fellini, 2003); however, questions of commonality add significant complexity, and it is a reasonable first step to rule out this possibility.

The component production allocation variable V_{jlmn}, represents the number of units of component l in design j on which operation n is performed by machine m. The production volume target V^{M}_j passed from marketing is achieved by producing enough of each component in the product to assemble V^{M}_j complete products. To do this, a decision variable V^{E}_j is added to represent the total number of products of type j produced, and this value is constrained, so that the manufacturing operations performed for each component V_{jlmn} are sufficient to generate the parts for V^{E}_j products.

$$\sum_m V_{jlmn} \geq \xi_l V_j^{\mathrm{E}}; \forall j, l, n. \tag{6.6}$$

Secondly, the total amount of time needed to execute manufacturing operations specified in V_{jlmn} must not exceed the amount of time T_{jm} allocated to product j on machine type m. If $r_{lmn}(\mathbf{x}_j)$ is a function specifying the time per component to execute operation n on component l with machine m for a design with variables $\mathbf{x}_j$, this constraint can be represented as

$$\sum_{l}\sum_{n} V_{jlmn} r_{lmn}\left(\mathbf{x}_j\right) \leq T_{jm}^{\mathrm{E}}; \forall j, m\,. \tag{6.7}$$

Finally, the volume V_{jlmn} of operation n performed on component l of product j by machine m may be greater than zero only if machine m has the capability to execute operation n on component l of product j. If $\mathbf{g}_{lmn}(\mathbf{x}_j)$ is a vector of constraint functions that define the feasibility of executing operation n on component l with machine m as a function of the design $\mathbf{x}_j$ of product j, then V_{jlmn} can be greater than zero only if $\mathbf{g}_{lmn}(\mathbf{x}_j) \leq \mathbf{0}$. If $\mathbf{g}_{lmn}(\mathbf{x}_j) > \mathbf{0}$, then the machine constraints are not satisfied by the product component, so operation n of component l cannot be performed on machine m, and V_{jlmn} must be exactly zero. This restriction can be represented by the following constraint

$$V_{jlmn}\mathbf{g}_{lmn}\left(\mathbf{x}_j\right) \leq \mathbf{0}; \forall j, l, m, n\,. \tag{6.8}$$

Taken in conjunction with the condition that $V_{jlmn} \geq 0$, this constraint ensures the specified relationship, allowing designs $\mathbf{x}_j$ to be altered to meet machine constraints and ensuring that components are not produced on machines if the design is not so altered. While this constraint can be implemented directly, it is advisable to implement it as a penalty function to avoid numerical problems with the near-colinearity of the gradients of Eq.(6.8) and the $V_{jlmn} \geq 0$ constraint for large values of $\mathbf{g}_{lmn}$.

The entire formulation for each engineering design subproblem, along with relationships to the other subproblems, are shown in Figure 6.2.

6.2 CASE STUDY

To demonstrate the methodology described in the previous section, the dial-readout case study is extended to develop manufacturing models and compare results from Chapter 5 with those achieved through coordination with manufacturing decisions. The software package DFMA: Design For Manufacture and Assembly, by Boothroyd Dewhurst (DFMA, 2004) was used to provide estimates of the manufacturing steps involved in producing the components of dial readout scales. For the case study, the scope was limited to the manufacture of five components: $l = 1$, the cover; $l = 2$, the base; $l = 3$, the (identical) long levers; $l = 4$, the (identical) short levers; and $l = 5$, the rack. There are two of each lever and one of each other component in each complete scale, so the number of components per product $\xi_l = \{1, 1, 2, 2, 1\}$ for $l = \{1, 2, 3, 4, 5\}$ respectively. Each of these components is produced with stamping machines. The cover and base require two operations ($N_1 = N_2 = 2$): a shearing operation ($n = 1$) followed by a bending operation ($n = 2$), each performed with a compound die. The levers and rack are each produced with a single shearing step in a progressive die ($N_3 = N_4 = N_5 = 1$). Material cost $c^S{}_l$ was also estimated per part. For simplicity, the unit material cost was treated here as constant, rather than as a function of the component dimensions; however, inclusion of unit material cost as a function of design dimensions is straightforward if data is available. Because the unit material cost is treated as constant in the case study, it need not be passed back and forth as a target, so the material cost calculation is included directly in the marketing subproblem here to reduce computational load. Finally, the force required to perform each operation was estimated based on the machine suggestions made by the software, and the time to load and unload each part was estimated by the software. These data are summarized in Table 6.1.

Table 6.1: Component and operation data

l	Part:	Parts per product	Material Cost ($/part)	n	Machine Operation	Force Required (tons)	Process	Strokes Per Part	Load / Unload Time Per Part (s)
1	Cover	1	$2.35	1	Shearing + Hole	100	Compound Die	3	8.35
				2	Bending	100	Compound Die	3	8.80
2	Base	1	$1.93	1	Shearing + Hole	100	Compound Die	3	8.32
				2	Bending	100	Compound Die	3	8.71
3	Long Lever	2	$0.28	1	Shearing	60	Progressive Die	1	NA
4	Short Lever	2	$0.16	1	Shearing	32	Progressive Die	1	NA
5	Rack	1	$0.07	1	Shearing	45	Progressive Die	1	NA

Secondly, a set of nine available machine alternatives ($M = 9$) was compiled using the software, which provided information on machine dimensions, force capacity, speed, and operating costs. Machine purchase cost estimates were obtained through informal discussions with Minster Machine Company. These machine data are summarized in Table 6.2.

Table 6.2: Machine characteristics

m	Machine	Bed Width (in)	Bed Length (in)	FORCE (tons)	Press Speed (strokes/ min)	Machine Rate ($/hr)	Operator Rate ($/hr)	Machine Cost
1	Minster P2H-160	33.5	63	180	40	$22.10	$25.00	$335,000
2	Minster P2H-100	26	48	112	60	$19.40	$25.00	$250,000
3	Minster OBI #4F	9	12	32	90	$16.30	$25.00	$75,000
4	Minster OBI #5F	12	16	45	85	$16.70	$25.00	$60,460
5	Minster OBI #6F	14	18	60	75	$17.40	$25.00	$90,000
6	Minster OBI #7F	14	19	75	70	$18.00	$25.00	$100,000
7	Minster E2-200	36	60	200	36	$22.80	$25.00	$200,000
8	Minster E2-300	42	96	300	36	$26.70	$25.00	$300,000
9	Minster E2-400	48	108	400	36	$30.60	$25.00	$400,000

Given these data, the rate function $\mathbf{r}_{lmn}$ can be calculated for each operation n on each machine m for each component l by dividing the number of strokes required per part by the machine press speed and adding the load / unload time. In general, $\mathbf{r}_{lmn}$ may be a function of the design variables $\mathbf{x}_j$; however, for simplicity in this case study it is taken to be constant with respect to $\mathbf{x}_j$. The time period of interest T is set to one year, encompassing 52 weeks, five days per week without holidays, and eight hours per day, for a total of 7,488,000 seconds of machine time per machine purchased. So, it is assumed that all machines are purchased in full at the beginning of the year for production during that year only. This is quite conservative, since most machines in industry are purchased with multiple years of production in mind; however, changing time periods or including machine leasing or resale options is straightforward. The machine constraints $\mathbf{g}_{lmn}$ are of two types: (1) ensure that the component is small enough to fit in the machine bed, and (2) ensure that the machine has sufficient force capacity to meet the component force requirements. Both of these conditions are enforced only for cases where $V_{jlmn} > 0$, as described previously. Specifically, the machine bed constraints applied to the cover, base, long lever, short lever, and rack respectively specify that

$$\begin{aligned} &l = 1: \quad x_{13}, x_{14} \leq \text{bed width} \\ &l = 2: \quad x_{13} - 2y_1, x_{14} - 2y_1 \leq \text{bed width} \\ &l = 3: \quad x_1 + x_2 \leq \text{bed width} \\ &l = 4: \quad x_3 + x_4 \leq \text{bed width} \\ &l = 5: \quad x_8 \leq \text{bed width} \end{aligned} \quad . \tag{6.9}$$

Secondly, the force capacity constraints specify that the machine force is greater than or equal to the component required force for each component, operation, and machine, using the relevant data from Table 6.1 and Table 6.2. Finally, the market potential s = 5 million, as in Chapter 5, and the value of the design parameters y are consistent with Chapter 5.

6.3 RESULTS

The ATC hierarchy was solved using as a starting point the solution from Chapter 5 with a value of zero for all machine purchases κ_m and time allocations (T^{M}_{jm}, T^{D}_{jm}, and

T^{P}_{jm}). The solution was obtained in three stages: First the relaxed problem (omitting Eq.(6.3) and Eq.(6.8)) was solved repeatedly while sequentially decreasing the weighting coefficient for the profit term of the objective function from 10^0 to 10^{-10} by powers of 10, solving the entire ATC hierarchy for each weighting value in turn to reach the *relaxed solution* with acceptable inconsistency between subproblems. Next, the penalty functions representing the machine feasibility constraints in Eq.(6.8) were added to the objective function with penalty terms increasing from 10^{-10} to 10^0 by powers of 10, gradually forcing the solution out of infeasible regions to achieve the *feasible solution*. Finally, the penalty function forcing κ to integer values (Eq.(6.3)) was added to the objective function with a scaling parameter which was similarly increased from 10^{-10} to 10^0 in powers of 10, gradually forcing κ to integer values and achieving the final *feasible integer solution*. The final resulting solution is not necessarily the global optimum, but it is a local optimum near the solution to the relaxed problem. Table 6.3 shows a comparison of the revenue, cost, profit, and machine purchase variables at each stage. The revenue, cost, and profit of the stage 2 feasible solution are very close to the stage 1 relaxed solution, suggesting that the relaxed solution had few bad assignments of components to machines incapable of making them, and was very near-feasible in this case. The profit of the final integer solution is lower than the stage 2 feasible solution, as expected since the stage 2 solution is an upper bound on the final integer solution. However, the resulting profit of the final solution is within 0.4% of the relaxed solution; therefore, the result is at least of high quality, and likely a global solution. In the first two stages the κ variables are real-valued, while in the final solution they are integers. Although the final solution in this case appears to be simply a rounding of the relaxed solution for each value of κ, this relationship does not hold in general, and alternative values of the parameters in the case study yield some solutions where the final integer solution is not equal to a rounding of the relaxed solution.

Table 6.3: Comparison of relaxed solution with final solution

	STAGE 1 Relaxed Soln	STAGE 2 Feasible Soln	STAGE 3 Integer Soln
revenue	$95,511,000	$95,511,000	$94,762,000
cost	$28,124,000	$28,125,000	$27,626,000
profit	$67,386,000	$67,387,000	$67,136,000
κ_1	0.00	0.00	0.00
κ_2	23.32	23.33	23.00
κ_3	0.18	0.15	0.00
κ_4	0.88	0.92	1.00
κ_5	0.78	0.78	1.00
κ_6	0.03	0.03	0.00
κ_7	0.05	0.04	0.00
κ_8	0.00	0.00	0.00
κ_9	0.00	0.00	0.00

The products that result from this optimization are shown in Table 6.4, along with their predicted market shares, production volumes, and selling prices. As before, the total market share for the line is less than 100% because of the existence of an outside good. Differences in design variables between the results in Chapter 5 and Chapter 6 occur because, as explained earlier, the design space in this problem does not map one to one with the product characteristics space, and multiple product designs exist that yield identical product characteristics: The specific design found by the algorithm on any given iteration is a matter of chance, but the coordination ensures that at least one feasible design exists that can attain the target product characteristics.

The variables associated with machine purchase and manufacturing allocation are provided in Table 6.5, where values near to zero are omitted for readability. In the table, the V_{jlmn} terms are shown in millions of units and the T_{jm} terms are shown in millions of seconds. In the final solution, the time available from purchased machines is allocated to the four products via the T_{jm} terms, and each product allocates its components to the most cost-effective available machines via the V_{jlmn} terms. To check this, observe that in the relaxed problem it so happens that the necessary total material, investment, and operating cost associated with executing V_{jlmn} repetitions of operation n on component l of product j with machine m is a linear function of V_{jlmn}

$$\left(c_j^{\mathrm{M}} + c_m^{\mathrm{I}} \frac{r_{lmn}}{T} + c_m^{\mathrm{O}} r_{lmn} \right) V_{jlmn} . \tag{6.10}$$

Table 6.6 shows the unit cost of each component and operation on each machine, where each shaded cell represents a machine unable to produce a particular component for any design variable values (insufficient force). In this linear case, it is easy to identify the most efficient machines for each component and operation, and these values are shown in bold. Table 6.5 shows that ATC was able to find these machines for efficient allocation.

Table 6.4: Product line design solution

	PRODUCT (j)			
	1	**2**	**3**	**4**
V^{M}_j (mil)	1.23	1.01	0.91	0.57
Share	25%	20%	18%	12%
z1	292	258	200	258
z2	0.980	1.155	0.924	0.975
z3	140	123	106	140
z4	0.103	0.119	0.121	0.115
z5	1.22	1.37	1.30	1.33
p	$24.13	$25.40	$24.57	$30.00
x1	11.78	0.125	9.351	9.846
x2	0.192	11.5	0.809	2.149
x3	3.364	5.981	5.745	3.778
x4	4.754	2.264	2.951	5.068
x5	0.125	0.186	0.125	0.135
x6	147.88	1.00	117.22	97.36
x7	0.50	0.50	0.50	0.50
x8	5.65	5.25	3.66	4.48
x9	0.353	0.766	0.478	0.345
x10	1.052	1.432	0.741	1.349
x11	1.696	1.776	1.696	1.878
x12	9.515	9.721	7.714	9.422
x13	11.71	11.92	9.914	11.68
x14	11.95	10.32	10.73	11.98

Table 6.5: Product line manufacturing investment and allocation solution

	j	l	n	MACHINE (m) 1	2	3	4	5	6	7	8	9
κ					23		1	1				
T_{jm} (million sec.)	**1**				5.68		0.25	0.20				
	2				4.66		0.20	0.16				
	3				4.22		0.18	0.15				
	4				2.67		0.11	0.09				
V_{jlmn} (million units)	**1**	**1**	**1**		1.23							
	1	**1**	**2**		1.23							
	1	**2**	**1**		1.23							
	1	**2**	**2**		1.23							
	1	**3**	**1**					2.41				
	1	**4**	**1**				2.32					
	1	**5**	**1**				1.21					
	2	**1**	**1**		1.01							
	2	**1**	**2**		1.01							
	2	**2**	**1**		1.01							
	2	**2**	**2**		1.01							
	2	**3**	**1**					1.97				
	2	**4**	**1**				1.92					
	2	**5**	**1**				0.98					
	3	**1**	**1**		0.91							
	3	**1**	**2**		0.91							
	3	**2**	**1**		0.91							
	3	**2**	**2**		0.91							
	3	**3**	**1**					1.78				
	3	**4**	**1**				1.69					
	3	**5**	**1**				0.90					
	4	**1**	**1**		0.57							
	4	**1**	**2**		0.57							
	4	**2**	**1**		0.57							
	4	**2**	**2**		0.57							
	4	**3**	**1**					1.11				
	4	**4**	**1**				1.03					
	4	**5**	**1**				0.56					

Table 6.6: Total cost per unit in the relaxed problem formulation

l	n	MACHINE (m) 1	2	3	4	5	6	7	8	9
1	**1**	$3.093	**$2.869**					$2.884	$3.077	$3.269
1	**2**	$3.119	**$2.890**					$2.902	$3.101	$3.300
2	**1**	$2.671	**$2.448**					$2.463	$2.655	$2.847
2	**2**	$2.694	**$2.465**					$2.478	$2.676	$2.874
3	**1**	$0.367	$0.326			**$0.299**	$0.302	$0.347	$0.371	$0.395
4	**1**	$0.247	$0.206	$0.174	**$0.174**	$0.179	$0.182	$0.227	$0.251	$0.275
5	**1**	$0.157	$0.116		**$0.084**	$0.089	$0.092	$0.137	$0.161	$0.185

As another check, the total V_{jlmn} for each component l in Table 6.5 should equal ξ_l times the production volume V^M_j listed in Table 6.4. For example, operation 1 of component 1 in product 1 has a production volume of 1.23 million on machine 2 in Table 6.5 to meet the demand for 1.23 million units of product 1 requested by marketing in Table 6.4. The values that nominally deviate are due to inconsistencies between subproblems, as discussed in Chapter 3. One limitation of ATC is that when more variables are shared, more places are created for the effects of inconsistencies between subproblems to build, and it becomes more computationally intensive to resolve inconsistencies to tight tolerances. Finally, multiplying r_{lmn} for each operation n and component l by the respective production volume V_{jlmn} values in Table 6.5 and summing for a particular product j and machine m yields the total machine time of machine m required by product j, which matches T_{jm} in Table 6.5.

6.4 CONCLUSIONS

The ATC methodology was applied to coordinate manufacturing investment decisions with marketing and product design decisions. Manufacturing decisions typically involve a number of inherently discrete decisions, such as how many machines of each type to purchase. Such discrete decisions introduce complexity. In this formulation, these discrete decisions were represented by relaxing the problem to a continuous space and imposing constraints to enforce solutions with discrete values; however, the formulation creates multiple local minima, and local search algorithms guarantee only local optimality of solutions. The strategy employed here is to solve the relaxed problem and then impose interior penalty functions to achieve a valid solution close to the relaxed solution. In the case study this strategy was successful, resulting in a final solution with a profitability within 0.4% of the relaxed solution; however, the application highlights the need for further research to extend the ATC methodology to problems with discrete variables so that mixed-integer programming can be utilized and more complex problems involving manufacturing can be solved.

The modularity of the ATC methodology allows additional considerations, such as the manufacturing subproblem introduced in this chapter, to be added to an existing hierarchy without starting from scratch. This modularity provides an opportunity for models in various disciplines to be built and used when available and when appropriate to the scope of questions of interest with minimal restructuring.

Finally, there exist alternative ways to decompose the marketing, engineering design, and manufacturing subproblems in this example. The formulation presented was designed to allocate as much complexity as possible to the engineering design subproblems in order to improve scalability to many products: With the inclusion of many products, the marketing and manufacturing subproblems grow in dimensionality; however, each engineering design subproblem remains constant in size. Additionally, in this formulation, as shown in Figure 6.2, the relaxed manufacturing subproblem is linear in constraints and quadratic in the objective function, so scalability is well achieved by this formulation. Additionally, choice of decomposition was made to minimize the number of variables shared among subproblems, improving computational properties. However, a more systematic methodology for determining how to decompose systems with multiple products for best scalability would be helpful.

This application of ATC has the potential to bridge gaps between design, manufacturing, and business perspectives of product development and production. The current model is static in the sense that market share is a deterministic function of the product characteristics and price, and demand does not vary over the time period in question. A number of potential extensions are possible such as modeling market dynamics by considering investment time (Georgiopoulos *et al.*, 2002) and demand fluctuation (Asl and Ulsoy, 2002a,b), or by including considerations of product life cycle economic modeling (Birge *et al.*, 1998), buyer-supplier relationships (Chick *et al.*, 1997), machine system configuration and adaptability (Maier-Speredelozzi and Hu, 2002; Zhong *et al.*, 2000) and machine reconfiguration (Koren *et al.,* 1999).

6.5 REFERENCES

[6.1] Asl, F.M., Ulsoy, A.G., (2002a) "Capacity management in reconfigurable manufacturing systems with non-stationary stochastic market demand," *ASME International Mechanical Engineering Congress and Exhibition*, New Orleans, LA.

[6.2] Asl, Farshid., Ulsoy, A.G., (2002b) "Capacity management via feedback control in reconfigurable manufacturing systems," *Proc. of the Japan-USA Symposium on Flexible Automation*, Hiroshima.

[6.3] Benjaafar, S., M. ElHafsi, and F. de Vericourt (2004) "Demand allocation in multiple-product, multiple-facility, make-to-stock systems," *Management Science* v50 n10 p1431-1448.

[6.4] Birge, J. R., Duenyas, I., Narongwanich, W. (1998) "When should one invest in reconfigurable capacity?" ERC/RMS Technical Report TR98-23, in review *Management Science*.

[6.5] Chick, S.E., Olsen, T., Sethuraman, K., Stecke, K., White, C.C. (1997) "A descriptive multi-attribute model for reconfigurable machining system selection that examines buyer-supplier relationships," Research Support, University of Michigan Business School, W.P. #9709-26, November. (in review *Journal of Operations Management*.

[6.6] DFMA: Design For Manufacture and Assembly, software by Boothroyd Dewhurst, accessed on 7-Nov-2004 *http://www.dfma.com*.

[6.7] Fellini, R. (2003) *A Model-Based Methodology for Product Family Design*, Ph.D. Dissertation, Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan, USA

[6.8] Georgiopoulos, P., Fellini, R., Sasena, M., Papalambros, P. (2002) "Optimal design decisions in product portfolio valuation," *Proceedings of the 2002 ASME Design Automation Conference*, DAC- 34097, Montreal, Quebec, Canada, September 29-October 2.

[6.9] Huang, Z., D. Yip-Hoi, J. Zhou (2003) "A graph-based approach for capturing the capability envelope of a machining process," *Journal of Manufacturing Science and Engineering*, v125 p272-288.

[6.10] Koren, Y., U. Heisel, F. Jovane, T. Moriwaki, G. Pritschow, G. Ulsoy and H. Van Brussel (1999) "Reconfigurable manufacturing systems," a keynote paper, *Annals of the CIRP*, v48 n2 p527-540.

[6.11] Kurnaz, S., A. Cohn, and Y. Koren (2005) "A framework for evaluating production policies to improve customer responsiveness," Working Paper, University of Michigan.

[6.12] Kota, S., K. Sethuraman and R. Miller (2000) "A metric for evaluating design commonality in product families" *Journal of Mechanical Design*, v122 p403-410

[6.13] Lee, B. (2001) *An Automatic Datum Definition Method for Product Families Based on Quantitative Manufacturing Feedback*, doctorial dissertation, University of Michigan.

[6.14] Maier-Speredelozzi, V., Hu, S.J. (2002) "Selecting manufacturing system configurations based on performance using AHP," *Transactions of NAMRI*, West Lafayette, Indiana.

[6.15] Michalek, J., O. Ceryan, and P. Papalambros (2005) "Manufacturing investment and allocation in product line design decision-making," working paper, University of Michigan.

[6.16] Qureshi, K. (2001) *Design for Facility (DFF) Approach for Commodity Parts*, doctorial dissertation, University of Michigan.

[6.17] Son, S-Y., Olsen, T.L., Yip-Hoi, D. (2001) "An approach to scalability and line balancing for reconfigurable manufacturing systems," *International Journal of Agile Management Systems*, v12 n7 p500-511.

[6.18] Spicer, P., Y. Koren, M. Shpitalni, D. Yip-Hoi (2002) "Design principles for machining system configurations," *Annals of the CIRP* v51.

[6.19] Sriraman, N., S. Imfeld and S. Swisher (2002) "Optimal part family and production module planning with scalable capacity" unpublished document, University of Michigan Department of Mechanical Engineering, Engineering Research Center for Reconfigurable Manufacturing Systems.

[6.20] Thonemann, Ulrich W. and Margaret L. Brandeau (2000) "Optimal commonality in component design," *Operations Research*, v48 n1 p1-19.

[6.21] Zhong, W., Maier-Speredelozzi, V., Bratzel, A., Yang, S., Chick, S., Hu, S.J. (2000) "Performance analysis of machining systems with different configurations," *Proc. Japan USA Symposium on Flexible Manufacturing*, Ann Arbor, Michigan.

CHAPTER 7

SOCIAL PREFERENCES: THE EFFECTS OF REGULATION POLICY IN A COMPETITIVE MARKET

The preceding chapters have considered modeling of preferences and design decisions from the perspective of a profit-seeking firm. Profit-seeking firms operate within a competitive, regulated marketplace, and the interactions of players in the market lead to specific decisions. In the literature, the ubiquitous objective of design is to "meet needs" of the user/consumer; however, profit-seeking enterprises, particularly corporations, are only driven to meet needs insofar as doing so is financially profitable. If, as designers, we are truly concerned about the degree to which our decisions meet needs and the way in which our decisions resolve conflicting stakeholder preferences, we should not take it for granted that the market system will resolve these tradeoffs perfectly and naturally.

In general, it is not simply profit for shareholders that drives designers: It is the desire to improve the world around us, and to the degree that markets accomplish this goal, we should use them as a tool. However, in theory and in practice, markets are not always aligned with public preferences (Chomsky, 1999). Regulation is used to constrain and direct markets to prevent great harm to society. From antitrust laws to environmental legislation to minimum wage, child labor, and safety laws, regulation is one corrective method used to ensure that a society organized around private property, private enterprise, and capital accumulation does not neglect social justice or self destruct. This is of utmost importance to design engineers interested in considering and modeling people's preferences because individuals as individuals acting in their own perceived short-term best interests have different preferences than those same individuals acting as part of a large, complex social system (Schumacher, 1973). The engineer's responsibility, as reflected in the codes of professional organizations such as the National Society of Professional Engineers (NSPE, 2004) the American Society of Mechanical Engineers (ASME, 2004), and the Institute of Electrical and Electronics Engineers (IEEE, 2004), is to both their professions and to the public, who put trust in engineers to make decisions that require technical training and analysis with the interests of the public in mind. Therefore, in pursuit of methods for modeling and coordination of stakeholder preferences with engineering design decisions, examination of the public interest is required.

One method for acting on public preferences is through legislation introduced and passed by elected representatives. While engineers traditionally have had relatively little involvement in the dialogue about such legislation, the resulting decisions affect engineering work directly. For example, recent environmental legislation, such as the European Union Directive on End-of-Life Vehicles and the Japanese Home Electric Appliances Recycling law, has had a major influence on product design from both an engineering and an economic perspective. By studying the interactions of markets and regulation with design decisions, engineers can improve their capability to predict consequences of their actions, and therefore can contribute meaningfully to the dialogue regarding legislation that directly affects the design process as well as the degree to which that process is aligned with the public interest.

This chapter presents a methodology for studying the effects of automobile fuel efficiency and emission policies on the long-term design decisions of profit-seeking automobile producers competing in an oligopoly market. Mathematical models of engineering performance, consumer demand, and manufacturing costs are developed for a specific market segment, and game theory is utilized to simulate competition among firms to predict design choices of producers at market equilibrium. Several policy scenarios are evaluated for the small car market, including corporate average fuel economy (CAFE) standards, carbon dioxide (CO_2) emissions taxes, and diesel technology quotas. The ability to compare regulations and achieve realistic trends suggests that including engineering design and performance considerations in policy analysis can yield useful predictive insight into the impact of government regulations on industry, consumers, and the environment. The material in this chapter is based on publications by Michalek, Papalambros and Skerlos (2005).

7.1 INTRODUCTION

In automotive manufacturing, profitability depends upon a vehicle's engineering performance and cost, as well as its appeal to consumers and the regulatory restrictions imposed by government. In this investigation, each of these points is considered and the impact of regulatory fuel-economy and emissions policies on the design decisions made by profit-seeking producers is evaluated.

Automobile producers provide private goods (vehicles) for private profit (investors), but externalities (emissions) are generated with costs that are publicly shared. For example, costs associated with driving high-emission vehicles in the southern coast of California can generate pollution costs estimated at $10,000 or more per year (Dixon, Garber, and Vaina, 1996). Despite regulatory enforcement over the past three decades, vehicle emissions still significantly impact U.S. air quality, accounting for up to 95% of city CO emissions, 32% of NO_x emissions, and 25% of volatile organic compound emissions (US EPA, 2001). These emissions create smog, increase atmospheric greenhouse gas concentrations, create human health risks, and damage agricultural, ecological, and urban infrastructure systems. Since the market in which goods are traded does not automatically provide individual incentives to reduce publicly shared environmental damage (the "tragedy of the commons"; Hardin, 1968), government regulatory policies have been imposed on vehicles at both national and state levels to provide emission reduction incentives. Examples include the Clean Air Act (US EPA, 1990), which regulates tailpipe emissions, corporate average fuel economy (CAFE) standards (US Congress, 1975), which require vehicle fleets to meet target average fuel efficiencies, and quotas for “cleaner” vehicles, such as California’s “zero emissions vehicle” (ZEV) regulation. While National Ambient Air Quality Standards established by the Clean Air Act still have not been achieved in many major U.S. cities, recent attempts to regulate further the vehicle design process toward producing “cleaner” vehicles have had only limited success. One example is California’s attempt to achieve 10% sales in ZEVs from its top seven automotive manufacturers by 2003 (Dixon, Garber and Vaina, 1996). The ZEV technology quota policy has suffered from the high cost (average purchase cost of $35,000) and poor range (approximately 90 miles) of electric vehicles (Gardner, 1996), resulting in limited consumer appeal. The policy is now under review, with low polluting gasoline and highly fuel efficient gasoline-electric hybrids likely to comprise the bulk of the 10% quota (Associated Press, 2003). The example demonstrates the importance of simultaneously considering technology capabilities, costs, and consumer preferences when developing environmental policies.

In this chapter, a quantitative methodology is developed for considering engineering design performance and constraints, producer objectives, consumer choice, and competition among producers in the analysis of environmental policy. This methodology permits specific policies to be analyzed in the context of their impacts on consumers, producers, and total air quality, leading to estimates of cost and effectiveness for different environmental policies under consideration.

7.2 BACKGROUND

Policy research related to the automotive industry has focused primarily on the effects of changing CAFE standards. One such study by the National Academy of Sciences (2002) identified technologies that could be implemented in all vehicles today, including estimated cost and fuel savings associated with each technology. Specifically, the effects of incremental changes in CAFE standards on vehicle price, performance, demand, and product mix were evaluated. While external factors such as gasoline price were also included in the assessment, the report considered only the inclusion of new technologies in existing engines. Longer term options to change new vehicle design decisions were not considered. The same is true for a recent European Union report on the Auto-Oil II Programme, which targets reductions in automobile emissions (2000). In the report, future vehicle emissions levels were forecast as functions of fuel quality using atmospheric emissions and impact models. Although alternative emissions policies were evaluated for their economic efficiency in reducing emissions, the option for producers to change design decisions in response to policy was not considered.

A different study by Greene and Hopson (2003) examined the impact of various regulatory strategies on average fuel economy using a mathematical programming model. Regulatory options included raising the CAFE standard, making a fuel-economy standard voluntary, and creating a weight-based metric. Although regulatory options were evaluated in the context of their impact on producers and consumers, the market positions of manufacturers were taken as constant, and few longer-term design changes were considered.

While these previous models analyze important aspects of emissions policies, there are opportunities to extend their scope of consideration. Previous investigations assume each manufacturer will maintain its current product mix, making only incremental technology improvements to existing products (e.g., direct injection, variable valve timing, etc). In contrast, this chapter provides an economic oligopoly analysis where each firm designs its product mix, changing design variables in response to regulations and competition. Previous studies also rely on assumptions about consumer willingness to pay for increased fuel economy rather than using attribute-based consumer choice models derived from past purchase data. This chapter uses an optimization framework to integrate quantitative models for each component, including emissions, engineering design, cost, consumer demand, and producer profit. The framework is modular and hence allows for the substitution of alternative models for any of the various models employed in this study. Moreover, the producers in this investigation are abstract; that is, the results obtained do not apply to a specific producer's actions, but rather represent the general market trend created by government incentives. Therefore the model created here is able to evaluate trends of cost and effectiveness created by alternative policies that aim to reduce automobile emissions through improved fuel economy.

The remainder of this chapter proceeds as follows. Section 7.3 describes the proposed policy analysis methodology, including the development of individual models for engineering performance, consumer demand, cost, producer profit, and regulation. The models are utilized to establish oligopoly market competition between firms, where policy impacts are analyzed at Nash equilibrium. The results of the investigation are summarized in Section 7.4.

7.3 METHODOLOGY

While it is possible to use ATC as before to coordinate this problem, ATC is abandoned here because of the relatively simple design variable and product characteristic space. For simplicity, a single optimization loop is used instead to optimize the decisions of a single profit-seeking firm, and the scope of focus is extended to include competition and regulation effects. Similarly, while it is possible to use a heterogeneous Bayesian model of demand, as in Chapter 5, a simple aggregate logit model from the literature is used here for simplicity.

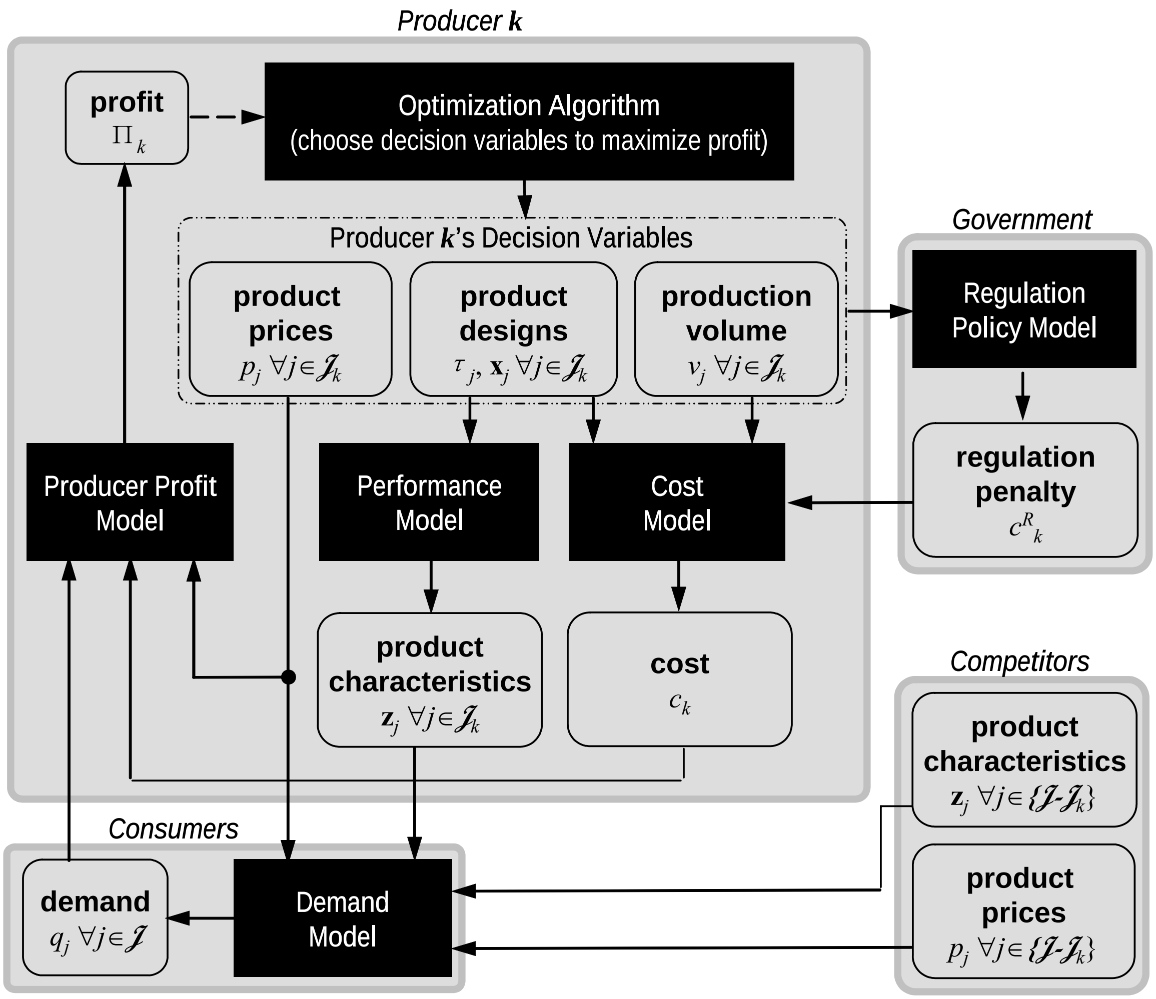


Figure 7.1: Overview of the modeling framework

The general modeling framework used to capture producer and consumer behavior in this study is shown in Figure 7.1, where individual analysis models are shown as black boxes. As before, producers are assumed to make product design and production decisions that maximize profit, and consumers are assumed to choose from the available alternatives those products that have maximum utility based on a model of their preferences. New components include modeling the effects of competitor products, which compete for market share, and regulation policy, which can influence these decisions by imposing penalties and incentives toward the modification of producer and consumer behavior. This investigation considers several policy scenarios that have direct impact on producer behavior such as CAFE standards, carbon dioxide (CO_2) emissions taxes, and diesel technology quotas.

In this framework, each producer k decides on a set of designs to produce $\mathcal{J}_k$ including design decisions, prices, and production volumes for each design. Design topology τ and design variables $\mathbf{x}$ (such as engine size) determine product characteristics $\mathbf{z}$ (such as fuel economy), calculated using an engineering performance analysis model. Design variables, production volume V, and regulation penalties c^R also determine producer cost c, calculated by the cost analysis model. The set of competitors' designs $\{\mathcal{J} - \mathcal{J}_k\}$ are viewed by producer k as static parameters, and consumers make purchasing choices among the set of producer and competitor products $\mathcal{J}$ based on product characteristics and prices p. Purchasing choices determine demand for each design q calculated by the demand model, and resulting profits Π are calculated in terms of p, q, and c. Resulting profit is used as the objective function for producer k's optimization model, and the dotted line in Figure 7.1 represents the feedback loop for iterations of the optimization algorithm. The optimization model represents each producer's attempt to maximize profit by making the best design, pricing, and production decisions. Government regulation can influence this process by imposing penalties on producers, thereby affecting production costs and design decisions. Note that this study is limited to government regulation directly affecting producers without impacting consumer behavior such as driving habits or preference structures.

In the present model, all producers are profit driven, so production volume will equal product demand at an optimum. This assertion is valid for continuous demand functions with negative price elasticities since any producer who wishes to produce a lower volume of a product (for example, because of capacity constraints or marginal cost curves) has no incentive to produce less volume than that for which there is demand. Instead, the producer can simply raise the price until demand is lowered to the desired production volume, so it is assumed that $V_j = q_j$ from this point forward.

The objective of each producer is modeled as profit maximization (Π, revenue minus cost) subject to engineering constraints as follows,

$$\begin{aligned}&\text{maximize } \Pi_k = \left(\sum_{j \in \mathcal{J}_k} q_j p_j \right) - c_k \\ &\text{with respect to } \{\tau_j, \mathbf{x}_j, p_j\} \ \forall j \in \mathcal{J}_k \\ &\text{subject to engineering constraints}\end{aligned} \tag{7.1}$$

Profit for each producer is calculated as a function of the producer decision variables by combining the engineering performance, consumer demand, cost, profit, and regulation models described in Sections 3.1–3.5. The size n_k of the set $\mathcal{J}_k$, is a variable in this formulation. For a fixed n_k and fixed engine types τ_j for each vehicle, the model (Eq.(7.1)) is a smooth, continuous optimization formulation that can be solved with gradient-based methods. To take advantage of this property, separate optimization runs are formulated for each combinatorial set of $n_k \in \{1, 2, \ldots, n_{\max}\}$ and τ_j $\forall j \in \mathcal{J}_k$, and gradient-based methods are used to determine the optimal solution for each value of n_k. The most profitable solution among these cases is then taken as the optimum solution.

While this modeling framework is presented as a single loop of sequential computation solved all-at-once, it is possible to break the problem into smaller pieces using multistage approaches (Li and Azarm, 2000, 2002) or decomposition and coordination optimization methods such as collaborative optimization (CO) (Balling and Sobieszczanski-Sobieski, 1994) and analytical target cascading (ATC) (Kim, 2003), as in previous chapters.

7.3.1 Engineering Performance Model

The engineering performance model takes design decisions $\mathbf{x}_j$ as input and predicts performance characteristics $\mathbf{z}_j$ that can be calculated for each design j. Several analysis models were explored for vehicle modeling, and ADVISOR (Markel *et al.,* 2002; Whitehead, 2001) was chosen because of its availability and appropriate level of detail for this study. ADVISOR contains models for conventional, electric, hybrid electric, compressed natural gas, and fuel cell vehicles. Experimentally-derived engine maps are used to estimate fuel economy and emissions characteristics across engine operating conditions. The vehicle is simulated through a driving cycle, and fuel economy, performance characteristics, and vehicle emissions are calculated for the cycle.

In this study, vehicles are assumed to differ only by engine design, so the default small car vehicle parameters were used in all simulations (based on the 1994 Saturn SL1), and only engine variables were changed. ADVISOR offers a set of nine gasoline and eleven diesel engine types. Each engine type has a base size b_τ, corresponding to the power output of a tested engine, which can be scaled to predict performance of larger or smaller engines. (ADVISOR allows scaling parameters between 0.75 and 1.50). The EPA Federal Test Procedure (FTP-75) driving cycle was used for all simulations. Two engine types, $\tau = \{\text{SI102, CI88}\}$, were utilized in this study with two design variables: the engine scaling parameter x_1 in the range [0.75, 1.50], and the final drive ratio x_2 in the range [0.2, 1.3]. The computed outputs (performance criteria) include the gas mileage (gasoline equivalent) z_1 in miles per gallon (mpg) and the time to accelerate from 0 to 60 mph, z_2, in seconds. The engine type τ = SI102 refers to a spark ignition (gasoline) engine with b_{SI102} = 102kW based on the 1991 Dodge Caravan 3.0L engine, while τ = CI88 refers to a compression ignition (diesel) engine with b_{CI88} = 90.5kW based on an Audi 2.5L engine. Other engine types were explored but turned out to be oversized or undersized for this study. For a particular choice of engine type τ, ADVISOR acts as a function f_τ mapping $\mathbf{x}$ to $\mathbf{z}$:

$$\mathbf{z} = f_\tau(\mathbf{x}) \tag{7.2}$$

where $\mathbf{z} = [z_1, z_2]^T$, and $\mathbf{x} = [x_1, x_2]^T$. ADVISOR simulations were computed for evenly spaced points in a 13 by 19 point grid covering the ranges of x_1 and x_2 respectively for each engine type, and the responses were used to create a set of surface splines as surrogate models for ease of computation during optimization. Sample contour plots of the simulation results are shown in Figure 7.2.

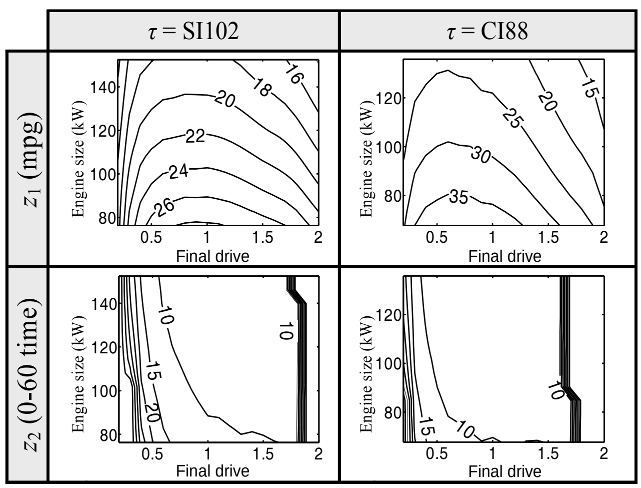


Figure 7.2: ADVISOR simulation result contour plots

7.3.2 Consumer Demand Model

The consumer demand model is based on discrete choice analysis (DCA), which presumes users make purchasing decisions based on the *utility* value of each product option. Utility u is measured in terms of an observable deterministic component v, which is taken to be a function of product characteristics, and a stochastic error component ε.

The probability P_j of choosing a particular product j from the set ϑ is calculated as the probability that product j has a higher utility value than all alternatives.

$$P_j = \Pr\left(v_j + \varepsilon_j \geq v_{j'} + \varepsilon_{j'}; \forall j' \in \mathcal{J}\right) \tag{7.3}$$

Various probabilistic choice models follow the DCA approach, including the logit model (McFadden, 1976) and the probit model (Currim, 1982). As before, assuming the double exponential distribution for the ε terms in Eq.(7.3), the probability P_j of choosing alternative j from set $\mathcal{J}$ is computed (Train, 2003) as

$$P_j = \frac{e^{v_j}}{\sum_{j' \in \mathcal{J}} e^{v_{j'}}}. \tag{7.4}$$

Each utility function v_j depends on the characteristics $\mathbf{z}_j$ and the price p_j of design j. Given a functional form for $v_j(\mathbf{z}_j)$ based on observed data, regression coefficients are found such that the likelihood of generating the sample data with the model is maximized. For example, Boyd and Mellman (1980) fit a simple logit model to automotive sales data based on price, fuel economy, and acceleration (among other vehicle factors). After an analysis of several other vehicle choice models (Berkovec, 1985; Brownstone *et al.,* 2000; Goldberg, 1996; Manrai, 1995; Wojcik, 2000; Yee, 1991), the Boyd and Mellman model was chosen for this study for the following reasons:

- The model is based on product characteristics that can be related to engineering design, as opposed to consumer demographics.
- The independent variables include the vehicle's price, fuel economy, and acceleration, which match the characteristics predicted by the engineering performance model under consideration in this study.

- The model was fit to a large volume of annual market data and validated using data from a subsequent year.

The utility equation developed by Boyd and Mellman[10] is:

$$v_j = \beta_1 p_j + \beta_2 \left(\frac{100}{z_{1j}} \right) + \beta_3 \left(\frac{60}{z_{2j}} \right) \tag{7.5}$$

where β_1 = -2.86·10^{-4}, β_2 = -0.339, β_3 = 0.375, p_j is the price of vehicle j, z_{1j} is the gas mileage of vehicle j, and z_{2j} is the 0-60 mph acceleration time of vehicle j. Although several other variables were included (e.g., vehicle style, noise, and reliability), these variables were assumed constant across all vehicles for this study. Since logit choice predictions depend on the differences between utility values, factors that are constant across alternatives do not affect predictions of choice, and they can be ignored. Other factors, such as advertising, promotions, aesthetics, and brand image were also assumed equal across alternatives. While the Boyd and Mellman demand model is adequate for a preliminary analysis, it does introduce several sources of error:

- The model was fit to purchase data from 1977-1978.
- The model utilizes purchase data only; consumers who chose not to purchase vehicles were not studied. Thus, we can predict only *which* vehicles consumers will purchase, not *whether* they will purchase, and the size of the purchasing population is treated as fixed, independent of vehicle prices (i.e., there is no outside good).

[10] The coefficients β_1 and β_2 were assumed here to be negative, even though they are listed as positive in the Boyd and Mellman article. In the text the authors describe the variables as having a negative relationship even though all coefficients are listed as positive in the regression summary.

- The model is an aggregate model, and therefore it does not account for different segments or consumer groups.
- The use of the logit model carries with it a property called *independence from irrelevant alternatives* (IIA), which implies that as one product's market share increases, the shares of all competitors are reduced in equal proportion (Train, 2003). For example, a model with the IIA property might predict that BMW competes as equally with Mercedes as with Chevrolet. In reality, different vehicles attract different kinds of consumers, and competition is not equal. In this investigation, predictive limitations of the IIA property are mitigated since the model is applied only to the small car market (a relatively homogeneous market) rather than to the entire spectrum of vehicles.

The demand model above was developed by Boyd and Mellman to study the effects of fuel economy standards on the market, and it should be sufficient to capture the trends important in a general analysis, even if the numbers vary for today's consumers. For the purposes of this study, the assumption was made that the size of the car-buying population *S* is 1.57 million people. This figure is based on 11 million people that bought cars in 1977 (US Dept. of Transportation, 2001) and an assumption that the size of the small car market was about 1/7 of the total market[11]. The Boyd and Mellman model was then applied to the small car sub-market, with recognition that this could introduce additional error since the model was developed based on the entire car market. Using the logit model with a fixed market size *S*, the demand q_j for product *j* is

$$q_j = SP_j = S \frac{e^{v_j}}{\sum_{j' \in \mathcal{J}} e^{v_{j'}}}, \qquad (7.6)$$

[11] Further research indicated that a better estimate of the size of the small car market may be 2/7 of the total market (US Dept. of Energy, 2001)

where v_j is defined by Eq.(7.5).

7.3.3 Cost Model

Production cost is modeled as a function of the vehicle design, and all producers are assumed to have the same manufacturing cost structure. In practice, differences in equipment, assets, suppliers, and expertise exist between manufacturers. However, assuming consistent production cost structures across manufacturers is appropriate for oligopoly analysis, and it is useful to analyze trends even if individual numbers differ between firms. In this analysis, the total cost to manufacture a vehicle c^P is decomposed into two components: the investment cost to set up the production line c^I and variable cost per vehicle c^V. The variable cost is comprised of the cost to manufacture the engine c^E and the cost to manufacture the rest of the vehicle c^B, so that $c^V = c^B + c^E$. The cost to manufacture q units of a vehicle with topology τ and design variables $\mathbf{x}$ is then

$$c^{\mathrm{P}}\left(\tau,\mathbf{x}\right)=c^{\mathrm{I}}+qc^{\mathrm{V}}\left(\tau,\mathbf{x}\right)=c^{\mathrm{I}}+q\left(c^{\mathrm{B}}+c^{\mathrm{E}}\left(\tau,\mathbf{x}\right)\right) \tag{7.7}$$

where it is assumed that c^{B} = \$7500 for all vehicles based on data for the Ford Taurus (Delucchi and Lipman, 2001), and c^{I} = \$550 million per vehicle design for all manufacturers based on an average of two figures for new production lines (Whitney, 2001). The cost to manufacture an engine is modeled as a function of engine power, as determined by a regression analysis of data obtained from manufacturing, wholesale, and rebuilt engine costs (American Speed Enterprises, 2003; E Diesels, 2003; Ford Motors, 2003; Hardy Diesels, 2003; National Engine, 2003; Rebuilt Auto Car Engines, 2003). Wholesale and rebuilt engine prices were assumed to be close to manufacturing prices, and these data fit the curve well. The resulting functions are,

$$c^{\mathrm{E}}(\tau,\mathbf{x})=\begin{cases}\beta_4\exp\left(\beta_5 b_M x_1\right) & \text{if } \tau\in\mathrm{SI}\\ \beta_6\left(b_M x_1\right)+\beta_7 & \text{if } \tau\in\mathrm{CI}\end{cases} \tag{7.8}$$

where β_4 = 670.51, β_5 = 0.0063, β_6 = 26.23 and β_7 = 1642.8. These functions are plotted in Figure 7.3, and all designs considered in this study fall within the range of the data. As expected, the cost associated with manufacturing diesel engines is higher than for gasoline engines. It is possible that increased diesel production volumes would change this cost structure, but this possibility was not explored in this study. Although both cost regression models rely on maximum engine power as the only dependant variable, Figure 7.3 demonstrates that the regressions fit the data well and predict realistic cost trends.

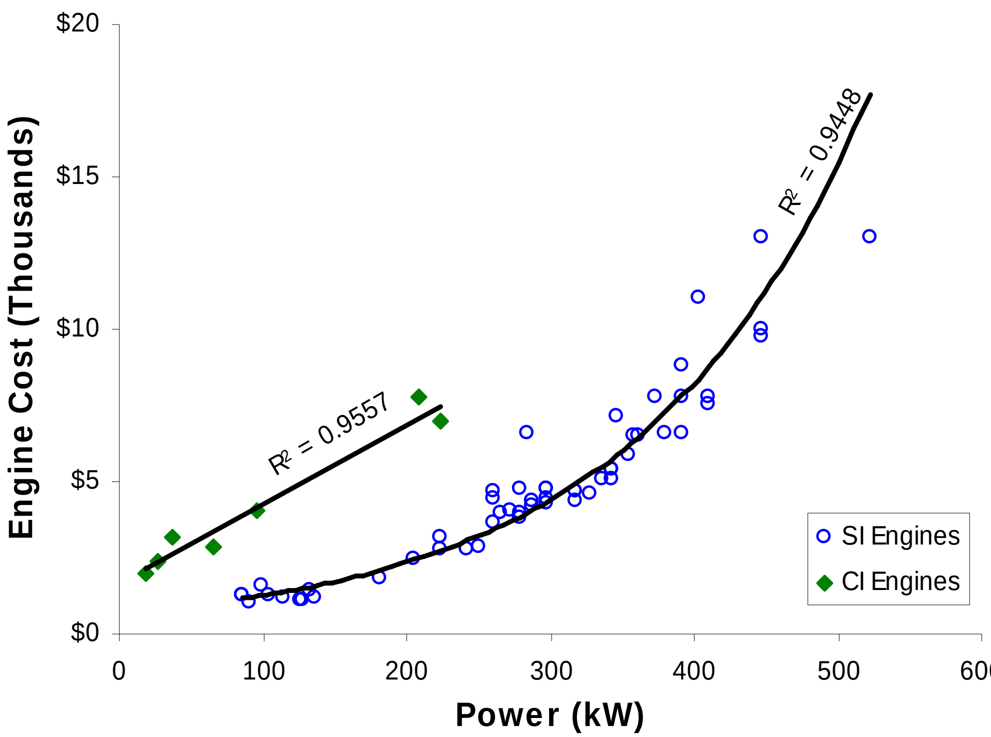


Figure 7.3: Manufacturing cost for SI and CI engines

The total cost to producer k is the sum of the production costs for each vehicle in k's product line and the regulation cost c^{R}, as described in Section 3.5.

$$c_k = \left(\sum_{j \in \mathcal{J}_k} c_j^{\mathrm{P}} \right) + c_k^{\mathrm{R}} \tag{7.9}$$

7.3.4 Profit Model

The profit model for each producer k is calculated simply as revenue minus cost:

$$\Pi_k = \left(\sum_{j \in \mathcal{J}_k} q_j p_j \right) - c_k = \left(\sum_{j \in \mathcal{J}_k} q_j \left(p_j - c_j^{\mathrm{V}} \right) - c^{\mathrm{I}} \right) - c_k^{\mathrm{R}} \tag{7.10}$$

where $c^{\mathrm{R}}{}_k$ is the regulation cost for producer k (defined in Section 3.5). The model assumes that all transactions happen instantaneously without consideration of the time value of money, opportunity costs, or changes in production loads over time. Demand is predicted over the course of one year, with all costs and revenue occurring during that year. The inclusion of dynamic time considerations brings with it a plethora of uncertainties and issues that are difficult to model, and is therefore left for future consideration. Note that it is assumed that the investment cost c_I is completely paid during this year. In practice, the investment cost associated with designing and building production lines and planning supply chains is spread over several years with only minor changes to the vehicles during those few years, implying that this model will tend to over-predict investment cost.

7.3.5 Regulation Policy

Four producer penalty policies were used to define c^{R}: the no-regulation base case ($c^{\mathrm{R}} = 0$), CAFE standards, CO_2 emission taxes, and diesel vehicle sales quotas. Each of these policies applies a penalty cost to the producer as a function of the fuel economy, emission properties, or fuel type of the producer's vehicles. The specific applications of the penalty policies are described below.

Corporate Average Fuel Economy (CAFE)

CAFE regulations establish minimum average fuel economy standards that each producer's vehicle fleet must meet to avoid penalties. To define a CAFE policy, both the fuel economy standard and the penalty must be specified. In this study, only a single market segment is utilized, although CAFE regulations in the United States apply to all passenger vehicle markets in which the producer operates. (Multiple market segments are left for future consideration.) The current CAFE fuel economy standard for cars, z_{CAFE} = 27.5 mpg, was used here, and two different penalty charges were explored: the current standard, ρ = \$55 per vehicle per mpg under the limit, and a hypothetical double-penalty scenario. Additional future credit for vehicle fleets with average fuel economies greater than the standard was not modeled. The total cost incurred by design j is therefore $\rho q_j(z_{\mathrm{CAFE}} - z_{1j})$, where ρ is the penalty, q_j is the number of vehicles of type j that are sold, z_{CAFE} is the CAFE limit, and z_{1j} is the fuel economy of vehicle j. The total regulation cost to producer k is then

$$c_k^{\mathrm{R}} = \max\left(0,\ \sum_{j \in \mathcal{J}_k} \rho q_j \left(z_{\mathrm{CAFE}} - z_{1j}\right)\right) \tag{7.11}$$

CO_2 Emission Tax

A vehicle emission valuation study (Matthews and Lave, 2000) was used to estimate the economic cost to society associated with environmental damage due to the release of each ton of CO_2. Using this valuation, a tax can be imposed on the manufacturer based on the estimated lifetime CO_2 emissions of each vehicle sold due to the burning of hydrocarbon fuel. Tax per vehicle sold can be calculated as $\upsilon d\alpha_\tau/z_1$, where υ is the dollar valuation of a ton of CO_2, d is the number of miles traveled in the vehicle's lifetime, α_τ is the number of tons of CO_2 produced by combusting a gallon of fuel for engine type τ, and z_1 is the fuel economy of the vehicle. The total regulation cost to the producer in this study is

$$c_k^{\mathrm{R}} = \sum_{j \in \mathcal{J}_k} q_j \frac{\upsilon d \alpha_\tau}{z_{1j}} \tag{7.12}$$

where d = 150,000 miles, α_τ is $9.94 \cdot 10^{-3}$ tons CO_2 per gallon for gasoline or $9.21 \cdot 10^{-3}$ tons CO_2 per gallon for diesel fuel (Bohac and Jacobs, 2002), and the value of υ was varied from \$2/ton to \$23/ton with a median estimation of \$14/ton.

Diesel Fuel Vehicle Sales Quotas

As a regulation method, quotas can be used to force more costly alternative fuel vehicles into the market (Associated Press, 2003). In this case, a hypothetical policy is considered that introduces a large penalty cost for violation of a quota on percent diesel sales as a way to enforce adoption of a higher fuel efficiency vehicle alternative. Diesels were selected due to data availability, their competitive fuel efficiency and acceleration characteristics, and their similarity to gasoline engines in unobserved characteristics such as range and existence of supporting infrastructure, which allows application of the demand model without introducing large errors. It is left for future work to consider regulation of emissions such as NO_x and particulate matter, which tend to be larger in diesel engines and which play a significant role in determining environmental tradeoffs between diesel and gasoline engines in practice. The regulation cost is modeled as

$$c_k^{\mathrm{R}} = \max\left(0,\ \rho\left(q_k^{\mathrm{SI}} - \left(1-\phi\right)\left(q_k^{\mathrm{SI}} + q_k^{\mathrm{CI}}\right)\right)\right) \tag{7.13}$$

where ρ is the penalty per gasoline vehicle over quota (\$1000), ϕ is the minimum diesel percentage required by the quota (40%), $q^{SI}{}_k$ is the total number of spark ignition (gasoline) engines sold by producer k, and $q^{CI}{}_k$ is the total number of compression ignition (diesel) engines sold by producer k.

7.3.6 Nash Equilibrium Solution Strategy

In a free market, manufacturers have economic incentives to produce and sell products only if there is an opportunity to make profit within the competitive market. To account for competition in the design of vehicles subject to government regulations, game theory was used to find the market (Nash) equilibrium among competing producers. In game theory, a set of actions is in Nash equilibrium if for each producer k = 1, 2, ... , K, given the actions of its rivals, the producer cannot increase its own profit by choosing any action other than its equilibrium action (Tirole, 1988). In the absence of a cartel agreement or strategic dynamic actions, game theory predicts that the market will stay stable at this point. It is assumed that this market equilibrium point can provide a reasonable prediction of which designs manufacturers are driven to produce under various regulation scenarios. It should be noted however that the Nash equilibrium does not model preemptive competitive strategies by producers. Instead, it assumes that each producer will move to increase its profit while treating competitor decisions as constant.

In order to search for the equilibrium point, an algorithm was employed in which each producer separately optimizes its own profit while holding all competitor producer decisions constant. Each producer's optimization model is solved sequentially, and the process is iterated across producers, in turn optimizing and updating each producer's decisions until all producers converge. Then, a parametric study on K is used to determine the largest value of K that produces a Nash equilibrium with positive producer profits, and this point is taken to be the market equilibrium.

Using the models developed in Sections 7.3.1 - 7.3.5, each producer k will individually attempt to maximize profit by solving the following optimization problem,

$$
\begin{aligned}
&\text{maximize } \left(\sum_{j \in \mathcal{J}_k} q_j \left(p_j - c_j^{\mathrm{V}} \right) - c^{\mathrm{I}} \right) - c_k^{\mathrm{R}} \\
&\text{with respect to } \left\{ \tau_j, x_{1j}, x_{2j}, p_j \right\} \forall j \in \mathcal{J}_k \\
&\text{subject to } 0.75 \le x_{1j} \le 1.50 \\
&\qquad 0.2 \le x_{2j} \le 1.3 \\
&\qquad q_j = s \frac{e^{v_j}}{\sum_{\hat{j} \in \mathcal{J}} e^{v_{\hat{j}}}} \\
&\qquad v_j = \beta_1 p_j + \beta_2 \left(\frac{100}{z_{1j}} \right) + \beta_3 \left(\frac{60}{z_{2j}} \right) \\
&\qquad \mathbf{z}_j = f_\tau \left(\mathbf{x}_j \right) \\
&\qquad c_j^{\mathrm{V}} = c^{\mathrm{B}} + \begin{cases} \beta_4 \exp\left(\beta_5 b_{\mathrm{SI102}} x_{1j} \right) & \text{if } \tau_j = \mathrm{SI102} \\ \beta_6 \left(b_{\mathrm{CI88}} x_{1j} \right) + \beta_7 & \text{if } \tau_j = \mathrm{CI88} \end{cases}
\end{aligned}
\tag{7.14}
$$

where $s = (11/7) \cdot 10^6$, $\beta_1 = -2.86 \cdot 10^{-4}$, $\beta_2 = -0.339$, $\beta_3 = 0.375$, $\beta_4 = 670.51$, $\beta_5 = 0.0063$, $\beta_6 = 26.23$, $\beta_7 = 1642.8$, $b_{\mathrm{SI102}} = 102\mathrm{kW}$, $b_{\mathrm{CI88}} = 90.5\mathrm{kW}$, $c^{\mathrm{B}} = \$7500$, $c^{\mathrm{I}} = \$550 \cdot 10^6$, and c^{R}_k is defined by Eq. (7.11), Eq. (7.12), Eq. (7.13) or zero, depending on which regulation scenario is used. For each producer, competitor products are represented by the set $\{\mathcal{J} - \mathcal{J}_k\}$, and are considered fixed parameters that affect demand (Eq.(7.6)). The first two constraints represent limits on the ability to model variables outside these ranges rather than physical feasibility limits. If these constraints were active, it would represent an inability to model the optimum solution (Papalambros and Wilde, 2000). However these constraints were not active in any of the results, indicating that the optima discussed here are all interior optima and the solutions are valid.

Despite the computational savings gained by creating surrogate models of the engineering performance simulations (splines), the computational burden is still significant. For each producer, separate optimization runs must be computed to determine which combination of vehicles is best for the product line. This combinatorial set of optimization problems is computed for each producer, and each producer model is then iterated several times in the Nash equilibrium solution strategy. In order to reduce the computational burden, the number of deigns per producer was limited to a maximum of two ($n_{\max} = 2$). It was shown that this assumption was reasonable because results of all runs indicate that each producer manufactures only one design, implying that there is a lack of incentive to produce multiple designs (except for the quota regulation case where each producer manufactures both an SI and a CI engine).

7.4 RESULTS AND DISCUSSION

The results of the investigation are summarized in Table 7.1, with a graphical summary of the resulting fuel economy and regulation cost per vehicle provided in Figure 7.4. For each regulation scenario, the table shows the maximum number of producers K that yields a positive-profit Nash equilibrium and the market share per producer design. The use of the aggregate demand model results in each producer making the same decisions at market equilibrium, so Table 7.1 summarizes the decision variables, product characteristics, costs, and profits for a typical producer in each scenario. The fact that all producers are driven to produce the same vehicle design facilitates comparison of the trends that result from each regulation scenario. Additionally, at equilibrium each producer manufactures only a single design rather than a product line (except in the quota case). This result could be changed by modeling cost savings due to economies of scope (Panzar and Willig, 1981), possible commonality among designs (Fellini, 2003), and the use of a heterogeneous model for demand, as in Chapter 5. From Table 7.1, it is also evident that the model predicts equal profits for all regulation scenarios (except the quota case), and all incurred costs are passed to the consumer at equilibrium. This is because the demand model assumes a fixed car-buying population (there is no option not to buy) and does not consider the utility of outside goods.

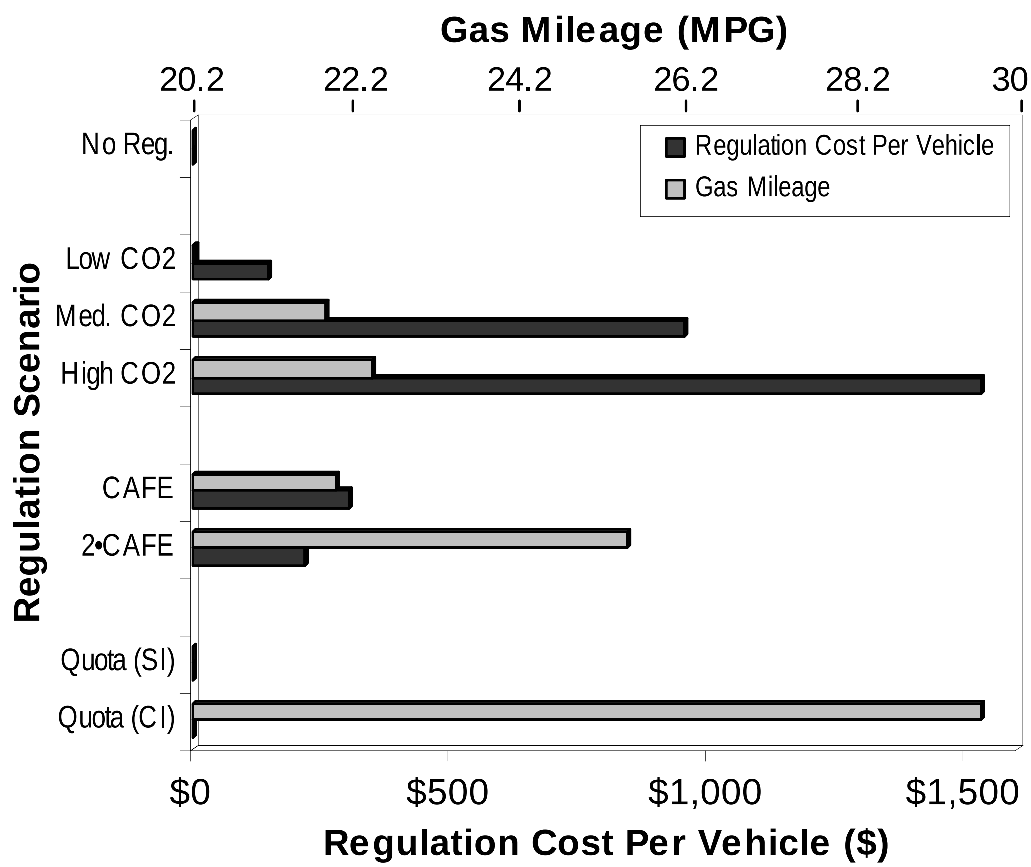


Figure 7.4: Resulting vehicle gas mileage and regulation cost per vehicle under each policy

It is important to take care when interpreting results of an optimization study that is based on a demand regression model. Even if the demand model succeeds in capturing important trends in consumer purchasing preferences according to measurable characteristics, the metrics do not capture purchasing criteria entirely, as the model ignores unmeasured and unobservable characteristics. For example, the model used in this study predicts a preference for vehicles with faster acceleration; therefore, a vehicle that dramatically sacrifices unmeasured characteristics such as maximum speed for a slight improvement of acceleration time will be preferred according to the model. However, in practice a consumer would observe the unmeasured limitations during a road test, especially if the limitations are extreme. To check for this issue, each optimum vehicle design was tested to ensure the vehicle's ability to follow the standard FTP driving cycle and achieve a speed of at least 110mph on a flat road. All vehicle designs in the study passed this test.

Table 7.1: Nash equilibrium results for each regulation scenario

		Regulation Type							
		No Reg.	Low CO_2	Med. CO_2	High CO_2	CAFE	2·CAFE	Quota	
# Producers	K	10	10	10	10	10	10	5	
Market share	q/s	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	11.9%	8.1%
Engine type	M	SI	SI	SI	SI	SI	SI	SI	CI
Engine size	$b_M x_1$	127.9	127.7	114.3	110.3	113.3	88.4	127.9	98.0
FD ratio	x_2	1.28	1.28	1.28	1.27	1.28	1.29	1.28	0.88
Price	p	\$12,886	\$13,031	\$13,719	\$14,259	\$13,058	\$12,772	\$13,372	\$16,083
Gas mileage	z_1	20.2	20.3	21.8	22.4	22.0	25.5	20.2	29.8
Accel. time	z_2	7.46	7.46	7.93	8.10	7.97	9.29	7.46	7.84
Investment cost	c_I	\$550 mil	\$550 mil	\$550 mil	\$550 mil	\$550 mil	\$550 mil	\$550 mil	\$550 mil
Var. cost/vehicle	c_V	\$9,001	\$8,999	\$8,878	\$8,844	\$8,869	\$8,670	\$9,001	\$11,713
Reg. cost/vehicle	c_R/q	\$0	\$147	\$956	\$1,530	\$304	\$217	\$0	\$0
Profit	Π	\$60.5 mil	\$60.5 mil	\$60.5 mil	\$60.5 mil	\$60.5 mil	\$60.5 mil	\$276 mil	\$6.5 mil

7.4.1 Base Case

As a comparative baseline, the no-regulation case was first analyzed ($c^R = 0$). Without regulation, the model predicts ten producers in the small car market. Each producer manufactures a single vehicle with design variables, product characteristics, and costs shown in Table 7.1.

7.4.2 Corporate Average Fuel Economy (CAFE)

Table 7.1 shows that the CAFE regulation results in increased fuel efficiency at a lower manufacturing cost relative to the base case; however, performance is sacrificed, and regulatory costs are incurred (see Figure 7.4). The "2•CAFE" case represents a hypothetical doubling of the penalty for CAFE violation, which resulted in improved fuel economy, *reduced* regulation costs, and reduced vehicle prices relative to CAFE. It was as if an occult hand had intervened to turn an increased penalty into cost reduction (Greenberg, 2004). In both cases, it is predicted that it is profitable for manufacturers to violate CAFE standards and take the penalty in order to increase market share. The model indicates that full compliance with CAFE is dangerous for producers because competitors can produce larger engines, which are in high demand, and capture market share. However, when CAFE penalties are increased, there is less danger of losing market share to a competitor who sells more powerful engines because all producers are subject to a more stringent penalty. Therefore all producers design smaller, cheaper engines with less risk.

In practice, many producers do not currently accrue CAFE penalties and instead treat the CAFE standard as a constraint (Georgiopoulos *et al.,* 2002). One reason for this is the non-modeled extra costs to the producer caused by violation, such as damage to the producer's reputation (which could affect demand), public and government relations, as well as making future compliance more difficult. The results of this study suggest that these non-modeled aspects may provide significant incentives worthy of further consideration.

7.4.3 CO_2 Emissions Tax

Comparing the CO_2 emissions tax to the base case, several trends can be observed. As the tax increases, producers tend to design smaller, more fuel-efficient engines while transferring the added regulation cost to the consumer through an increased vehicle price (Figure 7.4). A low valuation penalty ($2/ton) has little effect on fuel efficiency with the only significant effect being added regulation costs that are in turn passed on to consumers. The median valuation ($14/ton) has a larger impact, increasing fuel-efficiency by 3.3 mpg, while the high valuation ($22/ton) adds only slight improvement in fuel economy at a substantial regulation cost increase over the median case. These trends predict reasonable real-world scenarios, since regulation provides an incentive to produce smaller, more fuel-efficient engines. However, in practice such increases in vehicle costs could lower the demand and sales of vehicles relative to other modes of transportation or other market segments.

7.4.4 Diesel Fuel Sales Quota

In the quota policy, producers were forced to offer diesel engines as a minimum percentage of their vehicle fleet (ϕ = 40%). The results indicate that producers follow this regulation strictly to avoid expensive penalties, producing exactly the minimum required percentage of diesels in their product mix. Since each producer manufactures two vehicle designs, fewer producers result at the market equilibrium.

7.5 CONCLUSIONS

This chapter presented a methodology for analyzing the impact of fuel economy regulations on the design decisions made by automobile manufacturers. The approach integrates models for engineering design, production cost, consumer demand, producer profit, and producer competition toward predicting the impacts associated with different policies that aim to improve fuel economy. Several trends were observed in the policy scenarios examined in this study. One notable observation is that increased regulation penalties can result in cost savings for all parties (e.g., in CAFE scenarios): Without a regulatory standard, producers cannot afford to make smaller, cheaper engines due to competition; however, when all producers are subject to the same regulation costs, then all producers are driven to produce smaller engines with less risk. On the other hand, increased regulation penalties can also lead to diminishing returns in fuel economy improvement with increased regulation penalties (e.g., CO_2 taxation). The observed trends indicate that the cost-benefit characteristics of a given policy can be modeled in a realistic way, and that a holistic integration of costs, performance, consumer preference, and competition may be helpful for evaluating and selecting environmental policies, as well as for choosing regulatory parameter values.

The study also shows that regulation is necessary to provide incentives for producers to design alternative fuel vehicles (e.g., diesels) that cost more to produce. While diesel engines have better fuel efficiency per unit power, gasoline engines are cheaper to manufacture and are therefore preferred by the market. Future investigations that combine engineering, marketing, and policy models with models of changing consumer preferences and driving habits could be used to predict trends for the diffusion of alternative fuel vehicles, possibly avoiding costly investment in products that are unlikely to achieve wide acceptance and help to focus resources and incentives toward solutions that are likely to make the most impact in reducing environmental damage.

The demand model used in this study indicated that individual consumers prefer vehicle acceleration over fuel economy performance. However, as a society, the same individuals may place value on environmental protection, human health, and sustainability that is not captured in the market of individual decisions. For example, while increased CAFE penalties resulted in decreased costs to producers and consumers relative to other fuel economy policies, they also result in smaller, lower-performance vehicles, which are less preferred by individual consumers. Naturally, it will be necessary to balance social versus individual preferences. To quote from the National Academy of Sciences report (2002): "Selection of fuel economy targets will require uncertain and difficult trade-offs among environmental benefits, vehicle safety, cost, oil import dependence, and consumer preferences. The committee believes that these trade-offs rightfully reside with elected officials."

This chapter has taken a step towards developing modeling tools to inform such policy tradeoff decisions. Overall the models presented here were successful in predicting realistic long-term trends resulting from several regulation scenarios. Therefore, the abstract oligopoly analysis was able to provide a useful analytical perspective on market incentives resulting from regulation, demonstrating that policy models that include engineering design decisions can be used to improve our general understanding of the interactions between government policy, industry, consumers, and the environment.

7.6 REFERENCES

[7.1] American Speed Enterprises Ent. Racing Engines Homepage (2003) 23-Jan-2003, *http://www.amerspeed.com*.

[7.2] ASME (2004) American Society of Mechanical Engineers Society Policy on Ethics, Oct-2004, http*://www.asme.org/asme/policies/pdf/p15_7.pdf*.

[7.3] Associated Press (2003) "GM pulling plug on electric cars," *Cable News Network*, 9-Apr-2003, *http://www.cnn.com/2003/TECH/biztech/04/09/electric.cars.ap/*.

[7.4] Balling, R.J., and J. Sobieszczanski-Sobieski (1994) "Optimization of coupled systems: a critical overview of approaches," *AIAA*-94-4330-CP, *Proceedings of the 5th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Panama City, Florida, September.

[7.5] Berkovec, James (1985) "Forecasting automobile demand using disaggregate choice models," *Trans. Res. B.*, v19B n4 p315-29.

[7.6] Bohac, Stani and Tim Jacobs (2002) Unpublished research, office communication (15-Nov-2002) University of Michigan, 1105 WE Lay Auto Lab, 1231 Beal Ave. Ann Arbor, MI 48109-2121.

[7.7] Boyd, J. Hayden and Robert E. Mellman (1980) "The effect of fuel economy standards on the U.S. automotive market: an hedonic demand analysis," *Transpn Res.-A*, v14 p367-78.

[7.8] Brownstone, David, David Bunch, and Kenneth Train (2000) "Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles," *Trans. Res. B.*, v34 p315-38.

[7.9] Chomsky, N. (1999) *Profit Over People: Neoliberalism and Global Order*, Seven Stories Press, New York, NY.

[7.10] Commission of the European Communities (2000) "A review of the Auto-Oil II Programme" *Communication from the Commission* Brussels, 05.10.2000 Com 626.

[7.11] Currim, I., (1982) "Predictive testing of consumer choice models that are not subject to independence of irrelevant alternatives," *Journal of Marketing Research*, v19.

[7.12] Delucchi, Mark and Timothy Lipman (2001) "An analysis of the retail and lifecycle cost of battery-powered electric vehicles," *Transportation Research Part D: Transport and Environment,* v6 n6 p371-404.

[7.13] Diaz, Alejandro (1987) "Interactive solution to multiobjective optimization problems," *International Journal for Numerical Methods in Engineering*, v24 p1865-77.

[7.14] Dixon, L., S. Garber, and M. Vaina (1996) "California's ozone-reduction strategy for light-duty vehicles: an economic assessment," *Rand Publications*, 17-Feb-2004, *http://www.rand.org/publications/MR/MR695.1/.*

[7.15] E Diesels Ltd. Homepage - Diesel Engine Trader (2003) 23-Jan-2003, *http://www.dieselenginetrader.com* Glasgow, Scotland.

[7.16] Fellini, R. (2003) *A Model-Based Methodology for Product Family Design*, Ph.D. Thesis, Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan, USA.

[7.17] Ford Motors N' Engine Replacement Co. Homepage (2003) 23-Jan-2003, *http://www.ford-motors-engines.com/.*

[7.18] Gardner, G. (1996) "GM flips EV1 'on' switch," *Ward's AutoWorld*, March 1996, *http://waw.wardsauto.com/ar/auto_gm_flips_ev/.*

[7.19] Georgiopoulos, P., R. Fellini, M. Sasena and P. Papalambros (2002) "Optimal design decisions in product portfolio valuation," *Proceedings of the 2002 ASME Design Automation Conference*, DAC-34097, Montreal, Quebec, Canada, September 29-October 2.

[7.20] Goldberg, Pinelopi (1996) "The effects of the corporate average fuel efficiency standards" *National Bureau of Economic Research* Working Paper 5673, Yale University, New Haven, Connecticut.

[7.21] Greenberg, P. (2004) "Game uncovered, as if by occult hand," *Arizona Daily Star*, Opinion, 5-Aug-2004, *http://www.dailystar.com/dailystar/opinion/32927.php*.

[7.22] Greene, David and Janet Hopson (2003) "An analysis of alternative forms of automotive fuel economy standards," *for the United States*, Center for Transportation Analysis Publications, *http://www-cta.ornl.gov/cta/Publications/VolStdCh7-ABERDEEN.pdf*.

[7.23] Gu, X., J. Renaud, L. Ashe, S. Batill, A. Budhiraja, and L. Krajewski (2002) "Decision-based collaborative optimization," *ASME Journal of Mechanical Design*, v124 p1-13.

[7.24] Hardin, G. (1968) "The tragedy of the commons" *Science*, v162 p1243-48.

[7.25] Hardy Diesels & Equipment Homepage (2003) 23-Jan-2003, *http://www.hardydiesel.com/*, Jamul, CA.

[7.26] Hazelrigg, G. (1998) "A framework for decision-based engineering design" *ASME Journal of Mechanical Design*, v120 p653-658.

[7.27] IEEE (2004) Institute of Electrical and Electronics Engineers Code of Ethics, Oct-2004, *www.ieee.org*.

[7.28] Kim, H.M. (2003) "Target cascading in optimal system design," *ASME Journal of Mechanical Design* v125 n3 p474-480.

[7.29] Kim, H.M., D.G. Rideout, P.Y Papalambros, and J. L. Stein (2003) "Analytical target cascading in automotive vehicle design," *ASME Journal of Mechanical Design*, v125 n3 p481-489.

[7.30] Li, H. and S. Azarm (2000) "Product design selection under uncertainty and with competitive advantage," *ASME Journal of Mechanical Design*, v122 p411-418.

[7.31] Li, H. and S. Azarm (2002) "An approach for product line design selection under uncertainty and competition," *ASME Journal of Mechanical Design*, v124 p385-392.

[7.32] Manrai, Ajay (1995) "Mathematical models of brand choice behavior," *European Journal of Operational Research*, v82 p1-17.

[7.33] Markel, T., A. Brooker, T. Hendricks, V. Johnson, K. Kelly, B. Kramer, M. O'Keefe, S. Sprik and K. Wipke (2002) "ADVISOR: a systems analysis tool for advanced vehicle modeling" *Journal of Power Sources*, v110 n2 p255-266.

[7.34] Matthews, Scott and Lester Lave (2000) "Applications of environmental valuation for determining externality costs," *Environ. Sci. Tech.*, v34 p1390-95.

[7.35] McFadden, D., (1976) "Quantal choice analysis: a survey," *Annals of Economic and Social Measurement*, v5 p363-90.

[7.36] Michalek, J.J., F.M. Feinberg and P.Y. Papalambros (2004) "An optimal marketing and engineering design model for product development using analytical target cascading," *Proceedings of the Tools and Methods of Competitive Engineering Conference*, Lausanne, Switzerland, April 12-16.

[7.37] Michalek, J.J., F.M. Feinberg and P.Y. Papalambros (2005) "Linking marketing and engineering product design decisions via analytical target cascading," *Journal of Product Innovation Management*, v22 p42-62.

[7.38] Michalek, J.J., P.Y. Papalambros, and S.J. Skerlos (2005) "A study of fuel efficiency and emission policy impact on optimal vehicle design decisions," *Journal of Mechanical Design* (forthcoming).

[7.39] National Academy of Sciences Committee on the Effectiveness and Impact of Corporate Average Fuel Economy Standards, Board on Energy and Environmental Systems, Transportation Research Board, National Research Council (2002) *Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards*, National Academies Press.

[7.40] National Engine Homepage (2003) 23-Jan-2003, *http://www.nationalengine.com*.

[7.41] NSPE (2004) National Society of Professional Engineers Code of Ethics, Oct-2004, *http://www.nspe.org/ethics/eh1-code.asp*.

[7.42] Panzar, J. and R. Willig (1981) "Economies of scope," *The American Economic Review*, v71 n2 p268-72.

[7.43] Papalambros, P. and D. Wilde (2000) *Principles of Optimal Design : Modeling and Computation*, Second Edition, Cambridge University Press, Cambridge UK, 2000.

[7.44] Rebuilt Auto Car Engines *Co.* Homepage (2003) 23-Jan-2003, *http://www.rebuilt-auto-engines.com/*, Spokane, WA.

[7.45] Schumacher, E.F. (1973) *Small is Beautiful: Economics as if People Mattered*, Blond and Briggs Ltd., London.

[7.46] Tappeta, R., J. Renaud (2001) "Interactive multiobjective optimization design strategy for decision based design," *ASME Journal of Mechanical Design*, v123 p205-215.

[7.47] Tirole, Jean (1988) *The Theory of Industrial Organization*, The MIT Press, Cambridge, Massachusetts; London, England 1988.

[7.48] Train, K. (2003) *Discrete Choice Methods With Simulation*, Cambridge University Press.

[7.49] U.S. Congress (1975) *Energy Policy and Conservation Act of 1975* (P.L. 94-163).

[7.50] U.S. Department of Energy, Office of Transportation Technologies, *Energy Efficiency and Renewable Energy Network* (2001) Source: Oak Ridge National Laboratory, 10-Sept-2001 *http://www.ott.doe.gov/facts/archives/fotw108supp.shtml*.

[7.51] U.S. Department of Transportation, Federal Highway Administration, Office of Highway Policy Information (2001) Source: *Ward's 2001 Automotive Yearbook* and the annual *Ward's Motor Vehicle Facts and Figures Publication*, 23 Jan. 2003, *http://www.fhwa.dot.gov///////ohim/onh00/line4.htm*.

[7.52] U.S. EPA (1990) *Clean Air Act*, Environmental Protection Agency, Washington, DC, 15-Nov-2002, *http://www.epa.gov/oar/oaq_caa.html*.

[7.53] U.S. EPA (2001) *National Emission Inventory, 2000 Air Quality Trends Report*, Sept. 2001, Environmental Protection Agency, Washington, DC.

[7.54] Wassenaar, H. and W. Chen (2003) "An approach to decision-based design with discrete choice analysis for demand modeling," *ASME Journal of Mechanical Design*, v125 p490-497.

[7.55] Whitehead, J. W. (2001) *Design and Performance of Derivative-Free Optimization Algorithms Used with Hybrid Electric Vehicle Simulations*, MS Thesis, University of Michigan Department of Mechanical Engineering.

[7.56] Whitney, Daniel, (2001) "Cost performance of automobile engine plants," *International Motor Vehicle Program*, Working Papers, Massachusetts Institute of Technology, 22-Aug-2001 *http://web.mit.edu/ctpid/www/Whitney/morepapers/Engine.pdf*.

[7.57] Wojcik, Charlotte (2000) "Alternative models of demand for automobiles," *Economic Letters*, v68 p113-18.

[7.58] Yee, Jet (1991) "Effect of consumption standards on U.S. gasoline consumption," *Journal of Policy Modeling*, v13 n2 p205-27.

CHAPTER 8

CONCLUSIONS

8.1 SUMMARY

This dissertation has proposed and demonstrated methods for measuring and modeling preferences and coordinating them with engineering design decision models to resolve tradeoffs and achieve joint solutions.

First, the ATC methodology for coordinating decomposed hierarchies of complex systems was introduced in Chapter 3, where current convergence theory was reviewed. It was demonstrated that solving the ATC hierarchy will converge to an inconsistent solution whenever the top level targets are unattainable; however, it is possible to achieve a solution with arbitrarily small inconsistencies by selecting appropriate weighting coefficients. This is particularly important for the applications in the dissertation, since the top level objective does not include an attainable target. A weighting update methodology was introduced to find weighting coefficients necessary to achieve user-specified acceptable tolerances for inconsistency. This method was shown to improve computational time in some cases, although further research is necessary to generalize these results.

In Chapter 4, the ATC methodology was applied to coordinate marketing and engineering design models for product development, where the marketing subproblem is viewed as the top level system and the engineering design subproblem is viewed as the subsystem. In this setup marketing decision variables are price and target *product characteristics*, which are objective, measurable aspects of the product observable by the user. Engineering design decision variables are *design variables*, which are detailed engineering quantities that define the design, but are not necessarily directly observable by the user. In this model, the marketing model statement is to maximize profit with respect to price and target product characteristics, where demand is estimated as a function of the price and target product characteristics using a logit model fit to choice data obtained through an efficient conjoint analysis experimental design survey. The resulting target product characteristics are passed to engineering design. The engineering design subproblem statement is to minimize deviation between the target product characteristics set by the marketing subproblem and those product characteristics achieved by the engineering design with respect to design variables and subject to engineering constraints, where the product characteristics are functions of the design variables. The resulting achieved product characteristics are passed back to marketing. This process is repeated until the two subproblems converge to a joint solution. It was shown with a case study that this joint solution is substantially better, with respect to the profit objective, than a sequential approach where each model is optimized separately without iteration.

In Chapter 5, the ATC methodology was extended to the design of product lines by developing a more sophisticated hierarchical Bayesian mixture model to represent heterogeneity of preferences in the consumer population and fitting the model to choice-based conjoint survey data. Specifically, the marketing subproblem was extended to simultaneously choose target product characteristics for all products in the line, conditional on demand for each product calculated as a function of the price and target product characteristics of all products using draws from the Bayesian model. Then one copy of the engineering design subproblem was made for each product in the line, and the marketing parent subproblem was coordinated with all of the engineering design subproblems iteratively until convergence to a joint solution. Unlike the simple aggregate demand model used in Chapter 4, the heterogeneous demand model produced a multimodal profit surface, and a high quality local minimizer was found using multi-start with gradient-based algorithms. It was shown again that the joint solution produced by ATC coordination yields superior profit to the solution obtained through a sequential approach. Also, significantly different results were obtained when using different heterogeneity representations, even when designing a single product, suggesting a need to use the most general form of heterogeneity. The use of ATC to coordinate the separate, but related problems of designing multiple products in a line improves scalability of the problem by avoiding the high dimensionality and complexity of solving multiple product optimization problems simultaneously.

In Chapter 6, this model was extended further to include manufacturing investment and allocation decisions by allowing marketing to set cost and production volume targets to be achieved by manufacturing and product design subproblems. Some of the inherently discrete decisions relevant to manufacturing were formalized into a continuous space. In the case study this methodology enabled a local solution close to the solution of the relaxed problem to be found; however, the local minima created by the transformation of discrete decisions and constraints into a continuous space require a global search algorithm to find a true global solution in general. This application highlights the need for further work to extend the ATC methodology to problems with discrete variables in order to resolve tradeoffs without the computational burden of performing branch and bound, or other such techniques, on the entire ATC hierarchy.

While Chapters 4-6 focused on coordination within a single firm to achieve maximum profit, Chapter 7 expanded scope to consider the effects of multiple profit-seeking firms competing in a regulated marketplace. Individuals as consumers have different preferences than individuals as members of society, and unregulated marketplaces in many cases tend to respond more to private consumer preferences than to public social preferences. While an individual profit-seeking firm may choose to concern itself primarily with modeling consumer preferences, producing positive social results requires a balance between private and public interests. In particular, legislation of products affects and is affected by engineering design. By studying the interactions of markets and regulation with design decisions, engineers can improve their capability to predict consequences of their actions and therefore can contribute meaningfully to the dialogue regarding legislation that directly affects the design process as well as the degree to which that process is aligned with the public interest. Specifically, Chapter 7 abandons the ATC coordination because of the problem simplicity, using instead a single all-at-once optimization loop to represent decisions of an individual producer, and game theory is utilized to simulate competition among firms to predict design choices made at market equilibrium. A case study of vehicle design, regulated by fuel economy and emissions standards, is investigated, and the characteristics of the market equilibrium are compared under several alternative policy scenarios. Results show that while increasing penalties can lead to increased cost for some policies, it can also lead to decreased costs for other policies as a result of competitive interactions. Overall, the case study demonstrates a step toward developing models to inform policy tradeoff decisions.

While it is possible to integrate preference models and engineering decision models into a single optimization loop, as in Chapter 7, there are several advantages to the decomposition and coordination approach taken in the rest of the dissertation. First, ATC allows joint solutions to be achieved while maintaining disciplinary modeling focus, providing structure for organizing the different aspects affecting product development, rigorously defining interactions among these aspects, and relieving the need for a single modeler to become an expert in all areas. Second, ATC provides advantages for scalability. With large problems that contain many variables, the decision space of a problem solved all at once, in one large optimization loop, can have high dimensionality and be highly nonlinear and constrained. These issues can amplify practical issues associated with solving the problem, including scaling and numerical issues. ATC decouples the problem where possible, producing individual subproblems, each with relatively low dimensionality and nonlinearity, and with fewer constraints. This means that the individual subproblems are typically easier to solve, and some problems that cannot be solved otherwise can be solved through decomposition. While other decomposition methods exist, ATC is the only method with proven convergence properties capable of coordinating an arbitrarily large hierarchy of systems and subsystems. Third, ATC has been shown to improve computational time in some situations. This is not a universal claim, since many problems have been solved more quickly all at once; however, given the competitive computational efficiency and the opportunity to utilize distributed computational resources simultaneously, ATC proves to be a useful tool for many problems. In particular, product line design problems benefit from the separation of the highly decoupled design of individual products in the line. Finally, the modularity of the ATC hierarchy facilitates model changes and extensions easily, as demonstrated by the two extensions explored in this dissertation. Additional subproblems can be added, and existing subproblems can be altered in such a way that the only changes affecting a particular discipline are well-defined by the targets and linking variables shared with that discipline, facilitating concurrent design and modeling improvements.

8.2 SUMMARY OF CONTRIBUTIONS

Several distinct steps were taken toward the coordination of preference models with engineering design decision models.

1. The ATC convergence proof was clarified for cases with unattainable targets, and a weighting update routine was proposed for reducing resulting system inconsistency to a user-defined tolerance.
2. A methodology was proposed to apply ATC to coordinate marketing models of demand and profit with engineering design models of product feasibility and

performance to achieve joint solutions that are demonstrably superior to solutions arrived at independently. To marketing, the methodology offers the ability to indirectly accommodate technical constraints and ensure realizable solutions; and to engineering, the methodology offers a means to resolve technical objectives through coordination with downstream objectives.

3. An extension to the methodology was proposed to design lines of products via utilization of preference models that account for heterogeneity in the population and coordination of this model with a set of product design models. This methodology offers a new approach to product line design, accommodating variables with continuous domains, addressing non-monotonic product characteristics with ideal points that vary in the population, and enabling joint design of multiple products while mitigating problems associated with search of large, highly constrained design spaces by taking advantage of the coupling structures in the problem.
4. A second extension to the methodology was proposed to coordinate manufacturing investment and allocation decisions with design and marketing decisions. This model accounts for tradeoffs between cost reductions that may be obtained by compromising design of the product line to conform to the capabilities of existing or inexpensive equipment vs. the loss in revenue associated with failing to deliver the most desirable product.
5. Finally, a method was introduced to study the effects of policy for balancing public and private preferences on the resulting design decisions of profit-seeking firms in a competitive marketplace. For policymakers, the method provides a tool to predict effects and inform policy decisions; for engineering designers, the method provides a larger context for the effects of their decisions and gives them tools with which to participate in the ongoing dialogue about policy that directly affects their work.

8.3 OPEN QUESTIONS

A number of open questions were highlighted throughout the dissertation for future work. In Chapter 3, global convergence properties of ATC were explored and expanded upon; however, local convergence properties have yet to be well-defined with respect to decomposed systems, and much work remains to study these properties for ATC to determine general properties of computational efficiency. Also, as highlighted in Chapter 6, there exists a need to extend ATC theory for application to problems with discrete and categorical variables. Further examination is also needed to study the properties of ATC using non-gradient-based methods when solving each of the subproblems. Finally, the effects of using ATC on search of a space with multiple local minima deserve further study. In particular, the weighting update method allows quick movement with large inconsistencies for small weighting coefficients and slower, more precise movement for large weighting coefficients. With problems that are solved first with small weighting coefficients, the solution will move toward the optimum of the relaxed parent problem, even moving into regions of the space highly unachievable by subsystems. Depending upon the shape of the problem and the sequence of weighting coefficients used, the search procedure will have an effect on which local minima are located. Understanding this effect more precisely could lead to more directed research about when to use the method and how to generate good starting points.

In Chapter 4 the ATC methodology was applied successfully to coordinate engineering and marketing models and obtain joint solutions. While results of coordination are demonstrably better than solving the two problems independently, they do not resolve, and may in fact amplify, any problems with model validation. For example, while validation techniques for engineering performance models are commonplace, the optimization of an engineering model with respect to several specific product characteristics, while inevitably ignoring others, can lead to a solution strong in the modeled characteristics at the expense of unmodeled characteristics. Validation of a design achieved through model optimization must consider a holistic evaluation of the design to ensure that unmodeled characteristics are not drastically compromised at the solution. Also, the marketing demand models used in this dissertation are based on reported preference data. One method of validation for these models is to test predictions against a holdout sample, assessing the degree to which a model fit to a specific set of data will extrapolate to other choice situations. However, this test still provides only a measurement of stated choice, which can often differ dramatically from observed choices in the marketplace. Furthermore, and importantly, preference is not the same as demand since demand depends critically on other factors such as availability, distribution, familiarity, word of mouth, advertising, and store or shelf location. ATC offers a framework that could accommodate models of additional factors such as these for coordination; however, they were not modeled here. Due to these unmodeled factors, model predictions will differ from observed market behavior. Instead, the models predict the trends of demand, all else being equal. When all else is not equal, actual demand differs, and validation becomes very challenging. However limiting these issues may be, they do not result from the preference coordination introduced in the dissertation. In fact, the ATC preference coordination serves only to reduce the decision space of the marketing subproblem, as if projecting engineering constraints into the product characteristic space. So, applying preference coordination adds no new validation issues to a marketing optimization formulation; rather, it restricts the space to realizable options, which may, in fact, remove some otherwise unrealistic areas of the space and improve the model for validation relative to a pure marketing optimization problem.

Secondly is the problem of model bounds. The conjoint task covers a set of levels for each product characteristic. If the optimization were to yield a point outside of the bounds tested with conjoint, it would indicate a limitation of the model to predict preference in that region. In Chapter 4 the optimal solution is an interior solution, and model boundedness is not a problem. In Chapter 5, however, using the heterogeneous demand model, a few of the results, in particular price of the expensive product, were boundary solutions. This is expected since the ranges of levels in the conjoint task were based on products existing in a competitive market, and the optimization represented a monopolist case with a simple outside good. It is expected that monopolist prices will be higher than those resulting in a competitive marketplace. However, the other boundary characteristics suggest that the model contains individuals who prefer items outside the range observed in the market. This is an entirely plausible situation, since the market does not perfectly capture preferences, and since the survey respondents are not perfectly representative of the population; however, the solution lies outside the range of the model, and conclusions are difficult to make. It is important, therefore, when intending to use a model for optimization to study conjoint levels outside the range observed in the market in order to measure existing ambivalence or disdain for characteristics outside of that range and ensure interior solutions. Pretesting the conjoint task could facilitate choice of appropriate levels; however, more research on how best to select these levels is warranted.

The degree to which these models are sensitive to assumptions deserves further study as well, such as the assumption that utility is linear in part-worths and interaction terms are negligible. Chapter 5 demonstrated that assumptions about the shape of preference heterogeneity can strongly affect optimal line solutions. Other assumptions, such as externalities set as fixed parameters, may have more predictable effects. In particular, it would be worth further study to examine the effect of the model form defining the distribution of preferences over the population on the shape of the resulting demand surface. If some assumptions about this distribution lead to convenient properties for the demand surface, such as unimodality, then these assumptions should be used in cases where they are likely to hold. Furthermore, the demand surface in the marketing subproblem itself contains a significant amount of symmetry for multi-product cases.

Exploring methods to exploit this symmetry would improve the computational efficiency, and therefore the scalability of the method for large product lines.

As mentioned in Chapter 2, the perception of product characteristics is not explicitly modeled here. Modeling of the perception process, particularly when perception is heterogeneous across the population, could add realism and improve predictability of the models here.

In Chapters 5, 6, and 7, it was assumed that all products are designed and built from scratch with no shared components or other savings due to commonality. The addition of modeled cost savings due to commonality and resulting issues with product differentiation would extend the relevance of the product line studies. ATC preference coordination may provide a useful framework for studying product commonality, in particular for quantifying the tradeoff between cost savings resulting from part-sharing and the reduced demand resulting from a perceived lack of differentiation or from the inability to meet individual product targets under part-sharing constraints.

Further development of models to account for competition is also needed. Chapter 7 uses game theory to predict competitor decisions at market equilibrium, but market dynamics are ignored (the market is never truly at equilibrium), and common anticipatory competitive strategies are not considered. The alternative of modeling competitor products as fixed is also unrealistic, particularly for large design moves, which are likely to elicit competitive responses. Deeper examination of issues related to market competition would improve relevance to applied industrial problems.

Finally, the effects of policy modeled in Chapter 7 could be extended to examine alternative fuel vehicles and study the effects of CAFE standards for full product lines, where high fuel economy in one vehicle segment enables lower fuel economy in another. Automotive companies exert significant effort to comply with CAFE standards in strategic ways, and they could benefit from better modeling tools. As well, legislative bodies could benefit from tools better able to anticipate possible market responses to legislation such as the automaker push to sell SUVs to consumers as a way to take advantage of differential car and truck CAFE standards. Additionally, the ability to model effects of policy directly on consumer behavior, for example, the effects policy alternatives such as an increased gasoline tax on consumer purchase choices, driving behaviors, and the differential impact on lower-income individuals, would further assist policy analysis and decision-making.

In summary, this dissertation has taken steps toward development of tools for modeling the various preferences of product design stakeholders and coordinating these models with engineering design decision-making. It is the hope that these models, and others like them, will help design engineers and managers to understand better the relationship between their decisions and the networks of interests and preferences upon which they have impact so that wiser, more informed decision-making can be realized.

APPENDIX

DISCRETE CHOICE MODELS

This appendix serves to provide a detailed introduction to discrete choice models for readers new to the domain to supplement the material in the chapters.

Utility

Utility is a ubiquitous concept in economics as an abstract measurement of the degree of goal-attainment or want-satisfaction provided by a product or service relative to alternatives. We cannot measure directly how much utility a person may gain from a product; however, we can make inferences about utility based on the person's behavior, if we presume that people act *rationally*. In computer science, a *rational* agent is defined as one that acts to attain its goals. Likewise, in economics we assume that a rational person acts to increase her utility.

All else being equal, if a rational consumer is given a choice between product A, with utility $u_A = 1$, and product B, with utility $u_B = 2$, she will choose product B because it provides more utility. In general, given a set of alternatives $j = \{1, 2, ..., J\}$, a rational person will choose the alternative that provides the highest utility, so that alternative j is chosen if $u_j > \{u_{j'}\}_{\forall j' \neq j}$. This model does not take into account the degree to which the utility of one product exceeds the utility of another. For instance, if $u_A = 1$ then product B will be chosen if $u_B > 1$, regardless of whether $u_B = 1.0001$ or $u_B = 1000$. In reality, uncertainty in utility estimates would lead one to be more confident in predicting choice B if $u_B = 1000$ and less confident if $u_B = 1.0001$

Random Utility Discrete Choice Models

In general, we cannot measure utility (predict choices) exactly because, for example, we may not be able to observe or measure every characteristic of the individual, product, or choice situation that affects choice behavior; however, if we can observe some information about the individual, the product, or the choice situation, we can use that information to help predict choice. So, in random utility models we presume that the utility u_{ij} provided to individual i by product j is composed of a deterministic component v_{ij}, which can be calculated based on observed characteristics, and a stochastic error component ε_{ij}, which is unobserved, so that

$$u_{ij} = v_{ij} + \varepsilon_{ij}\,. \tag{A.1}$$

Later we will discuss how to estimate the observable component of utility v_{ij} for individual i and product j using data, but for now we take it as given. Because we never observe the error component ε_{ij}, we do not have enough information to predict a specific individual's choice on a specific choice occasion, but, as in regression, we can make predictions about the patterns of choices over many individuals and many choice occasions. The probability P_{ij} of individual i choosing product j from a set of products is

$$\begin{aligned} P_{ij} &= \Pr\left[u_{ij} > \left\{ u_{ij'} \right\}_{\forall j' \neq j} \right] \\ &= \Pr\left[v_{ij} + \varepsilon_{ij} > \left\{ v_{ij'} + \varepsilon_{ij'} \right\}_{\forall j' \neq j} \right] \end{aligned} \tag{A.2}$$

Distributions for the ε Error Terms

The ε error terms are unobserved random variables that are described by a probability distribution. In general, this may be a joint distribution of all the error terms, so we use the vector $\boldsymbol{\varepsilon}_i = [\varepsilon_{i1}\ \varepsilon_{i2}\ ...\ \varepsilon_{in}]^T$, which aggregates the error terms for all products, and describe its probability distribution by the cumulative distribution function (CDF) $F_\varepsilon(\boldsymbol{\varepsilon})$ and its corresponding probability density function (PDF) $f_\varepsilon(\boldsymbol{\varepsilon})$.

Let us examine a simple case where the choice set is composed of only two products, j and j', and we can generalize later. In this case

$$
\begin{aligned}
P_{ij} &= \Pr\left[v_{ij} + \varepsilon_{ij} > v_{ij'} + \varepsilon_{ij'} \right] \\
&= \Pr\left[\varepsilon_{ij'} < v_{ij} - v_{ij'} + \varepsilon_{ij} \right]
\end{aligned}
\tag{A.3}
$$

For a given value of ε_{ij} this is $F_\varepsilon(\varepsilon_{ij}, v_{ij}-v_{ij'}+\varepsilon_{ij})$: the CDF of the joint random variable distribution evaluated at the point $(\varepsilon_{ij}, v_{ij}-v_{ij'}+\varepsilon_{ij})$, i.e., the probability that the random variable $\varepsilon_{ij'}$ is less than the value $(v_{ij}-v_{ij'}+\varepsilon_{ij})$, given ε_{ij}. However, ε_{ij} is a not deterministic fixed value, but instead is itself described by a probability density function $f_\varepsilon(\varepsilon_{ij})$. Therefore, the probability can be calculated by integrating over all values of ε_{ij}

$$
\begin{aligned}
P_{ij} &= \int_{\varepsilon_{ij}=-\infty}^{\infty} F_\varepsilon\left(\varepsilon_{ij}, v_{ij} - v_{ij'} + \varepsilon_{ij}\right) d\varepsilon_{ij} \\
&= \int_{\varepsilon_{ij}=-\infty}^{\infty} \int_{\varepsilon_{ij'}=-\infty}^{v_{ij}-v_{ij'}+\varepsilon_{ij}} f_\varepsilon\left(\varepsilon_{ij}, \varepsilon_{ij'}\right) d\varepsilon_{ij'} d\varepsilon_{ij}
\end{aligned}
\tag{A.4}
$$

In general, for a set of products

$$
\begin{aligned}
P_{ij} &= \Pr\left[v_{ij} + \varepsilon_{ij} > \left\{ v_{ij'} + \varepsilon_{ij'} \right\}_{\forall j' \neq j} \right] \\
&= \Pr\left[\left\{ \varepsilon_{ij'} < v_{ij} - v_{ij'} + \varepsilon_{ij} \right\}_{\forall j' \neq j} \right] \\
&= \int_{\varepsilon_{ij}=-\infty}^{\infty} \left(\int_{\varepsilon_{i1}=-\infty}^{v_{ij}-v_{i1}+\varepsilon_{ij}} \int_{\varepsilon_{i2}=-\infty}^{v_{ij}-v_{i2}+\varepsilon_{ij}} \cdots \int_{\varepsilon_{iJ}=-\infty}^{v_{ij}-v_{iJ}+\varepsilon_{ij}} f_\varepsilon(\boldsymbol{\varepsilon})\, d\tilde{\boldsymbol{\varepsilon}}_{ij} \right) d\varepsilon_{ij} \\
&\quad \text{where } d\tilde{\boldsymbol{\varepsilon}}_{ij} = d\varepsilon_{iJ} \ldots d\varepsilon_{i(j+1)} d\varepsilon_{i(j-1)} \ldots d\varepsilon_{i2} d\varepsilon_{i1}
\end{aligned}
\tag{A.5}
$$

The Probit Model

Most commonly in statistics, unobserved random error terms are taken to be normally distributed (e.g., least squares regression, etc). The central limit theorem provides a theoretical justification for this choice in the absence of other information about distributional forms. If $f_\varepsilon(\boldsymbol{\varepsilon})$ is assumed to be a multivariate joint normal distribution with mean zero and covariance matrix Λ, this is called the *probit model*. The probit model allows for quite a general model; however, it does not yield a closed form solution, and it requires multidimensional integration.

Some econometricians have alternatively used a restricted form of the probit model where error terms are taken to be independently and identically distributed: i.e., the covariance matrix Λ is assumed to be diagonal. In this case, P_{ij} reduces to a single dimensional integral:

$$\begin{aligned} P_{ij} &= \Pr\left[\left\{\varepsilon_{ij'} < v_{ij} - v_{ij'} + \varepsilon_{ij}\right\}_{\forall j' \neq j}\right] \\ &= \prod_{j' \neq j} \Pr\left[\varepsilon_{ij'} < v_{ij} - v_{ij'} + \varepsilon_{ij}\right] \\ &= \int_{\varepsilon_{ij}=-\infty}^{\infty} \left(\prod_{j' \neq j} F_{\varepsilon}\left(v_{ij} - v_{ij'} + \varepsilon_{ij}\right)\right) dF_{\varepsilon}\left(\varepsilon_{ij}\right) \end{aligned} \tag{A.6}$$

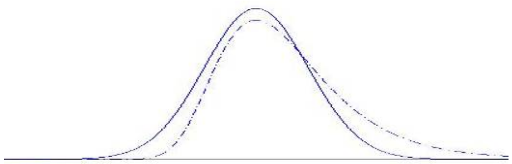
This simplified form is desirable; however, the assumption of independence of the error terms is a restriction that leads to specific implications, which will be discussed later.

The Logit Model

To simplify matters more, econometricians often use an alternative assumption for the distribution of the error terms: Instead of normal, error terms are assumed to be independently and identically distributed (iid) following the double exponential (Gumbel Type II extreme value) distribution:

$$\begin{aligned} F_{\varepsilon}\left(\varepsilon_{ij}\right) &= \exp\left(-e^{-\varepsilon_{ij}}\right) \\ f_{\varepsilon}\left(\varepsilon_{ij}\right) &= e^{-\varepsilon_{ij}} \cdot \exp\left(-e^{-\varepsilon_{ij}}\right) \end{aligned} \tag{A.7}$$

This assumption yields the *logit model*. Unlike the normal distribution, there is no theoretical reason to believe that the double exponential is a good assumption for the error terms; however, under this assumption P_{ij} reduces to a simple, explicit, usable form, and studies have shown that results obtained under this logit assumption are nearly indistinguishable from those produced by the probit model, except when large amounts of data are available. So, the logit assumption is a useful "engineering approximation." The standard normal and double exponential PDFs are shown below:



Plugging the double exponential distribution in for f_ε, we have

$$P_{ij} = \int_{\varepsilon_{ij}=-\infty}^{\infty} e^{-\varepsilon_{ij}} \exp\left(e^{-\varepsilon_{ij}}\right)\prod_{j'\neq j}\exp\left(-e^{v_{ij}-v_{ij'}+\varepsilon_{ij}}\right)d\varepsilon_{ij} \tag{A.8}$$

since $v_{ij} - v_{ij} = 0$, the exponential term can be brought inside the product, so that the expression is rewritten as

$$\begin{aligned} P_{ij} &= \int_{\varepsilon_{ij}=-\infty}^{\infty} e^{-\varepsilon_{ij}} \prod_{j'}\exp\left(-e^{-\left(v_{ij}-v_{ij'}+\varepsilon_{ij}\right)}\right)d\varepsilon_{ij} \\ &= \int_{\varepsilon_{ij}=-\infty}^{\infty} e^{-\varepsilon_{ij}} \exp\left(-\sum_{j'} e^{\left(v_{ij}-v_{ij'}+\varepsilon_{ij}\right)}\right)d\varepsilon_{ij} \\ &= \int_{\varepsilon_{ij}=-\infty}^{\infty} e^{-\varepsilon_{ij}} \exp\left(-e^{-\varepsilon_{ij}}\sum_{j'} e^{v_{ij'}-v_{ij}}\right)d\varepsilon_{ij} \end{aligned} \tag{A.9}$$

We can solve this integral with a change of variables. Let $t = \exp(-\varepsilon_{ij})$. Then $dt = -\exp(-\varepsilon_{ij})d\varepsilon_{ij}$ and $d\varepsilon_{ij} = -dt/t$. For the integration limits: as ε_{ij} approaches infinity, t approaches zero, and as ε_{ij} approaches negative infinity, t approaches infinity. Rewriting the integral in terms of t:

$$
\begin{aligned}
P_{ij} &= -\int_{t=\infty}^{0} \exp\left(-t\sum_{j'} e^{v_{ij'}-v_{ij}}\right)dt \\
&= \left(-\sum_{j'} e^{v_{ij'}-v_{ij}}\right)^{-1} \exp\left(-t\sum_{j'} e^{v_{ij'}-v_{ij}}\right)\Bigg|_{0}^{\infty} \\
&= \left(\sum_{j'} e^{v_{ij'}-v_{ij}}\right)^{-1} = \left(e^{-v_{ij}}\sum_{j'} e^{v_{ij'}}\right)^{-1}
\end{aligned}
\tag{A.10}
$$

$$
P_{ij} = \frac{e^{v_{ij}}}{\sum_{j'} e^{v_{ij'}}} \tag{A.11}
$$

The iid double exponential error term assumption has led to a very simple formula for choice probabilities with appropriate properties: choice probabilities range from zero to one and sum to one over all alternatives in the choice set.

Independence of Irrelevant Alternatives

It is important to be aware that assuming independence of the error terms (in both the logit and the restricted probit models) gives rise to a property called *independence from irrelevant alternatives*, or IIA. We know that if a new alternative product is added to the choice set, some individuals who would otherwise have chosen a product in the initial choice set will instead choose the new product. The IIA property means the *ratio* of choice probabilities between any two alternatives is unaffected by the presence of a third alternative, and any new alternative introduced to a choice set will take its choice share proportionally from all other alternatives in the choice set. For the logit model, this is easy to show:

$$
\frac{P_{i\mathrm{A}}}{P_{i\mathrm{B}}} = \frac{e^{v_{i\mathrm{A}}} \Big/ \sum_{j'} e^{v_{ij'}}}{e^{v_{iB}} \Big/ \sum_{j'} e^{v_{ij'}}} = \frac{e^{v_{i\mathrm{A}}}}{e^{v_{iB}}} \tag{A.12}
$$

The IIA property is also known as the "red bus, blue bus problem" because of a famous illustration of this property: Let's say commuters have the two options {car, blue bus} available to them and gain equal utility from each ($v_{\mathrm{CAR}} = v_{\mathrm{BLUEBUS}}$), therefore choosing each with probability 0.5. If a new product is added to the choice set that is very similar to one of the existing products {car, blue bus, red bus} with equal utility, the IIA property implies that the new product will draw choice proportionally from all other alternatives, so that $P_{\mathrm{CAR}} = P_{\mathrm{BLUEBUS}} = P_{\mathrm{REDBUS}} = 0.333$. In reality we would expect the red bus to draw far more commuters from the blue bus than from car travel since the two busses are very similar. Choice probabilities will likely be closer to $P_{\mathrm{CAR}} = 0.5$, $P_{\mathrm{BLUEBUS}} = P_{\mathrm{REDBUS}} = 0.25$. The IIA property also would imply, for instance, that the ratio of votes for Democratic and Republican candidates is unaffected by the presence of a third party candidate. Thus there are limitations to the applicability of models that possess the IIA property; however, a number of extensions exist to mitigate or eliminate this problem, and in many practical applications the IIA property is not problematic. For the remainder of this appendix the simple logit model will be used; however, more advanced models are explored in Chapter 5.

Functional Forms for the Observable Component of Utility *v*

The preceding discussion presumes that the observable component of utility v_{ij} is known for each individual i and each product j. We said v_{ij} is observable in that it is a function of the observable characteristics of the product, the individual, and the purchase situation. For now, we will limit our discussion so that v_j depends only on the characteristics of the product, i.e., all individuals have the same *observable* component of utility, individual differences are described only by the random error term, and the index i is dropped. The term *product characteristics* is used specifically to describe objective, measurable aspects of the product that are observed by and relevant to the consumer during the choice process. For example, fuel economy of a vehicle may be considered a product characteristic, but "sportyness" is not a characteristic because it is perceived subjectively, and transmission ratio is probably not a characteristic since it is generally not observed directly by customers (except for special cases), but rather by engineering designers. The value of the product characteristics of product j are written as the real-valued vector $\mathbf{z}_j$, and v_j is a function of $\mathbf{z}_j$ as well as the product's price p_j, which, by convention, is not included in $\mathbf{z}_j$.

Just as in regression, we do not know, in general, the functional form relating $\mathbf{z}_j$ and p_j to v_j; however, if we have experience with choice models and experience in the problem domain, we may be able to posit reasonable functional relationships that produce good predictions. For example, researchers Boyd and Mellman (1977) proposed a functional relationship for vehicles including price p_j, gas mileage z_{j1}, and performance measured as time to accelerate from 0-60 mph z_{j2}, among other characteristics. Their model proposed that

$$v_j = \beta_0 p_j + \beta_1 \left(\frac{1}{z_{j1}} \right) + \beta_2 \left(\frac{1}{z_{j2}} \right), \tag{A.13}$$

where β_0, β_1, and β_2 are coefficients. If we could observe v_j directly, then we could collect data for various values of p_j and $\mathbf{z}_j$ and perform an ordinary regression to find the best values for the β coefficients; however, v_j is not observed: Only choice is observed. Given past data on choices among vehicles with various values for p_j and $\mathbf{z}_j$, it is possible to find values for the β coefficients that result in choice predictions that best match the observed choice data, as we would do in simple regression, using a technique called *maximum likelihood*.

Maximum Likelihood

In this case, we have 1) assumed the distribution of the error terms (double exponential for logit), and 2) assumed the functional form of v_j with respect to observed characteristics. Now we want to find the best model parameters (β coefficients) to match observed data, given the model form. To do this we search for the coefficients that maximize the likelihood that the choice model (with coefficients β) would generate the data we observed: i.e., the model predicts choices probabilistically, and we want to maximize the likelihood that choices predicted by the model would be exactly those observed. On a specific choice occasion, the probability of the model predicting the same choice as the one observed for individual i is

$$\prod_j P_{ij}^{\Phi_{ij}} , \tag{A.14}$$

where $\Phi_{ij} = 1$ if individual i chooses product j, and $\Phi_{ij} = 0$ otherwise. If this process is repeated for many individuals, the total number of individuals choosing product j is given by $\Sigma_i \Phi_{ij}$, and the probability of the model generating the observed choices is

$$\prod_j P_j^{\sum_i \Phi_{ij}} . \tag{A.15}$$

We are searching for the values of β that maximize this quantity. To simplify calculations and avoid numerical difficulties, it is common practice to maximize the log of the likelihood, which has the same maximum, rather than maximizing the likelihood directly. This is called the *log-likelihood*, often written *LL*. The maximum (log) likelihood β terms are therefore:

$$\hat{\boldsymbol{\beta}} = \arg\max_{\boldsymbol{\beta}} \left(\sum_j \sum_i \Phi_{ij} \log P_j \right). \tag{A.16}$$

where P_{ij} is given by the logit model.

Example

Let's suppose our choice set consists of four vehicles with prices and characteristics shown below

	A	B	C	D
p_j (\$1000s)	15	15	20	20
z_{j1} (mpg)	25	35	25	35
z_{j2} (sec)	6	8	8	6

Suppose we ask 100 people which vehicle each would choose, and we find that 25 choose product A, 30 choose product B, 5 choose product C, and 40 choose product D. Using the logit model for choice probabilities P_j and the Boyd and Mellman model as the form of the utility function v_j we would solve for the β terms as:

$$\underset{\beta_0,\beta_1,\beta_2}{\text{maximize}}\left(25\log P_{\text{A}} + 30\log P_{\text{B}} + 5\log P_{\text{C}} + 40\log P_{\text{D}}\right)$$

$$\text{where } P_j = \frac{\exp\left(\beta_0 p_j + \beta_1\left(\frac{1}{z_{j1}}\right) + \beta_2\left(\frac{1}{z_{j2}}\right)\right)}{\sum_{j'}\exp\left(\beta_0 p_{j'} + \beta_1\left(\frac{1}{z_{j'1}}\right) + \beta_2\left(\frac{1}{z_{j'2}}\right)\right)} \qquad \text{(A.17)}$$

.

To find the maximum by hand, we can take the gradient of the function, set it equal to zero, and solve the resulting system of equations. Alternatively, we can use an optimization algorithm such as Excel Solver to find the values for the β terms that maximize the log likelihood. Using either technique, the solution is β_0 = -0.132, β_1 = -99.0, β_2 = 22.8. We see that β_0 is negative, indicating that increasing price will decrease utility, β_1 is negative, indicating that increasing fuel economy (decreasing $1/z_{j1}$) will increase utility, and β_2 is positive, indicating that increasing 0-60 time (decreasing $1/z_{j2}$) will decrease utility. Note that five individuals chose product C, even though it is more expensive, has worse fuel economy, and worse performance. While this goes against the utility trends in a deterministic utility model, random utility choice models, such as the logit model, allow for unobserved characteristics that may affect the decisions of individuals while still capturing the overall trends.

Using these newly obtained beta values, and the corresponding model of choice, we can now make predictions about new products or changes to existing products. Suppose we wanted to lower the price of product C to attract more buyers. How much would we have to lower the price to double market share (attract 10 out of 100 buyers instead of 5)? To make this prediction, we would simply solve

$$P_{\mathrm{C}} = \frac{\exp\left(\beta_0 p_{\mathrm{C}} + \beta_1 z_{\mathrm{C}1} + \beta_2 z_{\mathrm{C}2}\right)}{\sum_{j'} \exp\left(\beta_0 p_{j'} + \beta_1 z_{j'1} + \beta_2 z_{j'2}\right)} = 0.10\,. \tag{A.18}$$

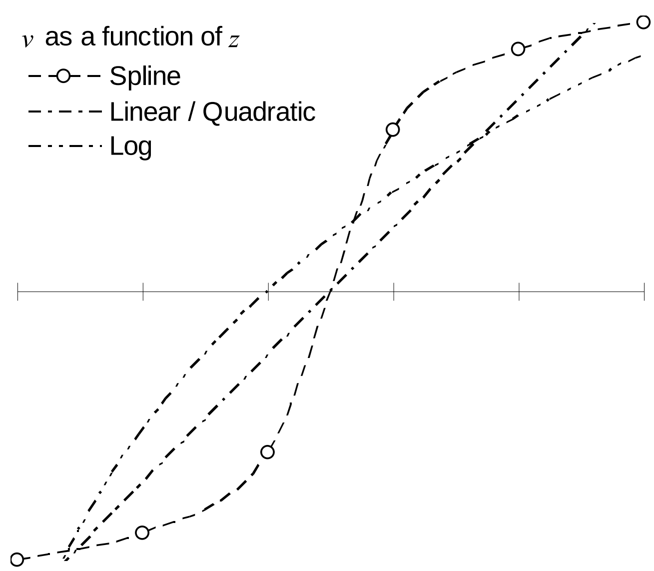
for p_{C} using the beta values and characteristic values from above. In this case the answer is \$14,300. So, vehicle C, with the least desirable characteristics, would have to drop its price below the prices of competitors in order to capture 10% of the market.

Functional Forms for the Observable Component of Utility *v*

In the previous example we used an assumed functional form for v established by experts. In general, how does one know what functional form to use, and what kind of functional form for v should be assumed when there is no prior knowledge about the relationship between p, v, and $\mathbf{z}$? In general we may not have good intuition about what functional forms to assume for a particular product and set of product characteristics. One method is to simply try different functional forms and see which one results in the highest likelihood value. However, this can be dangerous in the absence of information about the problem because more general forms (say, assuming a quadratic rather than linear relationship) will always yield higher likelihood than more restrictive forms; however, one must be wary of overfitting the data. So, in general, this might be a reasonable technique for testing whether the price relationship is linear or log; however, it is not a good idea to blindly test arbitrary functional form assumptions and pick the highest likelihood result.

Discretization of the Product Characteristics z and Price *p*

A more general technique is to divide the relevant range of each product characteristic in **z** and price p into discrete *levels*, capture the preference coefficients β at those discrete levels, and then interpolate for intermediate values. This allows the model to capture a wide variety of shapes with respect to the real-valued product characteristics **z** and price p. For example, the graph below shows a hypothetical case where the underlying relationship between v and a single product characteristic z is s-shaped. If we discretize z and obtain preference estimates at the discrete levels (shown as circles), we can interpolate the s-shaped curve. However, if we assume that v is a linear, quadratic, or log function of z, then we obtain a more restrictive estimate that does not capture all of the detail.



This technique of discretizing and interpolating may not be feasible using data from the market, since we may not be able to describe existing market products in terms of a small number of discrete levels of each characteristic. However, if we are collecting choice data using a designed survey, it is feasible and often desirable.

First, we divide each product characteristic **z** into discrete levels that span the relevant domain of characteristic values. If the product characteristics are indexed by ζ, we divide each characteristic z_ζ into levels indexed by $\omega = \{1, 2, 3, ..., \Omega_\zeta\}$. For example, characteristic $\zeta = 1$ is fuel economy, and if fuel economy z_1 ranges between say 10 mpg and 40 mpg, we might set levels at 10, 20, 30, and 40 mph, so that $\Omega_1 = 4$, and $\omega = \{1, 2, 3, 4\}$ refers to {10mpg, 20mpg, 30mpg, 40mpg} respectively.

Each product in the choice set must be coded with respect to these characteristic levels using *dummy variables*. Here we notate the dummy variables as $\delta_{j\zeta\omega}$, where $\delta_{j\zeta\omega} = 1$ if product characteristic ζ of product j is at level ω, and $\delta_{j\zeta\omega} = 0$ otherwise. We also include price in this set, with price indexed as $\zeta = 0$. Thus, any product j with product characteristics and price at the discrete levels can be coded as a set of 1's and 0's in $\delta_{j\zeta\omega}$ $\forall\zeta,\omega$. Assuming that preferences are linear in the discretized set (a main-effects model), we have

$$v_j = \sum_{\zeta} \sum_{\omega} \beta_{\zeta\omega} \delta_{j\zeta\omega} \ , \qquad \text{(A.20)}$$

where the coefficients $\beta_{\zeta\omega}$ are called *part-worths* because they describe the component of utility derived from characteristic ζ being at level ω. There may be cases where linearity of the characteristics cannot be assumed because of interaction effects, i.e., the shape of preferences for one characteristic may depend on the value of another characteristic. However, these are left as as advanced cases that are not addressed here.

Using the logit model, the probability of an individual choosing product j is then:

$$P_j = \frac{\exp(v_j)}{\sum_{j'} \exp(v_{j'})} = \frac{\exp\left(\sum_{\zeta}\sum_{\omega} \beta_{\zeta\omega}\delta_{j\zeta\omega}\right)}{\sum_{j'} \exp\left(\sum_{\zeta}\sum_{\omega} \beta_{\zeta\omega}\delta_{j\zeta\omega}\right)} \tag{A.21}$$

and the log likelihood that a model with part-worth coefficients $\beta_{\zeta\omega}$ will reproduce the observed data Φ_{ij}, where $\Phi_{ij} = 1$ if individual i chooses product j, and $\Phi_{ij} = 0$ otherwise, is

$$LL = \sum_j \sum_i \Phi_{ij} \ln P_j \tag{A.22}$$

as derived before. Given a set of observed choice data Φ_{ij} we can find the coefficients $\beta_{\zeta\omega}$ that maximize the log likelihood.

Example

In the vehicle example, we had

	A	B	C	D
p_j (\$1000s)	15	15	20	20
z_{j1} (mpg)	25	35	25	35
z_{j2} (sec)	6	8	8	6

where levels are defined as

ζ	symbol	level ω=1	level ω=2
0	p	\$15,000	\$20,000
1	z_1	25	35
2	z_2	6	8

The corresponding dummy variables $\delta_{j\zeta\omega}$ for these products are

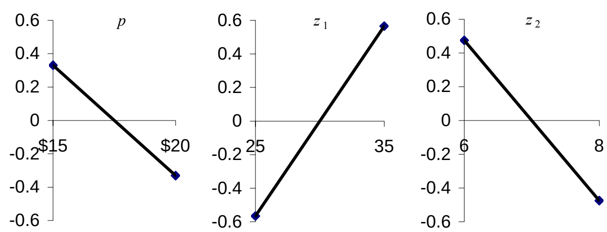
		j=A	j=B	j=C	j=D
ζ=0	ω=1	1	1	0	0
ζ=0	ω=2	0	0	1	1
ζ=1	ω=1	1	0	1	0
ζ=1	ω=2	0	1	0	1
ζ=2	ω=1	1	0	0	1
ζ=2	ω=2	0	1	1	0

As before, given this choice set suppose that 25 respondents choose product A, 30 choose product B, 5 choose product C, and 40 choose product D. If the log likelihood is maximized using Excel Solver, the resulting $\beta_{\zeta\omega}$ part-worths are:

$\beta_{\zeta\omega}$	ζ=0	ζ=1	ζ=2
ω=1	0.3304	-0.56544	0.47428
ω=2	-0.3304	0.56544	-0.47428

Actually, there are infinitely many sets of part worth coefficients that predict equivalent choice probabilities, and the results shown above are just one such set. This is because our model for v has extra degrees of freedom: i.e., there are more variables than equations in the system of equations. Any of the sets of betas that yield equivalent choice probabilities and log likelihood values are equivalent with respect to the choice model, and any can be used. If we wish to restrict the model to a single answer (this is called *model identification*), we can code the beta coefficients in terms of fewer variables ($1 + \Sigma_\zeta (\Omega_\zeta\text{-}1)$ variables are needed), or, equivalently, we can add extra constraints to restrict the solution to a particular set of beta values from the infinite set of equivalent values for easier interpretation. The solution shown above is the particular beta solution maximizing LL where the average β value of each characteristic ζ across all of its levels ω is zero.

The resulting beta values are plotted below for each characteristic and price. Each ζ is divided into only two levels, so we can use linear interpolation to estimate β values for intermediate levels, for example, a price of $18,000.



By including only two levels per ζ, the resulting interpolation shown in the graphs is linear with respect to the real-valued characteristics, and we have essentially assumed a linear relationship. The final interpolated relationship for intermediate values of v, using linear interpolation, is

$$\hat{v}_j = -0.132 p_j + 0.113 z_{1j} - 0.474 z_{2j} \tag{A.23}$$

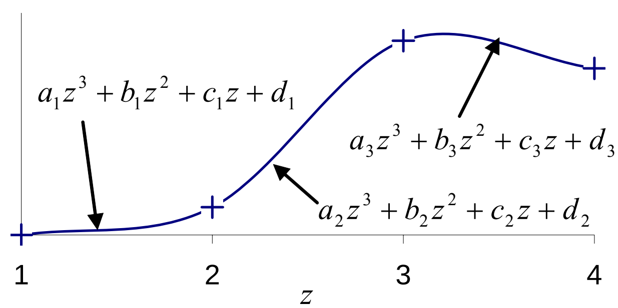
We see that the slope of the part worth for price (0.3304 – (-0.3304))/($20-$15) = 0.132 is the same value we obtained previously when we had assumed a linear functional form of p. The slopes of z_1 and z_2 are different because here we have only two levels, which implies a linear relationship, whereas the functional form assumed previously was inversely proportional to each. So, using only two levels for each ζ is not recommended unless the modeler is relatively certain that the relationship is nearly linear, or that a linear representation will suffice. Use of more than two levels allows more general spline interpolation, and can represent more complex relationships.

Interpolation of Part-Worth Coefficients Using Splines

In general, a spline can be fit through the part worth values $\beta_{\zeta\omega}$ of all levels ω in each ζ to interpolate intermediate values of ζ. It is possible to use many types of splines to interpolate the points; however, to facilitate optimization over the real-valued product characteristic values, it is desirable to interpolate using a spline function that is smooth and continuous over the domain. In particular, we will focus on natural cubic splines: a set of (Ω_ζ–1) cubic polynomials, each of which has a domain between two adjacent levels ω (one between $\omega = 1$ and $\omega = 2$, another between $\omega = 2$ and $\omega = 3$, etc), that:

- Match the value $\beta_{\zeta\omega}$ at each of the two domain endpoints ω,
- Match the first and second derivatives of the adjacent cubic polynomial at each domain endpoint
- Have second derivatives of zero at the extreme bounds of the spline: $\omega = 1$ and $\omega = \Omega_\zeta$.

An illustration is shown below with $\Omega_\zeta = 4$ four levels for hypothetical characteristic z:



It is possible to calculate the coefficients of the (Ω_ζ – 1) cubic polynomials in a spline for characteristic ζ given $\beta_{\zeta\omega}$ by solving a system of equations representing the three conditions; however, this detail is avoided here. Instead, software packages such as Excel or Matlab can be used to automatically calculate cubic splines given values for $\beta_{\zeta\omega}$. The cubic spline function for characteristic (or price) ζ that passes through the levels ω of $\beta_{\zeta\omega}$ will be notated as Ψ_ζ. The interpolated observable component of utility then involves the resulting spline function evaluated at the intermediate, real-valued product characteristics and price:

$$\hat{v}_j = \Psi_0\left(p_j\right) + \sum_{\zeta>0} \Psi_\zeta\left(z_{\zeta j}\right) \tag{A.24}$$

This interpolated value of v can then be used in the logit model to predict the choice probabilities of new products with intermediate product characteristic and price values.

Example

Suppose that we had included more levels in our earlier example

ζ	symbol	level ω=1	level ω=2	level ω=3
0	p	\$15,000	\$20,000	\$25,000
1	z_1	25	35	45
2	z_2	6	8	10

and the three separate choice sets below were provided to survey respondents, and their choices were recorded for each choice set.

Choice set		A	B	C	None
1	p_j (\$1000s)	15	20	25	-
	z_{j1} (mpg)	25	35	45	-
	z_{j2} (sec)	6	10	6	-
2	p_j (\$1000s)	15	20	25	-
	z_{j1} (mpg)	35	45	25	-
	z_{j2} (sec)	8	6	10	-
3	p_j (\$1000s)	15	20	25	-
	z_{j1} (mpg)	45	25	35	-
	z_{j2} (sec)	10	8	6	-

Suppose 100 people were given this survey and the number of people choosing each option in each set is given by:

Choice set	A	B	C	None	Total
1	45	5	45	5	100
2	40	55	0	5	100
3	30	25	30	15	100

Given these data, the partworths (centered around zero for each characteristic, as before) can be calculated as

ζ	symbol	level ω=1	level ω=2	level ω=3
0	p	0.64	-0.03	-0.61
1	z_1	-0.67	-0.07	0.74
2	z_2	0.74	0.57	-1.32

with the no-choice option utility value of -1.829. Interpolating a spline through the levels of price and each characteristic would enable estimate of the part worth of an intermediate level, and predictions of choice probabilities could be calculated with the logit model, as before.

These basic models are applied in Chapters 4-7 in the development of models to predict choice as a function of the prices and characteristics of available product alternatives. Given such a predictive model, the prices and characteristics of the alternatives are varied using optimization algorithms to find values that yield the most desirable predictions.