Proceedings of the ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2011
August 29-31, 2011, Washington, DC, USA

DETC2011-48409

ROBUST DESIGN FOR PROFIT MAXIMIZATION UNDER UNCERTAINTY OF CONSUMER CHOICE MODEL PARAMETERS USING THE DELTA METHOD

Camilo B. Resende
Graduate Student
Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA, United States

C. Grace Heckmann
Graduate Student
Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA, United States

Jeremy J. Michalek
Associate Professor
Mechanical Engineering
Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA, United States
Corresponding Author

ABSTRACT

In new product design, risk averse firms must consider downside risk in addition to expected profitability, since some designs are associated with greater market uncertainty than others. We propose an approach to robust optimal product design for profit maximization by introducing an α-profit metric to manage expected profitability vs. downside risk due to uncertainty in market share predictions. Our goal is to maximize profit at a firm-specified level of risk tolerance. Specifically, we find the design that maximizes the α-profit: the value that the firm has a (1-α) chance of exceeding, given the distribution of possible outcomes. The parameter α∈[0,1] is set by the firm to reflect sensitivity to downside risk (or upside gain), and parametric study of α reveals the sensitivity of optimal design choices to firm risk preference. We account here only for uncertainty of choice model parameter estimates due to finite data sampling when the choice model is assumed to be correctly specified (no misspecification error). We apply the delta method to estimate the mapping from uncertainty in discrete choice model parameters to uncertainty of profit outcomes and identify the estimated α-profit as a closed form function of design decision variables. This process is described for the multinomial logit model, and a case study demonstrates implementation of the method to find the optimal design characteristics of a midsize consumer automobile.

Keywords: Design for market systems, delta method, logit, design optimization, robust design, design under uncertainty, discrete choice model

NOMENCLATURE

c Variable cost
C Fixed cost
F_π Cumulative distribution function of profit estimate
g Mapping function for delta method
j Product index
J Number of products
m Market size
n Number of attributes per product
p_j Price of product j
s_j Point estimate market share for product j
$\hat{s}_j$ Random variable market share estimate for product j
s_j^α Market share of product j at risk level α
$\mathbf{v}$ Vector of point estimates of utility for all products
v_j Point estimate observable utility for product j
$\hat{v}_j$ Random variable observable utility estimate for product j
$\mathbf{x}_j$ Vector of attributes for product j
$\mathbf{x}_j^{\alpha*}$ Optimal product attributes at level α
$\mathbf{X}$ Matrix of attributes for all products
t_j Technology implementation design variable for product j in optimization model
w_j Engine size constant in optimization model
α Profit risk tolerance parameter
$\boldsymbol{\beta}$ Vector of choice model parameter point estimates
$\hat{\boldsymbol{\beta}}$ Random vector of choice model parameter estimates
$\bar{\boldsymbol{\beta}}$ Mean of $\hat{\boldsymbol{\beta}}$ distribution
$\boldsymbol{\Sigma}$ Covariance matrix of $\hat{\boldsymbol{\beta}}$ distribution
π_j Point estimate of profit for product j
$\hat{\pi}_j$ Random variable profit estimate for product j
π_j^α Profit of product j at a level α
Φ Standard normal cumulative distribution function

1. INTRODUCTION

Over the last three decades, a significant portion of the new product development (NPD) literature has been dedicated to the integration of engineering design and marketing processes for differentiated markets. Simple models to determine the most profitable characteristics of a single new product [1-2] have progressed to account for issues such as product-line design and preference heterogeneity [3-7], competitor reactions [8-10], cost structure [11-12], distribution channels [9,13-16], choice-set-dependent preferences [17], and coordination with constrained engineering design decisions [18-26].

As Hsu and Wilcox [27] argue, the trend towards estimating marketing models at lower levels of aggregation that are more structural[1] in consumer behavior representation has led to models with many parameters and consequently greater uncertainty of those parameters. However, despite the advances in NPD methods, the research has not given much consideration to the intrinsic parameter uncertainty of the demand models. Demand uncertainty directly affects the risk of introducing a new product into the market, and firms evaluate potential projects not only in terms of expected return, but also in terms of risk.

The purpose of this work is threefold. First, we define a robust *a*-profit metric and propose a general framework to incorporate demand uncertainty arising from choice model parameter estimation into the design decision process such that it accounts for varying levels of loss tolerance. Second, we apply the delta method to approximate the *a*-profit function in closed form for multinomial logit (MNL) demand models to be used efficiently in numerical optimization routines. Finally, we show how ignoring demand uncertainty can lead to suboptimal decisions for risk averse firms.

We do not intend to consider all the various sources of demand model uncertainty [28], and several questions will remain open. In particular, we assume the discrete choice model is correctly specified and ignore uncertainty due to model misspecification, and we assume that the model parameters do not change over time or from the context in which the data were collected to the context in which predictions will be made. Nevertheless, the proposed methodology can be useful, and it serves as a first step in addressing design for profit maximization under demand model uncertainty. Design decision makers will be able to use the popular multinomial logit demand models to develop new products based on their expected profit and also to account for the inherent uncertainty present in any statistical estimation process.

This paper begins by discussing the relevant literature on product design and pricing under uncertainty in Section 2. Section 3 describes the proposed methodology for finding optimal designs for varying levels of tolerated product profit uncertainty and applies it to multinomial logit demand models. Section 4 presents a case study using the multinomial logit demand model to determine the optimal attributes of a midsize consumer vehicle. Section 5 discusses conclusions, limitations, and future work.

[1]"Econometric models that are based explicitly on the consumer's maximization problem and whose parameters are parameters of the consumers' utility functions or of their constraints are referred to as structural models." [50]

2. LITERATURE REVIEW: PRODUCT DESIGN AND PRICING UNDER DEMAND UNCERTAINTY

Demand uncertainty is caused by several factors such as preference dynamics [29], demand model misspecification [30-31], choice context [32-33], response variability [34-35], and sampling errors associated with the estimation procedure [36]. As a result, several researchers have considered the impact of demand uncertainty on optimal pricing strategies [29-30,37-38]. However, in contrast to prices, design decisions are difficult to change post hoc, especially in durable-goods markets. Products with high start-up capital costs can have virtually unchangeable characteristics, and producers are incentivized to consider demand uncertainty during the initial stages of the design process (e.g., car manufacturers invest a significant portion of capital up front in production equipment, and changing a characteristic such as the footprint of a car leads to very high costs).

A few papers [21,36,39] have addressed product demand uncertainty resulting from variation in engineering design model parameters (e.g. due to manufacturing variability or usage conditions), and two of them also account for uncertainty in the marketing model parameters. Luo et al. [36] and Besharati et al. [21], both address demand uncertainty using discrete intervals.

Luo et al. [36] use the parameter covariance matrix of part-worth utility point estimates to obtain 95% confidence intervals around the point estimates from the design parameter best- and worst- case scenarios for a set of product alternatives under consideration. The greatest utility under the best-case scenario and lowest utility under the worst-case scenario within the confidence interval are compared to the similarly constructed estimates of utility for competitor products. The highest own-utility is compared to the sum of the lowest competitor-utilities and vice-versa to construct interval estimates of market shares (these no longer represent statistical confidence intervals for market share). They then use pair-wise comparisons to eliminate dominated alternatives (defined as alternatives that have a best-case market share worse than another's worst-case market share, perform worse on worst-case performance, and have higher performance variability). All non-dominated designs are then considered for prototyping and further subjective evaluation.

Besharati et al. [21] use a framework similar to Luo et al. [36], but they change the optimization criteria arguing that looking for the best performance on the worst case condition might be too conservative. Alternatively, they replace the design objectives of worst-case performance and performance variability with multi-objective optimization of nominal performance characteristics. The marketing model is also treated as a multi-objective optimization problem of maximizing nominal market share and minimizing the market share variance (both positive and negative) resulting from uncertainty in both engineering design parameters and part-worth utility estimates. Finally, they develop a ranking system for pair-wise comparison of designs on the design and marketing criteria.

Hsu and Wilcox [27] use the estimation error associated with the parameter estimates to find the stochastic market share prediction in a multinomial logit framework. They use a simulation-based approach for approximating the exact distribution in an efficient way.

Table 1 compares the above papers that consider the uncertainty in demand model parameters as a source of demand uncertainty and positions our contribution against this prior work. We address variance of profit estimates but do not seek to minimize it as a means to improve robustness because profit uncertainty is harmful to a firm only in the negative tail – i.e. when product demand is less than expected – and we avoid penalizing uncertainty that could lead to higher than expected profits.

We apply an α-profit metric in conjunction with discrete choice models as a means to incorporate firm risk tolerance into the new product design optimization process. This allows us to develop a framework to find optimal product characteristics and price in a continuous domain, instead of requiring a discrete set of product alternatives; and in contrast to Luo et al. [36] and Besharati et al. [21], we can treat demand uncertainty as a continuous probability distribution instead of representing it as an interval. We use the delta method to derive a closed-form approximation for the market share distribution, since a simulation-based approach such as the one used by Hsu and Wilcox [27], though efficient for estimating the stochastic distribution of a single design, would be computationally expensive and noisy when used as an intermediate function in a numerical optimization loop. Our framework focuses on demand models derived from random utility theory, particularly multinomial logit models (MNL) [40].

The α-profit methodology can be extended to multinomial probit (MNP) [41], mixed logit (MIXL) [42], and the recent generalized logit (G-MNL) [43] models; however, any functional forms that require numerical simulation to compute may be computationally burdensome and introduce potential numerical issues when embedded within an optimization loop.

3. THE PROPOSED METHODOLOGY

We want to find the characteristics of a new product in order to maximize a firm's profit; however, the uncertainty present in the demand model parameter estimates will result in uncertainty about the predicted profit, which we model as a distribution of potential profit outcomes for each design alternative. (A similar framework can also be used for maximizing alternative objective functions, such as market share.)

3.1. General mathematical formulation

Our goal is to find the design whose predicted profit distribution maximizes the α-profit: the value below which less than α of the cumulative profit distribution falls. The parameter α is set by the firm to reflect sensitivity to downside risk (or upside gain), and parametric study of α reveals the sensitivity of optimal design choices to firm risk preference. Formally, we define the α-profit, $\pi_j^\alpha(\hat{\boldsymbol{\beta}},\mathbf{X})$, as the value of the profit distribution at level $\alpha\in[0,1]$ for product $j\in\{1,2,\ldots J\}$ given the vector of random variables $\hat{\boldsymbol{\beta}} \sim N(\bar{\boldsymbol{\beta}},\boldsymbol{\Sigma})$ that define the parameter estimates of the choice model and the values of the n attributes (including price) for each of the J products available in the market $\mathbf{X} = [\mathbf{x}_1,\ldots,\mathbf{x}_j,\ldots,\mathbf{x}_J] \in \mathbb{R}^{Jn}$. Specifically, if $\hat{\pi}_j$ is a random variable with cumulative distribution function $\mathsf{F}_{\hat{\pi}|\hat{\boldsymbol{\beta}},\mathbf{X}}(\pi)$ representing the distribution of profit outcomes conditional on $\hat{\boldsymbol{\beta}}$ and $\mathbf{X}$, then $\pi_j^\alpha(\hat{\boldsymbol{\beta}},\mathbf{X})$ is the maximum value of π_j for which $\mathsf{F}_{\hat{\pi}|\hat{\boldsymbol{\beta}},\mathbf{X}}(\pi) \le \alpha$, i.e.: for which $\Pr(\hat{\pi}_j < \pi_j) \le \alpha$ (see Figure 1). If $\mathsf{F}_{\hat{\pi}|\hat{\boldsymbol{\beta}},\mathbf{X}}(\pi)$ is continuous and invertible, then $\pi_j^\alpha(\hat{\boldsymbol{\beta}},\mathbf{X}) = \mathsf{F}^{-1}_{\hat{\pi}|\hat{\boldsymbol{\beta}},\mathbf{X}}(\alpha)$.

TABLE 1 – PAPERS THAT CONSIDER CHOICE MODEL PARAMETER UNCERTAINTY AS A SOURCE OF DEMAND UNCERTAINTY

Ref.	Treats demand uncertainty as:	Design attributes	Design objective(s)
Hsu and Wilcox [27]	Probability distribution of market share obtained by simulation	NA	NA
Luo et al. [36]	Interval estimates of market shares obtained using 95% confidence levels for the utility function	Discrete	Maximize nominal market share Minimize performance variance Maximize worst-case performance
Besharati et al. [21]	Interval estimates of market shares obtained using 95% confidence levels for the utility function	Discrete	Maximize nominal share Minimize share variance Maximize nominal performance
This paper	Probability distribution of market share estimated by delta method	Continuous	Maximize profit at specified downside risk tolerance level

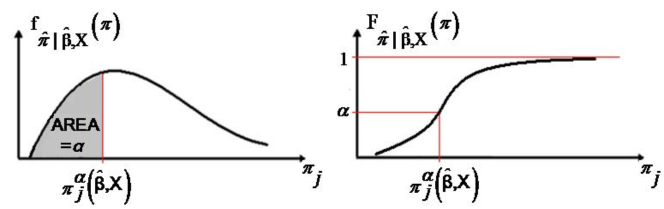

FIGURE 1 – α-PROFIT SHOWN FOR (A) PROBABILITY DENSITY FUNCTION OF PROFIT AND (B) CUMULATIVE DISTRIBUTION FUNCTION OF PROFIT

Our objective is to find the product attributes and price that maximize the robust profit given the α level that reflects firm sensitivity to downside risk. That is, we seek the robust optimal new product characteristics $\mathbf{x}_j^{\alpha*}$ at level α, where $\mathbf{x}_j^{\alpha*} = \text{argmax}_{\mathbf{x}_j}(\pi_j^\alpha(\hat{\boldsymbol{\beta}},\mathbf{X}))$; i.e. $\mathbf{x}_j^{\alpha*}$ is the design that maximizes the value of profit which the model predicts a (1-α) chance of exceeding. For illustration, Figure 2 shows the probability density function of profit for two alternative designs. Design 1 is preferred over Design 2 when optimizing for the expected value of profit. However, Design 1 has more downside risk, and a risk averse firm optimizing for the α-profit with small α would prefer Design 2.

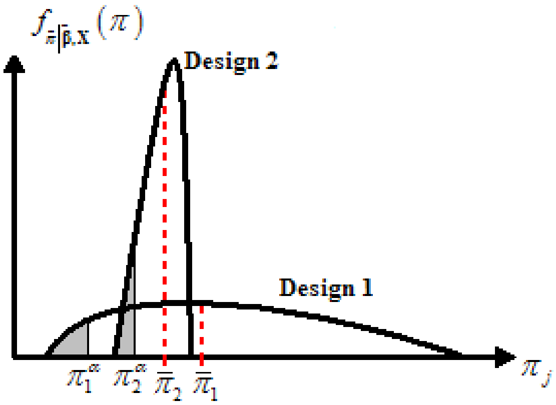

FIGURE 2 – EXPECTED PROFIT VS. DOWNSIDE RISK: THE EXPECTED PROFIT FOR DESIGN 1 IS HIGHER THAN THE EXPECTED PROFIT FOR DESIGN 2 ($\bar{\pi}_1 > \bar{\pi}_2$); HOWEVER, DESIGN 2 HAS A HIGHER PROFIT AT THE α-LEVEL THAN DESIGN 1 ($\pi_2^\alpha > \pi_1^\alpha$)

Defining for product j the random variable describing the distribution of market share outcomes, $\hat{s}_j$; market share at level α, s_j^α; price, p_j; variable cost, $c_j = f_{\mathrm{VC}}(\mathbf{x}_j)$; fixed cost, C_j, and total market size, m; we have $\hat{\pi}_j = m(p_j - c_j)\hat{s}_j - C_j$ and $\pi_j^\alpha = m(p_j - c_j)s_j^\alpha - C_j$ Assuming that there is no uncertainty on product price and costs and that $p_j > c_j$:

$$\Pr(\hat{\pi}_j < \pi_j^\alpha) = \Pr\left(m(p_j - c_j)\hat{s}_j < m(p_j - c_j)s_j^\alpha\right) = \Pr(\hat{s}_j < s_j^\alpha) \quad (1)$$

Therefore:

$$\Pr(\hat{\pi}_j < \pi_j^\alpha) \le \alpha \Leftrightarrow \Pr(\hat{s}_j < s_j^\alpha) \le \alpha \quad (2)$$

In the following sections, we will show how to find the robust optimal new product attributes, as defined in this section, for the multinomial logit (MNL) demand model, given uncertainty in the estimated parameters.

3.2. Application to Multinomial Logit Demand Model

We apply the proposed methodology to the MNL model for several reasons: (1) it is the most widely used discrete choice model, especially due to its closed form choice probabilities and interpretability [44]; (2) several discrete choice models evolved from MNL, such as MIXL and G-MNL, and a better understanding of how uncertainty affects NPD under MNL models can lead to general conclusions and intuition about the effects of uncertainty under different models; and (3) even though MNL was introduced more than thirty years ago by McFadden [40], it is still widely used in the NPD literature [14,45-46]. For the purposes of this paper, we assume that the model is correct and that the uncertainty arises from the parameter estimation and not model misspecification.

In a multinomial logit model, given some competitive set of J products, the predicted market share s_j for product j can be computed as

$$s_j(\mathbf{v}) = \frac{e^{v_j}}{\sum_{k=1}^{J} e^{v_k}} \quad (3)$$

where $\mathbf{v} = (v_1, \dots, v_J)$ is the vector of observable utility point estimates of the respective products.

The utility function is often specified to be linear in parameters: $v_k = \beta^{\mathrm{T}}\mathbf{x}_k$, resulting in predicted market share $s_j(\mathbf{X})$ for some product $j \in \{1, \dots, N\}$:

$$s_j(\mathbf{X}) = \frac{e^{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j}}{\sum_{k=1}^{J} e^{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_k}} \quad (4)$$

Ignoring constant fixed costs, in a multinomial logit demand model the predicted profit π_j can be computed as:

$$\pi_j(\mathbf{X}) = m(p_j - c_j)s_j(\mathbf{X}) = m(p_j - c_j)\frac{e^{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j}}{\sum_{k=1}^{J} e^{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_k}} \quad (5)$$

The classical practice is to use maximum likelihood methods to estimate the parameters $\boldsymbol{\beta}$ in multinomial logit models [27]. Train [44] notes that the estimates are easily obtained since the log-likelihood function is concave for linear utility specifications, and Wooldridge [47] proves that the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ is asymptotically normally distributed with distribution $\hat{\boldsymbol{\beta}} \sim N(\bar{\boldsymbol{\beta}}, \boldsymbol{\Sigma})$, where $\bar{\boldsymbol{\beta}}$ is the vector of means and $\boldsymbol{\Sigma}$ is the covariance matrix, implying that $\hat{v}_j \sim N(\bar{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}_j, \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{x}_j)$.

The exact distribution of $\hat{s}_j$ is unknown, but the delta method enables analytic approximation of a transformed distribution using a linear approximation of the mapping function. This frees us from the computational burden of simulating a market share distribution for each choice of product attributes in the optimization loop, as would be required by the method in Hsu and Wilcox [27]. The delta method states that any function of a normally distributed random variable (in this case the estimated parameters) converges asymptotically to a normal distribution ([47], see Appendix B for proof). The delta method relies on a Taylor series expansion of the mapping function g. If the function of the expected value of the parameters is $\mathrm{g}(\bar{\boldsymbol{\beta}})$, then $\mathrm{g}(\hat{\boldsymbol{\beta}}) \cong \mathrm{g}(\bar{\boldsymbol{\beta}}) + \nabla\mathrm{g}(\bar{\boldsymbol{\beta}}) \cdot (\hat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}})$. The mean and variance of $\mathrm{g}(\hat{\boldsymbol{\beta}})$ can be calculated as:

$$\mathrm{E}[\mathrm{g}(\hat{\boldsymbol{\beta}})] \cong \mathrm{E}[\mathrm{g}(\bar{\boldsymbol{\beta}}) + \nabla\mathrm{g}(\bar{\boldsymbol{\beta}}) \cdot (\hat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}})] = \mathrm{g}(\bar{\boldsymbol{\beta}}) \quad (6)$$

$$\mathrm{Var}[\mathrm{g}(\hat{\boldsymbol{\beta}})] = \mathrm{E}\left[\left(\mathrm{g}(\hat{\boldsymbol{\beta}}) - \mathrm{g}(\bar{\boldsymbol{\beta}})\right)^2\right] \cong \mathrm{E}\left[\left(\nabla(\bar{\boldsymbol{\beta}}) \cdot (\hat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}})\right)^2\right] = \mathrm{E}\left[\left(\nabla\mathrm{g}(\bar{\boldsymbol{\beta}})\right)^2 \cdot (\hat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}})^2\right] = \nabla\mathrm{g}(\bar{\boldsymbol{\beta}})^2 \cdot \mathrm{Var}(\hat{\boldsymbol{\beta}}) \quad (7)$$

As with any linear function approximation to a nonlinear function, it may lead to significant distortion of the function outside the neighborhood of $\mathrm{g}(\bar{\boldsymbol{\beta}})$.

The quantity of interest $\hat{s}_j$ is itself a function of $\hat{\boldsymbol{\beta}}$, but $s_j \in [0,1]$, which does not match the domain of the normal distribution. Instead, we select the intermediate function

$g(\boldsymbol{\beta}) = \ln(1/s_j - 1) \in (-\infty, +\infty)$ so that $g(\boldsymbol{\beta})$ has the same domain of a normal distribution, and so that in the case of a monopolistic single product firm with an outside good, the approximation leads to the exact distribution of the profit.

$$g(\boldsymbol{\beta}) = \ln\left(\frac{1}{s_j} - 1\right) = \ln\left(\frac{\sum_{k=1}^{J} e^{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_k}}{e^{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_j}} - 1\right) = \ln\left(\sum_{k \in J \setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}_j)}\right) \tag{8}$$

By the delta method, we know that

$$g(\hat{\boldsymbol{\beta}}) \overset{a}{\sim} N\left(g(\bar{\boldsymbol{\beta}}), \left(\frac{\partial g}{\partial \boldsymbol{\beta}^{\mathrm{T}}} \boldsymbol{\Sigma} \frac{\partial g}{\partial \boldsymbol{\beta}}\right)\Bigg|_{\bar{\boldsymbol{\beta}}}\right) \tag{9}$$

Since

$$\frac{\partial g}{\partial \boldsymbol{\beta}} = \frac{\sum_{k \in J \setminus j} (\mathbf{x}_k - \mathbf{x}_j) e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}_j)}}{\sum_{k \in J \setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}_j)}} \tag{10}$$

(see Appendix A for details) we can approximate the variance of g for any given $\mathbf{X}$. Because

$$\hat{s}_j < s_j^{\alpha} \iff \left(\frac{1}{\hat{s}_j} - 1\right) > \left(\frac{1}{s_j^{\alpha}} - 1\right) \iff g(\hat{\boldsymbol{\beta}}) > \ln\left(\frac{1}{s_j^{\alpha}} - 1\right) \tag{11}$$

we can calculate

$$\Pr\left(\hat{s}_j < s_j^{\alpha}\right) = \alpha \iff \Pr\left(g(\hat{\boldsymbol{\beta}}) > \ln\left(\frac{1}{s_j^{\alpha}} - 1\right)\right) = \alpha \tag{12}$$

Normalizing the right hand equation

$$\Rightarrow \Pr\left(\frac{g(\bar{\boldsymbol{\beta}}) - g(\hat{\boldsymbol{\beta}})}{\left(\frac{\partial g}{\partial \boldsymbol{\beta}^{T}} \boldsymbol{\Sigma} \frac{\partial g}{\partial \boldsymbol{\beta}}\right)^{\frac{1}{2}}\Bigg|_{\bar{\boldsymbol{\beta}}}} < \frac{g(\bar{\boldsymbol{\beta}}) - \ln\left(\frac{1}{s_j^{\alpha}} - 1\right)}{\left(\frac{\partial g}{\partial \boldsymbol{\beta}^{T}} \boldsymbol{\Sigma} \frac{\partial g}{\partial \boldsymbol{\beta}}\right)^{\frac{1}{2}}\Bigg|_{\bar{\boldsymbol{\beta}}}}\right) = \alpha \tag{13}$$

Since:

$$\left(\frac{g(\bar{\boldsymbol{\beta}}) - g(\hat{\boldsymbol{\beta}})}{\left(\frac{\partial g}{\partial \boldsymbol{\beta}^{T}} \boldsymbol{\Sigma} \frac{\partial g}{\partial \boldsymbol{\beta}}\right)^{\frac{1}{2}}\Bigg|_{\bar{\boldsymbol{\beta}}}}\right) \overset{a}{\sim} N(0,1) \tag{14}$$

The probability expression is the cumulative distribution of a standard normal, thus

$$\Pr\left(\hat{s}_j < s_j^{\alpha}\right) = \alpha \iff \Phi\left(\frac{g(\bar{\boldsymbol{\beta}}) - \ln\left(\frac{1}{s_j^{\alpha}} - 1\right)}{\left(\frac{\partial g}{\partial \boldsymbol{\beta}^{T}} \boldsymbol{\Sigma} \frac{\partial g}{\partial \boldsymbol{\beta}}\right)^{\frac{1}{2}}\Bigg|_{\bar{\boldsymbol{\beta}}}}\right) = \alpha \tag{15}$$

where Φ is the cumulative distribution function of the standard normal distribution. Solving for s_j^{α}

$$s_j^{\alpha} = \left(1 + \exp\left(g(\bar{\boldsymbol{\beta}}) - \Phi^{-1}(\alpha)\left(\frac{\partial g}{\partial \boldsymbol{\beta}^{T}} \boldsymbol{\Sigma} \frac{\partial g}{\partial \boldsymbol{\beta}}\right)^{\frac{1}{2}}\Bigg|_{\bar{\boldsymbol{\beta}}}\right)\right)^{-1} \tag{16}$$

Equation (16) enables a modeler to compute the estimated market share at the α risk level as a closed form deterministic function of the decision variables using only the mean $\bar{\boldsymbol{\beta}}$ and covariance matrix $\boldsymbol{\Sigma}$ defining the choice model parameter estimates. Both $\bar{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}$ are available from standard estimation procedures. The α-profit can then be computed as $\pi_j^{\alpha} = m(p_j - c_j)s_j^{\alpha} - C_j$. In the special case of a monopolistic single product firm, this framework leads to the exact distribution of the profit, since

$$\begin{aligned} g(\hat{\boldsymbol{\beta}}) &= \ln\left(e^{-\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}_j}\right) \\ &= -\hat{\boldsymbol{\beta}}^{T}\mathbf{x}_j \sim N\left(-\bar{\boldsymbol{\beta}}^{T}\mathbf{x}_j, \left(\frac{\partial g}{\partial \boldsymbol{\beta}^{T}} \boldsymbol{\Sigma} \frac{\partial g}{\partial \boldsymbol{\beta}}\right)\Bigg|_{\bar{\boldsymbol{\beta}}}\right) \end{aligned} \tag{17}$$

Figure 3 illustrates the mapping for a model with a single parameter showing the normal distribution of the estimated model coefficient, $\hat{\boldsymbol{\beta}}$, the resulting distribution of $g(\hat{\boldsymbol{\beta}})$ and its normally distributed approximation via the delta method, and the resulting distribution of $\hat{s}_j$ and its (non-normal) approximation via the delta method.

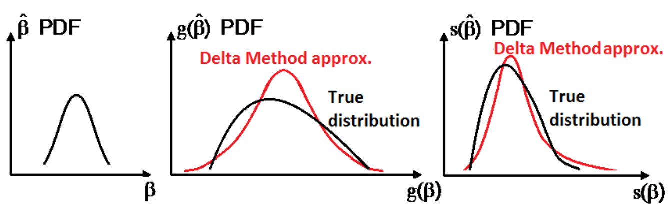

FIGURE 3 – ILLUSTRATION OF PROBABILITY DISTRIBUTION FUNCTIONS AND THEIR APPROXIMATIONS USING THE DELTA METHOD

As a result of the delta method formulation, the distribution of $\hat{s}_j$ depends on the variance of $g(\hat{\boldsymbol{\beta}})$, which depends on $\partial g/\partial \boldsymbol{\beta}$. Because $\partial g/\partial \boldsymbol{\beta}$ is proportional to $\sum_{k\in J\setminus j}(\mathbf{x}_k - \mathbf{x}_j)e^{\boldsymbol{\beta}^T(\mathbf{x}_k-\mathbf{x}_j)}$, the distribution of $\hat{s}_j$ is influenced by the differentiation of the new product attributes from each of the existing product attributes. All else being equal, greater differentiation implies higher uncertainty in market share predictions.

4. CASE STUDY

In this section, we examine the application of this method to the optimal design of a midsize vehicle for an automotive manufacturer. It is assumed that the manufacturer is operating as a single-product firm, and that other manufacturers do not redesign their products in response to the introduction of the new vehicle. The application to manufacturers that offer diversified product lines or operate in dynamic market conditions is left as future work.

4.1. Demand side

We estimate a logit model to describe vehicle choice using the fuel economy, price, and sales data from Ward's Automotive Index [48] for 44 midsize vehicle models (See Appendix C for vehicle details). We assume for simplicity that each of vehicles purchased in the data represents a consumer who considered the 44 midsize vehicle alternatives and purchased the identified vehicle at its manufacturer suggested retail price (MSRP). This data set has over 3 million midsize purchases – enough to identify logit model parameters with high certainty. To better illustrate the approach, we divide the total number of observations by 30,000 and round the resulting sales of each alternative to the nearest integer to simulate the uncertainty associated with a smaller data set of ~100 observations.

Two explanatory variables are considered: price (unit: \$10,000), and operating cost (unit: cents per mile), and we assume a linear utility function. If this model is misspecified, there will be additional uncertainty associated with model prediction, but we assume here correct specification and focus on uncertainty of parameter estimates due to missing data. Using the maximum-likelihood method for coefficient estimation, we obtain the results in Table 2.

TABLE 2 – MULTINOMIAL LOGIT MODEL PARAMETER COEFFICIENTS

	$n \approx 100$			$n \approx 1000$		
Observed variable	**Coef.**	**Std. Error**	**t-stat**	**Coef.**	**Std. Error**	**t-stat**
Price, \$10,000	-0. 354	0.126	-2.81	-0. 342	0.039	-8.68
Op. cost 0.01\$/mile	-0.114	0.063	-1.81	-0.119	0.020	-6.00

The coefficients suggest that consumers in the midsize segment value lower-priced vehicles with lower operating costs, as expected. The inverse of the Hessian, the information matrix, obtained from the maximum-likelihood optimization problem is shown in Table 3. In the case of maximum-likelihood estimators, the information matrix is also the variance-covariance matrix of the estimators, $\boldsymbol{\Sigma}$ (see Appendix D for proof).

TABLE 3 – VARIANCE-COVARIANCE MATRIX OF LOGIT COEFFICIENTS (INFORMATION MATRIX) FOR N=100

	Price (\$10,000)	Operating cost (0.01\$/mile)
Price (\$10,000)	0.0158	-0.0040
Operating cost (0.01\$/mile)	-0.0040	0.0040

4.2. Supply side

Following Shiau et al. [49], a midsize vehicle j with a gasoline engine is represented by an engine scaling variable w_j, a technology implementation variable t_j, and price p_j. Here w_j represents the power of the engine and t_j represents the level of implementation of fuel-saving technologies (e.g. low-friction lubricants or electric power steering). We set $w_j = 0.8$ for simplicity, reducing the problem to two decision variables. The mapping from design variable t_j to the observed product attribute of operating cost is given by equation 18. For this study, we assume a gasoline price of p_G = \$2.85, the average daily high price during the year the data were collected (2007).

$$\text{Operating cost (\$/mile)} = x_j = \frac{(1-t_j)\,p_G}{32.1760} \quad \text{where } p_G = \$2.85 \tag{18}$$

The marginal cost to the manufacturer per vehicle c_j is calculated as the summation of the vehicle base cost c_B engine cost c_E and fuel-saving technology cost c_T such that the production cost of a single vehicle is defined as $c_j=c_B+c_E+c_{Tj}$. We assume c_B is fixed at \$7,836, c_E is fixed at \$1,131.50, and c_{Tj} is given by equation 19:

$$c_{Tj}=b_3t_j^2+b_4t_j \quad \text{where } b_3 = \$85,936 \text{ and } b_4 = -\$2,177 \tag{19}$$

For more detail on the estimation of the operating cost and vehicle manufacturing cost models see Shiau et al. [49].

4.3. Optimization Results

The new vehicle is optimized according to the following formulation:

maximize $\pi_j^\alpha = m(p_j - c_j)s_j^\alpha - C_j$
with respect to $t_j,\ p_j$
subject to $0 \le t_j \le 1$; $10{,}000 \le p_j \le 90{,}000$
where

$$s_j^\alpha = \left(1+\exp\left(g(\bar{\boldsymbol{\beta}}) - \Phi^{-1}(\alpha)\left(\frac{\partial g}{\partial \boldsymbol{\beta}^T}\boldsymbol{\Sigma}\frac{\partial g}{\partial \boldsymbol{\beta}}\right)^{\frac{1}{2}}\Bigg|_{\bar{\boldsymbol{\beta}}}\right)\right)^{-1}$$

$$g(\boldsymbol{\beta}) = \ln\left(\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^T(\mathbf{x}_k-\mathbf{x}_j)}\right)$$

$$c_j = c_B + c_E + b_3t_j^2 + b_4t_j$$

Matlab's *fmincon* function was used to solve the problem, and the results are found in Table 4. Because optimization results are independent of the constants for fixed cost C and market size m we report the profit factor, defined as $s_j(p_j-c_j)$, which represents profit for a market size of one and fixed cost of zero. Profit for other values of these constants can be computed from the profit factor post hoc.

TABLE 4 - OPTIMAL PRODUCT CHARACTERISTICS

α (%)	Technology t_j	Price ($)	Op. cost (cents/mi.)	Var. Cost ($)	Market share at α level (%)	Profit factor at α level	Exp. market share (%)	Exp. Profit factor
10	0.05	30,298	8.37	9,105	1.91	406	2.83	599
15	0.07	31,516	8.21	9,271	1.93	430	2.75	613
20	0.09	32,652	8.06	9,464	1.95	452	2.69	624
25	0.10	33,768	7.93	9,683	1.96	473	2.63	633
30	0.12	34,903	7.80	9,930	1.98	494	2.55	636
35	0.13	36,086	7.67	10,211	1.99	516	2.48	642
40	0.15	37,343	7.55	10,532	2.01	539	2.43	650
45	0.16	38,704	7.41	10,903	2.03	563	2.35	653
50	0.18	40,206	7.27	11,336	2.04	590	2.28	659
55	0.20	41,896	7.12	11,852	2.07	621	2.20	661
60	0.22	43,839	6.95	12,478	2.09	655	2.11	661
65	0.24	46,132	6.76	13,256	2.12	696	2.00	657
70	0.26	48,928	6.55	14,256	2.15	746	1.88	652
75	0.29	52,486	6.28	15,600	2.19	809	1.73	637
80	0.33	57,293	5.95	17,525	2.25	895	1.58	629
85	0.38	64,420	5.49	20,577	2.34	1,024	1.37	602
90	0.46	76,876	4.76	26,361	2.48	1,254	1.07	543

Table 4 reveals that, for this case, as the firm's risk aversion increases (smaller α) it implements less fuel saving technology, resulting in lower production cost and higher operating cost, while pricing the vehicle lower and sacrificing some expected profit for reduced downside risk of profit at the α-level. As the firm becomes risk seeking (larger α) it implements more fuel-saving technology, resulting in higher production cost and lower operating cost, while pricing the vehicle substantially higher and sacrificing some expected profit for the small chance of high realized profits. An intermediate α-level is associated with highest expected profit, representing risk neutrality.

For comparison purposes, the competitor vehicle design statistics are summarized in Table 5, with details in Appendix C. The optimal price for the new vehicle at each α-level is greater than the competitor averages but within the range of competitor prices with the exception of high-risk cases α=85% and 90%. The operating cost for the new vehicle at all α-levels is lower than the averages for the competitor vehicles but within the range of competitor values with the exception of high-risk case α=90%. For the more risk averse α-levels, the optimal solution is closer to the competitor averages and within the range of competitor prices and operating costs. At the risk-seeking end of the spectrum, the new vehicle diverges from the average competitor designs, and at high risk-tolerance levels is even out of the range of the market-tested competitor designs.

TABLE 5 – COMPETITOR VEHICLE STATISTICS

	Price ($)	Fuel Economy (mi./gal.)	Op. cost (cents/mi.)
Minimum	10,415	21	5.21
Maximum	61,715	55	13.66
Sales weighted avg.	22,494	28	10.58

Figure 4 shows the cumulative profit distribution plots for $\mathbf{x}_j^{\alpha*}$ at α=25% and 75% and the optimal solution maximizing the value of profit at the expected value of $\boldsymbol{\beta}$. The optimal design at α=25% has lower profit at α=75% and vice versa. Thus the optimal design depends on a firm's sensitivity to downside risk: a risk averse firm would prefer the design resulting in the blue α=25% curve because there is less loss associated with downside risk. A risk-seeking firm would prefer the design resulting in the green α=75% curve because it has the greatest upside potential (fatter tails). The red curve is the design resulting from optimizing for the point estimate of market share and ignoring the uncertainty in $\boldsymbol{\beta}$. As seen in Figure 4, at α=25%, the optimal product designed for α=25% has profit factor approximately 25% larger than the optimal product designed for α=75%.

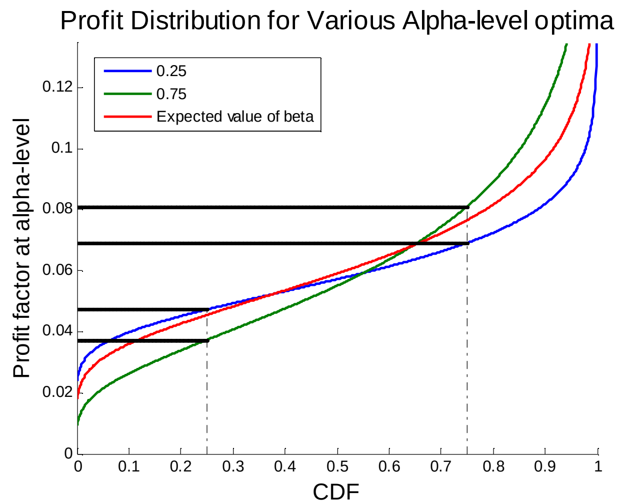

FIGURE 4 – CDF OF PROFIT DISTRIBUTIONS ILLUSTRATING THAT DIFFERENT DESIGNS ARE PREFERRED FOR α =0.25 VS. α = 0.75

Varying t_j to maximize profit results in a trade-off for the manufacturer between increased utility (demand) from improved consumer operating cost vs. higher manufacturing costs. In this study, lower levels of fuel-saving technology implementation are optimal for lower α-levels. As t_j is increased, the vehicle is increasingly differentiated from other vehicles in the data set. This change disproportionately affects the mean and variance of $g(\hat{\boldsymbol{\beta}})$. Larger values of t_j cause greater variation in the profit distribution, so that with more extreme product positioning, there is greater potential for upside but also more downside risk associated with the design. A sensitivity case using more optimistic estimates for technology costs resulted in different values of t_j but a similar trend with respect to α.

4.4. Assessing the delta method approximation

In order to check the quality of the delta method approximation, we compare the distribution obtained for the optimal design found at α=25% using a Monte Carlo simulation vs. the delta method.

First, we take 50,000 draws of the coefficients using the covariance matrix obtained in the logit estimation. The simulated distributions of the parameters are shown in Figure

5, which displays both a contour plot of the multivariate distribution and a set of random draws from the distribution. We use these simulated draws to find a simulated distribution of the g function (equation 9) and compare it with the delta method approximation. Using the g function distribution, we can also compute the market share distribution since $\hat{s}_j = \left(1 + e^{g(\hat{\boldsymbol{\beta}},\mathbf{X})}\right)^{-1}$.

Figures 6 and 7 show, respectively, the comparisons of the simulated g function and market share distributions with the ones obtained by the delta method. The delta method approximation yields high accuracy in this example.

5. CONCLUSIONS

Uncertainty in consumer choice model predictions implies uncertainty about the profit a product would generate. We propose a method for incorporating discrete choice model parameter uncertainty in the design decision problem and for determining the optimal design of a product given a specified level of risk tolerance. In the proposed method, the modeler specifies the level of sensitivity to downside risk by setting α. Specifically, π_j^α is defined as the value below which $\alpha \epsilon [0,1]$ of the profit distribution $\hat{\pi}_j$ lies, and the design is optimized to maximize π_j^α, rather than the expected value of profit. We apply the delta method to derive an estimated closed-form function for π_j^α in the case of the multinomial logit model. The closed-form function enables the optimization problem to be computationally efficient, and it is preferable over methods requiring a simulation-based approach when applicable.

We demonstrate the method in a simple vehicle design case-study, where the delta method is shown to yield a close approximation to the true distribution. We find that the optimal solution varies with α, and the optimal solution at one α-level may be significantly less profitable at another α-level. Thus, optimal design choices depend on risk preference. For the design of the new vehicle in the case study, we find that as level of risk tolerance increases, the optimal profit at the α-level is obtained by increasing the level of fuel-saving technology implementation and differentiating the new vehicle attributes from the average of the competitor attribute values.

The proposed methodology addresses only the uncertainty of model parameter estimates caused by missing data; therefore, it is useful in situations with limited data where model specification can be assumed to be correct, such as some conjoint experiments. Future work may expand the method to be used with other choice models and address other sources of uncertainty, such as model misspecification.

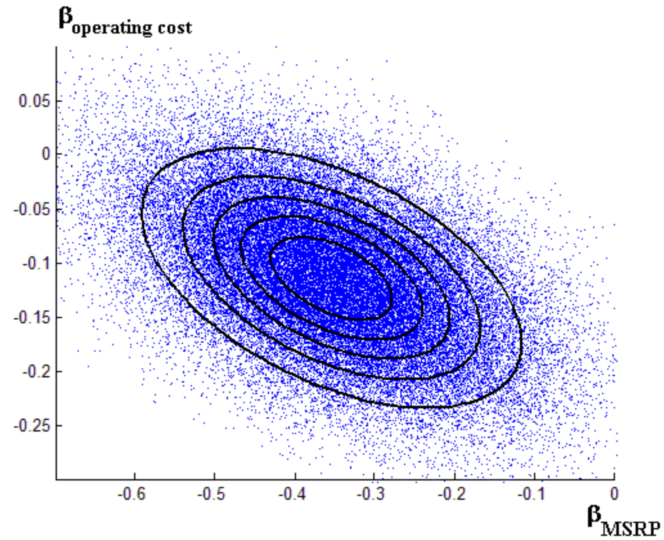

FIGURE 5 – DISTRIBUTIONS OF SIMULATED BETA COEFFICIENTS

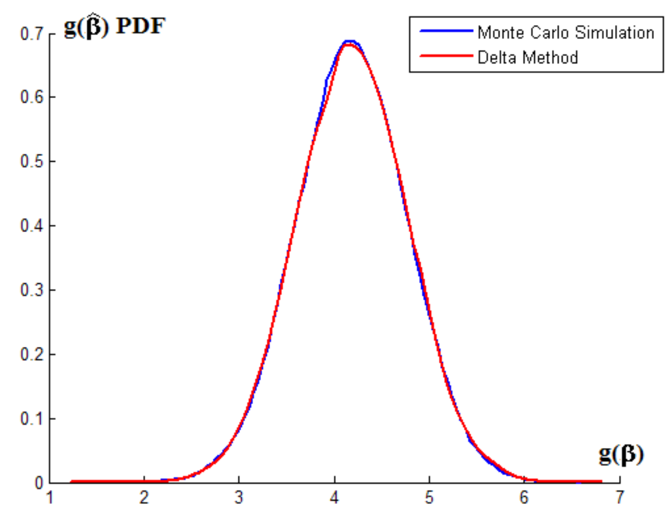

FIGURE 6 – COMPARISON OF SIMULATED AND APPROXIMATED g FUNCTION

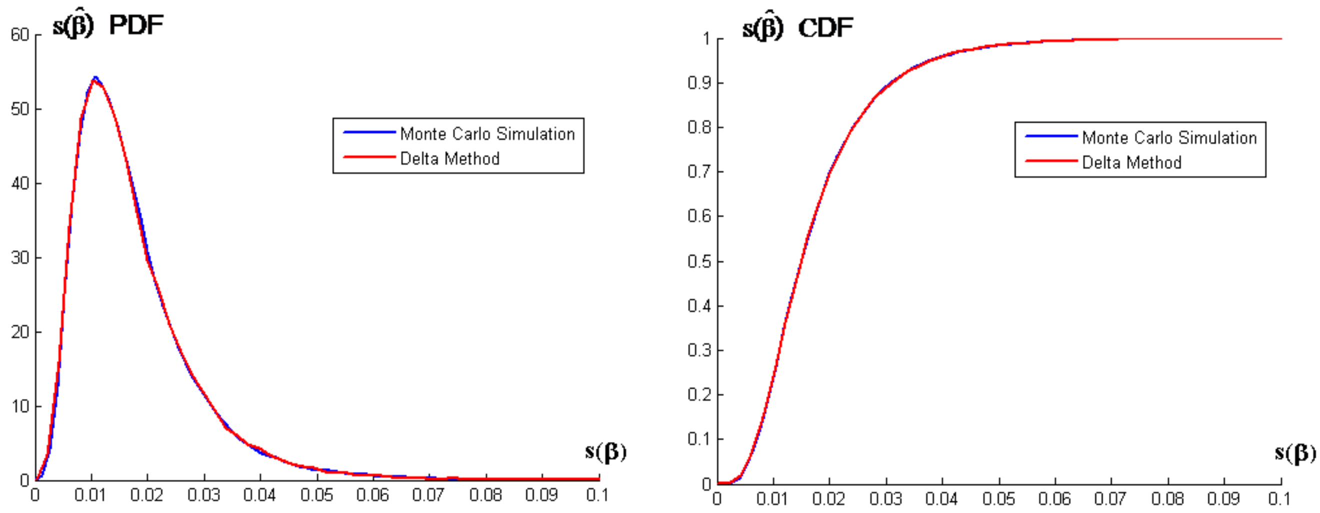

FIGURE 7 – COMPARISON OF SIMULATED AND APPROXIMATED MARKET SHARE

REFERENCES

[1] Green, P. E., Carroll, J. D., and Goldberg, S. M., 1981. "A General Approach to Product Design Optimization via Conjoint Analysis". *Journal of Marketing*, **45**(3), pp. 17-37.

[2] Gavish, B., Horsky, D., and Srikanth, K., 1983. "An Approach to the Optimal Positioning of a New Product". *Management Science,* **29**(11), pp. 1277-1297.

[3] Green, P. E. and Krieger, A. M., 1985. "Models and Heuristics for Product Line Selection". *Marketing Science*, **4**(1), pp. 1-19.

[4] Dobson, G. and Kalish, S., 1988. "Positioning and Pricing a Product Line". *Marketing Science*, **7**(2), pp. 107-125.

[5] Gonul, F. and Srinivasan, K., 1993, "Modeling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues". *Marketing Science*, **12**(3), pp. 213-229.

[6] D'Souza, B. and Simpson, T. W., 2003. "A Genetic Algorithm Based Method for Product Family Design Optimization". *Engineering Optimization*, **35**(1), pp. 1-18.

[7] Heese, H.S. and Swaminathan, J. M., 2006. "Product Line Design with Component Commonality and Cost-Reduction Effort". *Manufacturing & Service Operations Management*, **8**(2), pp. 206-219.

[8] Choi, S. C., Desarbo, W. S., and Harker, P. T., 1991. "Product Positioning under Price Competition". *Management Science*, **36**(2), pp. 175-199.

[9] Shiau, C.-S. and Michalek, J.J., 2009. "Should designers worry about market systems?". *ASME Journal of Mechanical Design*, **131**(1), pp. 1-9.

[10] Shiau, C.-S. and Michalek, J.J., 2009. "Optimal product design under price competition". *ASME Journal of Mechanical Design*, **131**(7), pp. 1-10.

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

[12] Tyagi, R. K., 2000. "Sequential Product Positioning Under Differential Costs". *Management Science*, **46**(7), pp. 928-940.

[13] Villas-Boas, J. M., 1998. "Product Line Design for a Distribution Channel". *Marketing Science*, **17**(2), pp. 156-169.

[14] Luo, L., Kannan, P. K., and Ratchford, B. T., 2007. "New Product Development Under Channel Acceptance". *Marketing Science*, **26**(2), pp. 149-163.

[15] Williams, N., Azarm, S., and Kannan, P. K., 2008. "Engineering Product Design Optimization for Retail Channel Acceptance". *ASME Journal of Mechanical Design*, **130**(06), pp. 1-10.

[16] Williams, N., Kannan, P.K., and Azarm, S., 2011. "Retail Channel Structure Impact on Strategic Engineering Product Design". *Management Science*, **57**(5), pp. 897-914.

[17] Orhun, A. Y., 2009. "Optimal Product Line Design When Consumers Exhibit Choice Set Dependent Preferences". *Marketing Science*, **28**(5), pp. 868-886.

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

[19] Wassenaar, H. J. and Chen, W., 2003. "An Approach to Decision Based Design with Discrete Choice Analysis for Demand Modeling". *ASME Journal of Mechanical Design*, **125**(3), pp. 490-497.

[20] Michalek, J.J., Feinberg, F. M., and Papalambros, P. Y., 2005. "Linking Marketing and Engineering Product Design Decisions via Analytical Target Cascading". *Journal of Product Innovation Management*, **22**, pp. 42-62.

[21] Besharati, B., Luo L., Azarm, S., and Kannan, P. K., 2006. "Multi-Objective Single Product Robust Optimization: An Integrated Design and Marketing Approach". *ASME Journal of Mechanical Design*, **128**(4), pp. 884-892.

[22] Michalek, J.J., Ceryan, O., Papalambros, P. Y., and Koren, Y., 2006. "Balancing Marketing and Manufacturing Objectives in Product Line Design". *Journal of Mechanical Design*, **128**(6), pp. 1196-1204.

[23] Dolan, B., and Lewis, K., 2008. "Robust Product Family Consolidation and Selection". *Journal of Engineering Design*, **19**(6), pp. 553-569.

[24] Kumar, D., Chen, W., and Simpson, T. W., 2009. "A Market-driven Approach to Product Family Design". *International Journal of Production Research*, **47**(1), pp. 71-104.

[25] Luo, L., 2010. "Product Line Design for Consumer Durables: An Integrated Marketing and Engineering Approach." *Journal of Marketing Research*, to appear.

[26] Michalek, J.J., Feinberg, F.M., Ebbes, P., Adigüzel, and Papalambros, P. Y., 2011. "Enhancing Marketing with Engineering: Optimal Product Line Design for Heterogeneous Markets". *International Journal of Research in Marketing*, **28**, pp. 1-12.

[27] Hsu, A. and Wilcox, R. T., 2000. "Stochastic Prediction in Multinomial Logit Models". *Management Science*, **46**(8), pp. 1137-1144.

[28] Frischknecht, B., Whitefoot K., and Papalambros, P.Y., 2010. "On the Suitability of Econometric Demand Models in Design for Market Systems". *Journal of Mechanical Design*, **132**(12), pp. 57-68.

[29] Raman, K. and Chatterjee, R., 1995. "Optimal Monopolist Pricing under Demand Uncertainty in Dynamic Markets". *Management Science*, **41**(1), pp. 144-162.

[30] Montgomery, A. L. and Bradlow, E. T., 1999. "Why Analyst Overconfidence about the Functional Form of Demand Models Can Lead to Overpricing". *Marketing Science*, **18**(4), pp. 569-583.

[31] Abramson, C., Andrews, R. L., Currim, I. S., and Jones, M., 2000. "Parameter Bias from Unobserved Effects in the Multinomial Logit Model of Consumer Choice". *Journal of Marketing Research*, **37**(4), pp. 410-426.

[32] Louviere, J. J., Street, D., Carson, R., Ainslie, A., Deshazo, J. R., Cameron, T., Hensher, D., Kohn, R., and Marley, T., 2002. "Dissecting the Random component of Utility". *Marketing Letters*, **13**(3), pp. 177-193.

[33] Swait, J., Adamowicz, W., Hanemann, M., Diederich, A., Krosnick, J., Layton, D., Provencher, W., Schkade, D., and Tourangeau, R., 2002. "Context Dependence and Aggregation in Disaggregate Choice Analysis". *Marketing Letters*, **13**(3), pp. 195-205.

[34] Louviere, J. J., 2001. "What if Consumer Experiments Impact Variances as well as Means? Response variability as a Behavioral Phenomenon". *Journal of Consumer Research*, **28**(3), pp. 506–511.

[35] Salisbury, L. C., and Feinberg, F. M., 2010. "Alleviating the Constant Stochastic Variance Assumption in Decision Research: Theory, Measurement, and Experimental Test". *Marketing Science*, **29**(1), pp. 1-17.

[36] Luo, L., Kannan, P. K., Besharati, B., and Azarm, S., 2005. **"**Design of Robust New Products under Variability: Marketing Meets Design". *Journal of Product Innovation Management*, **22**(2), pp. 177-192.

[37] Kalyanam, K., 1996. "Pricing Decisions under Demand Uncertainty: A Bayesian Mixture Model Approach". *Marketing Science*, **15**(3), pp. 207-221.

[38] Hitsch, G. J., 2006. "An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty". *Marketing Science*, **25**(1), pp. 25-50.

[39] Wassenaar, H.J., Chen, W., and Cheng, J., 2005. "Enhancing Discrete Choice Demand Modeling for Decision-Based Design". *Journal of Mechanical Design*, **127**(4), pp. 514-524.

[40] McFadden, D., 1974, *Frontiers in Econometrics*, Academic Press, New York, NY, USA, pp. 105-142, Chap. 4.

[41] Hausman, J. A. and Wise, D. A., 1978. “A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences”. *Econometrica*, **46**(2), pp. 403-426.

[42] McFadden, D., and Train, K., 2000. “Mixed MNL Models of Discrete Response”. *Journal of Applied Econometrics*, **15**(5), pp. 447-470.

[43] Fiebig, D. G., Keane, M. P., Louviere, J., and Wasi, N., 2009. “The Generalized Multinomial Logit Model.” *Marketing Science*, **29**(3), pp. 393-421.

[44] Train, K. E., 2009, *Discrete Choice Models with Simulation*, Cambridge University Press, New York, NY, USA.

[45] Ofek, E. and Srinivasan, V., 2002. “How Much Does the Market Value an Improvement in a Product Attribute?”. *Marketing Science*, **21**(4), pp. 398-411.

[46] Bucklin, R. E., Siddarth, S., and Silva-Risso, J. M., 2008. “Distribution Intensity and New Car Choice”. *Journal of Marketing Research*, **45**(4), pp. 473-486.

[47] Wooldridge, J. M., 2002, *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, MA, USA.

[48] Ward’s Automotive Group, 2008. Detailed 2007 US Automotive Sales Data. On the WWW. URL http://WardsAuto.com.

[49] Shiau, C.-S., Michalek, J. J., and Hendrickson, C., 2009. “A Structural Analysis of Vehicle Design Responses to Corporate Average Fuel Economy Policy”. *Transportation Research Part A: Policy and Practice*, **43**(9-10), pp. 814-828.

[50] Erdem, T. and Keane, M.P., 1996. “Decision-making Under Uncertainty: Capturing Dynamic Choice Processes in Turbulent Consumer Goods Markets”. *Marketing Science*, **15**(1), pp. 1-20.

Appendix A: Finding the Derivative $\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$ for a Multinomial Logit Model without Outside Good

Since $g(\boldsymbol{\beta}) = \ln\left(\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}\right)$

$$\Rightarrow \quad \frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} \ln\left(\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}\right) \tag{20}$$

$$= \frac{1}{\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}} \frac{\partial}{\partial \boldsymbol{\beta}}\left(\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}\right)$$

$$= \frac{1}{\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}} \left(\sum_{k\in J\setminus j} \frac{\partial}{\partial \boldsymbol{\beta}} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}\right)$$

$$= \frac{1}{\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}} \left(\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)} \frac{\partial}{\partial \boldsymbol{\beta}}\left(\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)\right)\right)$$

$$= \frac{1}{\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}} \left(\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)} (\mathbf{x}_k-\mathbf{x}_j)\right)$$

$$= \frac{\sum_{k\in J\setminus j} (\mathbf{x}_k-\mathbf{x}_j) e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}}{\sum_{k\in J\setminus j} e^{\boldsymbol{\beta}^{\mathrm{T}}(\mathbf{x}_k-\mathbf{x}_j)}} \tag{21}$$

Appendix B: Delta Method Proof

Adapted from Wooldridge 2002

Theorems and properties utilized:

Asymptotic properties of estimators: Let $\{\beta_N : N = 1,2,\ldots\}$ be a sequence of estimators of the P x 1 vector $\beta_0 \in \mathrm{B}$. If β_N is an unbiased and consistent estimator of β_0, then by the multivariate central limit theorem $\sqrt{N}(\beta_N - \beta_0) \xrightarrow{a} Normal(0, V)$ for any β_0 where V is a P x P positive semidefinite matrix. We say that β_N is $\sqrt{N}$ - asymptotically normally distributed and V is the asymptotic variance of $\sqrt{N}(\beta_N - \beta_0)$ denoted $\mathrm{Avar}\left(\sqrt{N}(\beta_N - \beta_0)\right) = V$. ($V$ is necessarily positive semidefinite because $\mathrm{Avar}\left(\sqrt{N}(\beta_N - \beta_0)\right) = \mathrm{E}[\sqrt{N}(\beta_N - \beta_0)\cdot\sqrt{N}(\beta_N - \beta_0)']$.)

Variance property: For any nonstochastic Q x P matrix R, with $\mathrm{rank}(R) = Q$, $\sqrt{N}R(\beta_N - \beta_0) \overset{a}{\sim} Normal(0, RVR')$

Asymptotic equivalence lemma: $\{x_N\}$ and $\{z_N\}$ are sequences of K x 1 random vectors. If $z_N \xrightarrow{d} z$ and $x_N - z_N \xrightarrow{p} 0$, then $x_N \xrightarrow{d} z$.

Mean value theorem: $f'(c) = \frac{f(b)-f(a)}{b-a}$ for some c: b > c > a

Slutsky's theorem: If $f(\cdot)$ is continuous, then $f\{x_N\} \xrightarrow{d} f(x)$ when random vectors $\{x_N\} \xrightarrow{d} x$

Landau Symbols and convergence properties: Let x be a continuous variable tending to some limit and let $\phi(x)$ be a positive function and $f(x)$ be any function. Then define $f = \mathrm{O}(\phi)$ to mean that $|f| < A\phi$ for some constant A and all values x, $f = \mathrm{o}(\phi)$ to mean that $\frac{f}{\phi} \longrightarrow 0$ as x approaches infinity. $f = \mathrm{o}(\phi)$ implies and is stronger than $f = \mathrm{O}(\phi)$.

Delta method theorem:

Let $\mathrm{f}: \mathrm{B} \to \mathbb{R}^Q$ be a continuously differentiable function on parameter space $\mathrm{B} \subset \mathbb{R}^P$ where $Q \le P$ and assume β is on the interior of the parameter space. Define $F(\beta) = \nabla_\theta \mathrm{f}(\beta)$ as the Q x P Jacobian of f(β).

If

$$\sqrt{N}(\beta_N - \beta_0) \xrightarrow{d} Normal(0, V) \tag{22}$$

and V is positive semidefinite,

then:

$$\sqrt{N}(\mathrm{f}(\beta_N) - \mathrm{f}(\beta_0)) \xrightarrow{a} Normal(0, \mathrm{F}(\beta_0)V\mathrm{F}(\beta_0)'). \tag{23}$$

Proof:

Because β_0 is in the interior of B and because β_N is a consistent estimator such that $\mathrm{plim}\,\beta_N = \beta_0$, $\hat{\beta}_N$ is an open convex subset of B containing β_0 with probability approaching 1. Therefore, with probability approaching 1, we can use the mean value theorem to write:

$$\mathrm{f}(\beta_N) = \mathrm{f}(\beta_0) + \ddot{\mathrm{F}}_N(\beta_N - \beta_0) \tag{24}$$

Where $\ddot{\mathrm{F}}_N$ denotes the Jacobian matrix $\mathrm{F}_N(\beta)$ evaluated at mean values between β_N and β_0. Rearranging the equation:

$$\sqrt{N}(\mathrm{f}(\beta_N) - \mathrm{f}(\beta_0)) = \ddot{\mathrm{F}}_N\sqrt{N}(\beta_N - \beta_0) \tag{25}$$

Adding and subtracting $\mathrm{F}(\beta_0)\sqrt{N}(\beta_N - \beta_0)$ to the right hand side of the equation yields:

$$\sqrt{N}(\mathrm{f}(\beta_N) - \mathrm{f}(\beta_0)) = \mathrm{F}(\beta_0)\sqrt{N}(\beta_N - \beta_0) + \left(\ddot{\mathrm{F}}_N - \mathrm{F}(\beta_0)\right)\sqrt{N}(\beta_N - \beta_0) \tag{26}$$

Because the mean values are between β_N and β_0, they converge in probability to β_0 and we can apply Slutsky's theorem, $\ddot{\mathrm{F}}_N \xrightarrow{p} \mathrm{F}(\beta_0)$, which implies that $\ddot{\mathrm{F}}_N - \mathrm{F}(\beta_0)$ is $o_p(1)$ and the equation becomes:

$$\begin{aligned}
&\sqrt{N}\big(\mathrm{f}(\beta_N) - \mathrm{f}(\beta_0)\big) \\
&= \mathrm{F}(\beta_0)\sqrt{N}(\beta_N - \beta_0) + \mathrm{o}_p(1)\cdot\sqrt{N}(\beta_N - \beta_0) \\
&= \mathrm{F}(\beta_0)\sqrt{N}(\beta_N - \beta_0) + \mathrm{o}_p(1)\cdot \mathrm{O}_p(1) \\
&= \mathrm{F}(\beta_0)\sqrt{N}(\beta_N - \beta_0) + \mathrm{o}_p(1) \\
&\xrightarrow{p} \mathrm{F}(\beta_0)\sqrt{N}(\beta_N - \beta_0)
\end{aligned} \tag{27}$$

By the asymptotic equivalence lemma and property of variance stated above with $\mathrm{F}(\beta_0)=R$:

$$\sqrt{N}\mathrm{F}(\beta_0)(\beta_N - \beta_0) \xrightarrow{a} Normal(0, \mathrm{F}(\beta_0)V\mathrm{F}(\beta_0)') \tag{28}$$

Appendix C: Competitor Vehicle Characteristics

Data from Ward's Automotive [48]

Model #	Price ($)	City MPG	HWY MPG	Harmonic average	Fuel consumption (gal./mi.)	Op cost ($/mile)
1	18,850	23	33	28	0.04	$0.10
2	19,220	24	34	29	0.03	$0.10
3	18,565	22	28	25	0.04	$0.11
4	22,305	60	51	55	0.02	$0.05
5	17,995	23	31	27	0.04	$0.11
6	15,365	28	34	31	0.03	$0.09
7	22,315	19	28	23	0.04	$0.12
8	10,415	28	36	32	0.03	$0.09
9	33,885	21	30	25	0.04	$0.11
10	20,305	23	33	28	0.04	$0.10
11	21,454	24	32	28	0.04	$0.10
12	13,495	27	35	31	0.03	$0.09
13	13,065	28	35	31	0.03	$0.09
14	32,150	19	27	23	0.04	$0.13
15	34,295	20	29	24	0.04	$0.12
16	13,820	22	30	26	0.04	$0.11
17	20,595	20	30	24	0.04	$0.12
18	20,915	19	27	23	0.04	$0.13
19	19,525	24	31	27	0.04	$0.10
20	44,195	20	30	24	0.04	$0.12
21	28,655	21	28	24	0.04	$0.12
22	22,915	19	27	23	0.04	$0.13
23	52,325	18	24	21	0.05	$0.14
24	30,405	18	27	22	0.05	$0.13
25	16,955	22	30	26	0.04	$0.11
26	23,590	22	31	26	0.04	$0.11
27	61,715	19	27	23	0.04	$0.13
28	19,445	23	31	27	0.04	$0.11
29	29,890	18	26	22	0.05	$0.13
30	20,720	24	34	29	0.03	$0.10
31	19,899	23	30	26	0.04	$0.11
32	44,865	21	29	25	0.04	$0.12
33	21,515	21	31	26	0.04	$0.11
34	42,765	18	27	22	0.05	$0.13
35	18,995	24	32	28	0.04	$0.10
36	42,670	19	27	23	0.04	$0.13
37	39,400	18	27	22	0.05	$0.13
38	46,450	18	26	22	0.05	$0.13
39	22,135	20	27	23	0.04	$0.12
40	35,115	21	29	25	0.04	$0.12
41	49,000	19	28	23	0.04	$0.12
42	27,385	22	31	26	0.04	$0.11
43	20,825	22	30	26	0.04	$0.11
44	53,090	18	25	21	0.05	$0.13

Appendix D: The Inverse of the Information Matrix is the Asymptotic Variance of Maximum Likelihood Estimators

Adapted from Wooldridge 2002

Definitions:

- An M-estimator solves the problem: $\max_{\beta\in\mathrm{B}} N^{-1}\sum_{i=1}^{N} q(w_i, \beta)$, where $w=(x,y)$ are the data, β is the parameter vector, and q is the quality function associated with the estimator (e.g. an error function). The parameter vector β_0 is assumed to uniquely solve the population problem $\max_{\beta\in\mathrm{B}} \mathrm{E}[q(w,\beta)]$.

- In the case of maximum likelihood M-estimators (MLE), $q(w_i,\beta) = \log\big(f(y_i|x_i;\beta)\big)$, where f is the likelihood function. Because it is a maximization problem, the expected value of the hessian of the objective function $\nabla_\beta^2 \log\big(f_i(\beta)\big)$, denoted H($\beta$), is negative definite at $\beta=\beta_0$.

- $\{f(\cdot\,|x,\beta): x\in X, \beta\in B\}$ denotes the parametric model of conditional density

- $\nu(dy)$ is a σ-finite measure, which for the purposes of this proof just denotes the increment over which the conditional density can be integrated

- $l_i \equiv l(y_i, x_i, \beta) \equiv \log\big(f(y_i|x_i;\beta)\big) \equiv$ log likelihood

- Score of the log likelihood:
$$\equiv s_i(\beta) \equiv \nabla_\beta l_i(\beta)' = \left(\frac{\partial l_i(\beta)}{\partial \beta_1}, \frac{\partial l_i(\beta)}{\partial \beta_2}, \ldots \frac{\partial l_i(\beta)}{\partial \beta_P}\right)'$$

Theorems and properties utilized:

- For M-estimators of β_0, $\sqrt{N}(\beta_N - \beta_0) \xrightarrow{a} Normal(0, A_0^{-1}B_0A_0^{-1})$ where:

  $A_0 \equiv -\mathrm{E}[\mathrm{H}(\beta_0)]$ (for maximization problems where the expected value of H is negative definite)

  $B_0 \equiv \mathrm{Var}[s_i(\beta_0)] = \mathrm{E}[s_i(\beta_0)s_i(\beta_0)']$

- For M-estimators, $\mathrm{E}[s_i(\beta_0)]=0$ is a necessary condition of asymptotic normality

- $l_i = \ln(f_i) \rightarrow f_i = e^{l_i} \rightarrow \nabla f_i = \nabla e^{l_i} = e^{l_i}\cdot\nabla l_i = f_i\cdot\nabla l_i = f_i\cdot s_i$

- Law of iterated expectations: $\mathrm{E}[x]=\mathrm{E}_y[\mathrm{E}[x|y]]$

Information matrix theorem:

Let $\{(x_i,y_i)\}$: i=1,2,…} be a random sample with $x_i \in X \in \mathbb{R}^K$ and $y_i \in Y \in \mathbb{R}^G$, $\mathrm{B}\in\mathbb{R}^P$ be the parameter set and $\{f(\cdot\,|x;\beta): x\in X, \beta\in \mathrm{B}\}$ denote the parametric model of conditional density. Under standard regularity conditions for asymptotic M-estimators:

$$\sqrt{N}(\beta_N - \beta_0) \xrightarrow{d} Normal(0, A_0^{-1} B_0 A_0^{-1}) \qquad (29)$$

If it is assumed that:

1. $f(\cdot\,|x,\beta)$ is a true density with respect to $\upsilon(dy)$ for all x and β so that $\int_Y f(y|x)\upsilon(dy) = 1$

2. For some $\beta_0 \in \mathrm{B}$, $f_0(\cdot\,|x) = f(\cdot\,|x;\beta_0)$ for all $x \in X$, and β_0 is the unique solution to $\max_{\beta\in \mathrm{B}} \mathrm{E}[l_i(\beta)]$

3. β_0 is on the interior of compact set B

4. For each $\beta \in \mathrm{B}$, $l(\cdot\,|\beta)$ is a Borel measurable function on Y x X

5. For each $(y,x) \in Y \times X, l(y,x,\cdot)$ is twice differentiable on int(B)

6. The elements of $\nabla_\beta^2\big(l(y,x,\beta)\big)$ are bounded in absolute value by a function *b(y,x)* with finite expectation

7. The interchanges of ∇ and $\int$ hold for all $\beta \in \mathrm{int(B)}$

8. $A_0 \equiv -\mathrm{E}[\mathrm{H}_i(\beta)]$ is positive definite

Then for maximum likelihood M-estimators:

$$\sqrt{N}(\beta - \beta_0) \xrightarrow{d} Normal(0, A_0^{-1}) \qquad (30)$$

and therefore the asymptotic variance, $\mathrm{Avar}(\beta)$, is equal to $\frac{A_0^{-1}}{N}$ where $A_0 \equiv -\mathrm{E}[\mathrm{H_i}(\beta_0)]$ is the information matrix.

Proof:

Let $\mathrm{E}_\beta[\cdot\,|x_i]$ denote conditional expectation with respect to the density $f(\cdot\,|x_i,\beta)$ for any $\beta \in \mathrm{B}$. Then, by definition:

$$\mathrm{E}_\beta[s_i(\beta)|x_i] = \int_\mathrm{Y} s(y,x_i,\beta) f(y|x_i;\beta)\upsilon(dy) \qquad (31)$$

Assuming the validity of interchanging integration and differentiation on the int(B) for all $x_i \in X, \beta \in \mathrm{int(B)}$:

$$\nabla_\beta \int_\mathrm{Y} f(y|x_i;\beta)\upsilon(dy) = \int_\mathrm{Y} \nabla_\beta f(y|x_i;\beta)\upsilon(dy) = 0 \qquad (32)$$

since $\int_\mathrm{Y} f(y|x_i;\beta)\upsilon(dy)$ is unity for all β. Therefore, the partial derivatives with respect to β must be identically zero.
Rewriting $\nabla_\beta f(y|x_i;\beta)$ as

$$\nabla_\beta l(y|x_i;\beta)\cdot f(y|x_i;\beta) = s(y|x_i;\beta)\cdot f(y|x_i;\beta) \qquad (33)$$

yields:

$$\int_\mathrm{Y} s_i(\beta) f(y|x_i;\beta)\upsilon(dy) = \mathrm{E}_\theta[s_i(\beta)|x_i] = 0 \qquad (34)$$

Taking the derivative and again interchanging integration and differentiation results in:

$$\nabla\left(\int_\mathrm{Y} s_i(\beta) f(y|x_i;\beta)\upsilon(dy)\right) = \int_\mathrm{Y} \nabla\big(s_i(\beta) f(y|x_i;\beta)\upsilon(dy)\big) \qquad (35)$$

$$= \int_\mathrm{Y} \nabla s_i(\beta) f(y|x_i;\beta)\upsilon(dy) + \int_\mathrm{Y} s_i(\beta)\nabla f(y|x_i;\beta)\upsilon(dy)$$

$$= \mathrm{E}[\nabla s_i(\beta)|x_i] + \int_\mathrm{Y} s_i(\beta)s_i(\beta) f(y|x_i;\beta)\upsilon(dy)$$

$$= \mathrm{E}[\nabla s_i(\beta)|x_i] + \int_\mathrm{Y} s_i^2(\beta) f(y|x_i;\beta)\upsilon(dy)$$

$$= \mathrm{E}[\nabla s_i(\beta)|x_i] + \mathrm{E}[s_i^2(\beta)|x_i]$$

$$= \mathrm{E}\big[\nabla_\theta^2\big(l_i(\beta)\big)|x_i\big] + \mathrm{Var}[s_i(\beta)|x_i]$$

$$= \mathrm{E}[\mathrm{H_i}(\beta)|x_i] + \mathrm{Var}[s_i(\beta)|x_i] = 0$$

$$\rightarrow -\mathrm{E}[\mathrm{H_i}(\beta)|x_i] = \mathrm{Var}[s_i(\beta)|x_i] \qquad (36)$$

for all $\beta \in \mathrm{B}$. Substituting $\beta = \beta_0$ yields:

$$-\mathrm{E}[\mathrm{H_i}(\beta_0)|x_i] = \mathrm{E}[s_i(\beta_0)s_i(\beta_0)'|x_i] \qquad (37)$$

Taking the expectation value with respect to the distribution of x:

$$-\mathrm{E_x}\big[\mathrm{E}[\mathrm{H_i}(\beta_0)|x_i]\big] = \mathrm{E_x}\big[\mathrm{E}[s_i(\beta_0)s_i(\beta_0)'|x_i]\big] \qquad (38)$$

Using the law of iterated expectations:

$$-\mathrm{E}[\mathrm{H_i}(\beta_0)] = \mathrm{E}[s_i(\beta_0)s_i(\beta_0)'] \qquad (39)$$

or

$$A_0 = B_0 \qquad (40)$$

This implies:

$$\sqrt{N}(\beta_N - \beta_0) \xrightarrow{a} Normal(0, A_0^{-1}A_0A_0^{-1})$$
$$= Normal(0, IA_0^{-1}) = Normal(0, \mathrm{E}[\mathrm{H}_0]^{-1}) \qquad (41)$$