An Empirical Examination of Maximum Entropy Estimation.

Randall Charles Campbell
Louisiana State University and Agricultural & Mechanical College

AN EMPIRICAL EXAMINATION OF MAXIMUM ENTROPY ESTIMATION

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Economics

by
Randall C. Campbell
B.S., Purdue University, 1989
M.S., Louisiana State University, 1994
May 1999

UMI Number: 9926415

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

DEDICATION

This work is dedicated to my parents, Jean A. and Charles S. Campbell, and to my wife, Rebecca J. Campbell.

ACKNOWLEDGEMENTS

I wish to acknowledge and thank the many individuals who have supported me and helped me to achieve my goal of a doctorate in economics. R. Carter Hill inspired my interest in econometrics through his excellence in the classroom. He has been an outstanding major professor in addition to being a good friend. He has generously given his time and shared his ideas while teaching me how to conduct research.

I want to thank all of my professors at LSU for their hard work and the knowledge they have given me. I particularly want to thank the other members of my committee for their time and insight: M. Dek Terrell, Doug McMillin, Luis Escobar, and Charlene Henderson.

Finally, I wish to thank my family. My mother and father have been supportive throughout my academic career and have inspired me to always want to learn. My brothers, Scott and Darin, and my sister Kathy, who visited me every year while I was in school, have encouraged me to finish this work. Finally, my wife Rebecca J. Campbell, whom I met at LSU, has sacrificed in many ways so that I could finish. She has made the entire process of completing this dissertation a very enjoyable one by being beside me and supporting me every day.

TABLE OF CONTENTS

ABSTRACT

Maximum entropy estimation is a relatively new estimation technique in econometrics. We carry out several Monte Carlo experiments using real data as a basis in order to understand the properties of the maximum entropy estimator. We compare the maximum entropy and generalized maximum entropy estimators to traditional estimation techniques in linear regression, binary choice, and multinomial choice models. In addition, we discuss maximum entropy estimation in censored and truncated regression models.

We find that the generalized maximum entropy estimator dominates the logit estimator and the multinomial logit estimator in Monte Carlo experiments. The generalized maximum entropy estimator in discrete choice models allows us to jointly estimate the unknown probabilities and the unknown errors resulting in more uniform predicted probabilities and reducing the variance of the parameter estimates. In the linear regression problem, the generalized maximum entropy estimator allows us to impose nonsample information about the unknown parameters and errors. However, we must impose a set of support points for unknown parameters and errors, which is not always an easy thing to do. We find that when we do specify nonsample information that is correct, the generalized maximum entropy estimator has lower risk than either the ordinary least squares or the inequality restricted least squares estimators. From our sampling experiments using real data, we find that maximum entropy estimation is a viable estimation technique in several econometric models.

CHAPTER 1

MAXIMUM ENTROPY ESTIMATION

1.1 Introduction

This dissertation examines statistical inference and estimation through the method of maximum entropy (ME). Shannon (1948) defines the entropy function as a means of measuring uncertainty in a discrete probability distribution. Jaynes (1957a, 1957b) introduces the maximum entropy method of estimation. Jaynes argues that when estimating an unknown probability distribution we should choose the distribution that maximizes entropy and is compatible with our prior information. Shore and Johnson (1980) and Skilling (1988) provide an axiomatic justification for using ME to estimate the unknown probabilities in a discrete probability distribution. Golan, Judge, and Miller (1996) apply ME estimation to a wide range of econometric problems.

Denzau, Gibbons, and Greenberg (1989), Soofi (1992), and Golan, Judge, and Perloff (1996) discuss ME estimation of multinomial choice models. Golan, Judge, and Miller (1996) extend the ME methodology to estimate linear regression, SUR, simultaneous equations, and multinomial choice models. They estimate these models both with and without an error term included in the entropy function and its constraints. Golan, Judge, and Perloff (1997) apply ME to both censored and ordered multinomial choice regression models. We examine ME estimation in these and other contexts.

We contribute to the literature by comparing empirical results from ME estimation to those of traditional estimation techniques. In addition, we empirically examine small sample properties of ME estimation. Golan, Judge, and Miller (1996), and Golan, Judge, and Perloff (1996) examine large sample properties and small sample performance of the ME estimator in terms of risk functions within Monte Carlo sampling experiments. We examine small sample properties of the ME estimator in the context of several statistical models using real data as a basis for Monte Carlo experiments, as opposed to previous studies that use artificial data.

We estimate several econometric models using ME techniques and compare the ME estimator to traditional estimators. We examine ME estimation in binary choice, multinomial choice and linear regression models, with and without sign and other inequality restrictions on the parameters using real data as a basis for Monte Carlo sampling experiments. Additionally, we discuss ME estimation in censored and truncated samples.

1.2 Maximum Entropy Estimation

Jaynes (1957a, 1957b) proposes the maximum entropy method for estimating the unknown probabilities of a discrete probability distribution. The estimation problem that Jaynes describes involves an observable random variable, x, which assumes the discrete values x_i $(i = 1,2,\ldots,N)$ with corresponding probabilities p_i, which are unknown and unobservable. The term entropy refers to the amount of uncertainty, with regard to predicting future outcomes, represented by a discrete probability distribution. Thus, a distribution that assigns a large probability to certain outcomes has smaller entropy or uncertainty than a distribution in which each outcome has a relatively small probability of occurrence.

Shannon (1948) defines the entropy of a discrete probability distribution as

$$H(p) = -\sum_{i=1}^{N} p_i \ln(p_i), \tag{1.1}$$

where $0 \cdot \ln(0) \equiv 0$. Thus, for a degenerate probability distribution, the entropy is $H(1) = 0$, and there is no uncertainty in predicting future outcomes. For a discrete uniform distribution $H(p) = \ln(N)$, which is the maximum value of the entropy function.

Jaynes (1957a) argues that maximization of Shannon's (1948) entropy measure provides a unique criterion for estimating an unknown probability distribution that does not impose any information beyond what we know. Any other distribution compatible with the observed data has lower entropy, or less uncertainty, in predicting outcomes. Therefore, distributions other than the ME distribution assume information that is not explicitly known. For this reason, Jaynes states that maximum entropy is "maximally noncommittal with regard to missing information" (1957a, p. 620). While it may seem counterintuitive to maximize uncertainty, we are only maximizing the uncertainty represented by the unknown probability distribution. By maximizing entropy subject to our prior

information we select the distribution whose entropy or uncertainty reflects our prior information and no additional information. In addition, by maximizing entropy we assign positive probability to every possible outcome.

The unknown probabilities are estimated by maximizing the entropy function subject to the constraint that the probabilities must sum to one as follows

$$\max H(p) = -\sum_{i=1}^{N} p_i \ln(p_i) \tag{1.2}$$

subject to

$$\sum_{i=1}^{N} p_i = 1. \tag{1.3}$$

The Lagrangian for this ME problem is

$$\mathcal{L} = -\sum_{i=1}^{N} p_i \ln(p_i) + \lambda(1 - \sum_{i=1}^{N} p_i),$$

and the resulting first-order conditions (FOC) are

$$\frac{\partial \mathcal{L}}{\partial p_i} = -1 - \ln(p_i) - \lambda = 0 \qquad \forall i$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 1 - \sum_{i=1}^{N} p_i = 0.$$

Solving the N first-order conditions for the unknown probabilities yields

$$\ln(p_i) = -1 - \lambda,$$

which implies

$$\hat{p}_i = \exp(-1 - \lambda) \qquad \forall i.$$

Since the probabilities must sum to one, we can rewrite the optimal ME probabilities as

$$\hat{p}_i = \frac{\exp(-1-\lambda)}{\sum_{i=1}^{N} \exp(-1-\lambda)} = \frac{1}{N}, \tag{1.4}$$

which is a discrete uniform distribution. Taking the second derivative of the Lagrangian with respect to p_i yields a negative definite Hessian matrix, which implies that the entropy function is globally concave and the ME solution is unique.

ME allows us to uniquely estimate the N unknown probabilities without imposing any constraints, except that the predicted probabilities must sum to one. This is an important property of ME estimation since traditional estimation techniques require that the number of observations be greater than the number of parameters in order to obtain unique estimates. Additionally, traditional estimation techniques often require us to make strong assumptions about the data or the errors in order to obtain a unique solution. ME estimation is a nonparametric estimation technique since it does not require any assumptions to be made about the parameters or error distributions. However, any prior information about the parameter or error values that the researcher has can be incorporated in ME by imposing constraints in the entropy maximization problem. The ME estimates are the most uniform probabilities compatible with the available prior information. A simple example using ME to solve a problem with prior information is illustrated in the following section.

1.3 Jaynes' Dice Problem

Jaynes (1963) considers estimating the probabilities of each possible outcome from a single roll of a six-sided die, given the prior information that the average outcome from a large number of independent rolls is equal to y. ME allows us to obtain unique estimates of the six unknown probabilities using this single piece of prior information as well as the fact that the probabilities must sum to one. Using the ME principle, we choose the unknown probabilities in order to

$$\max H(p) = -\sum_{i=1}^{6} p_i \ln(p_i) \tag{1.5}$$

subject to

$$\sum_{i=1}^{6} p_i x_i = y \tag{1.6}$$

$$\sum_{i=1}^{6} p_i = 1, \tag{1.7}$$

where $x_i = i$ $(i = 1,...,6)$ and $p_i = \Pr(x = x_i)$. The first constraint is the data constraint incorporating the prior information and the second constraint is the additivity constraint, which requires that the probabilities must sum to one. The Lagrangian for the dice problem is given by

$$\mathcal{L} = -\sum_{i=1}^{6} p_i \ln(p_i) + \lambda\left(y - \sum_{i=1}^{6} p_i x_i\right) + \gamma\left(1 - \sum_{i=1}^{6} p_i\right),$$

and the resulting first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial p_i} = -1 - \ln(p_i) - x_i\lambda - \gamma = 0 \qquad \forall i$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = y - \sum_{i=1}^{6} p_i x_i = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = 1 - \sum_{i=1}^{6} p_i = 0.$$

Solving the FOC's yields the optimal ME probabilities

$$\hat{p}_i = \exp(-1 - x_i\lambda - \gamma) \qquad \forall i.$$

Because the additivity constraint requires that the probabilities must sum to one, the above equation can be rewritten as

$$\hat{p}_i = \frac{\exp(-x_i\lambda)}{\sum_{j=1}^{6} \exp(-x_j\lambda)} = \frac{\exp(-x_i\lambda)}{\Omega(\lambda)}, \qquad (1.8)$$

where $\Omega(\lambda)$ is known as the partition function since it is the sum of the relative probabilities. The partition function allows us to partition the relative probabilities into six absolute probabilities, which must sum to one. Since the optimal probabilities are a function of λ, the entropy maximization problem does not have a closed form solution. Therefore, we must use numerical optimization techniques to obtain the ME solution. Using different values of y, representing different prior information, Table 1.1 presents the ME solution to the dice problem. We obtain the ME solution using the GAUSS constrained optimization module. Program 1, the GAUSS program for the dice problem, is included in Appendix B.

Table 1.1 Maximum Entropy Probabilities for Jaynes' Dice Problem

Y	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$H(p)$
1.0	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1.5	0.6637	0.2238	0.0755	0.0255	0.0086	0.0029	0.9534
2.0	0.4781	0.2548	0.1357	0.0723	0.0385	0.0205	1.3675
2.5	0.3475	0.2398	0.1654	0.1142	0.0788	0.0544	1.6136
3.0	0.2468	0.2072	0.1740	0.1461	0.1227	0.1031	1.7485
3.5	0.1667	0.1667	0.1667	0.1667	0.1667	0.1667	1.7918
4.0	0.1031	0.1227	0.1461	0.1740	0.2072	0.2468	1.7485
4.5	0.0544	0.0788	0.1142	0.1654	0.2398	0.3475	1.6136
5.0	0.0205	0.0385	0.0723	0.1357	0.2548	0.4781	1.3675
5.5	0.0029	0.0086	0.0255	0.0755	0.2238	0.6637	0.9534
6.0	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000

The results for the dice problem show that ME selects the most uniform distribution compatible with the given information. If the observed average is $y = 3.5$, a discrete uniform distribution is obtained. However, if the observed average is $y = 5.0$, ME assigns more weight to the higher values of x because this is compatible with the prior information. Additionally, we find that the ME probability table is symmetric. For example, the value of the entropy function is the same for an observed average of $y = 2.0$ as it is for an observed average of $y = 5.0$ since these two values can occur in exactly the same number of ways.

1.4 The Unconstrained Maximum Entropy Dual Problem

Since $\hat{p}$ is a function of λ, we can concentrate the Lagrangian for the ME problem by substituting $\hat{p}(\lambda)$ into the original Lagrangian. This substitution leads to an unconstrained optimization problem with λ as the choice variable. Golan, Judge, and Miller (1996) show that minimization of the concentrated Lagrangian yields the same solution as maximization of the entropy function, and is thus termed the "dual problem."

The ME solution is obtained by maximizing a real-valued concave function, $H(p)$, subject to a set of constraints which are a function of p. A corollary to the Arrow-Hurwicz-Uzawa theorem (Takayama 1985, pp. 94-95) states that $\hat{p}$ will maximize the entropy function subject to the constraints if and only if there exists a $\hat{\lambda} \geq 0$ such that $(\hat{p}, \hat{\lambda})$ is a saddle-point of the Lagrangian. This means that

$$\mathcal{L}(p, \hat{\lambda}) \leq \mathcal{L}(\hat{p}, \hat{\lambda}) \leq \mathcal{L}(\hat{p}, \lambda), \tag{1.9}$$

where $\mathcal{L}(\hat{p}, \lambda)$ represents the concentrated Lagrangian denoted $M(\lambda)$ by Golan, Judge, and Miller (1996). In terms of $M(\lambda)$ the saddle-point property (1.9) implies that

$$M(\hat{\lambda}) \leq M(\lambda). \tag{1.10}$$

Therefore, $\hat{\lambda}$ yields the smallest possible value of $M(\lambda)$, and $\hat{p}(\hat{\lambda})$ maximizes the Lagrangian. Thus, we may obtain the ME probabilities by choosing $\hat{\lambda}$ to minimize $M(\lambda)$ and substituting this value into the optimal probabilities in equation (1.8). This result holds for all of the ME problems that we consider. It is computationally simpler to obtain the ME dual solution rather than the ME primal solution since there are generally fewer parameters to estimate. In the ME dual problem we estimate K unknown parameters while in the ME primal problem we estimate N unknown probabilities. Program 2, in Appendix B, is the GAUSS program to solve the dice problem using the unconstrained dual formulation.

1.5 Outline of the Remaining Chapters

In this chapter, we described maximum entropy estimation and illustrated it using a simple example. In Chapter 2, we discuss maximum entropy (ME) and generalized maximum entropy (GME) estimation in both binary and multinomial choice models. We estimate binary choice models using both credit-scoring data and data on the choice of whether or not to attend post-secondary education (PSE). In addition, we estimate multinomial choice models using the PSE data. We carry out several Monte Carlo experiments using real data sets both with and without sign and other restrictions placed on the parameters. We compare the ME and GME estimators to traditional discrete choice estimators including logit, probit, linear discriminant analysis, and multinomial logit.

In Chapter 3, we discuss GME estimation in linear regression models. We carry out Monte Carlo experiments comparing GME to OLS. We estimate the models both with and without inequality restrictions placed on the parameters. In addition, we develop a new method for imposing inequality restrictions in GME. In Chapter 4, we consider the cost of imposing incorrect inequality restrictions on the GME estimator in the linear regression model. We examine censored regression models and sample selectivity issues in Chapter 5 and give conclusions in Chapter 6.

1.6 References

Denzau, A. T., P. C. Gibbons, E. Greenberg. (1989). Bayesian estimation of proportions with a cross-entropy prior, *Communications in Statistics-Theory and Methods* 18: 1843-1861.

Golan, A., G. Judge, D. Miller. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* New York: John Wiley and Sons.

Golan, A., G. Judge, J. M. Perloff. (1996). A maximum entropy approach to recovering information from multinomial response data, *Journal of the American Statistical Association* 91 (June): 841-853.

________. (1997). Estimation and inference with censored and ordered multinomial response data, *Journal of Econometrics* 79: 23-51.

Jaynes, E. T. (1957a). Information theory and statistical mechanics, *Physical Review* 106: 620-630.

________. (1957b). Information theory and statistical mechanics II, *Physical Review* 108: 171-190.

________. (1963). Information theory and statistical mechanics. In K. W. Ford (Ed.) *Statistical Physics.* New York: W. A. Benjamin, 181-218.

Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Technical Journal* 27: 379-423.

Shore, J. E., R. W. Johnson. (1980). Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, *IEEE Transactions on Information Theory* 26(1): 26-37.

Skilling, J. (1988). The axioms of maximum entropy. In J. Skilling (Ed.) *Maximum Entropy and Bayesian Methods in Science and Engineering.* Dordrecht: Kluwer, 173-187.

Soofi, E. S. (1992). A generalizable formulation of conditional logit with diagnostics, *Journal of the American Statistical Association* 87(Sept.): 812-816.

Takayama, A. (1985). *Mathematical Economics.* New York: Cambridge University Press.

CHAPTER 2

MONTE CARLO EXPERIMENTS COMPARING THE MAXIMUM ENTROPY ESTIMATOR TO TRADITIONAL ESTIMATORS IN DISCRETE CHOICE MODELS

2.1 Introduction

In this chapter, we use maximum entropy (ME) and generalized maximum entropy (GME) to estimate discrete choice models. We carry out several Monte Carlo experiments comparing the ME and GME estimators to traditional estimators: probit, logit, linear discriminant analysis, and multinomial logit. We compare the estimators based on their ability to predict outcomes, the mean squared error (MSE) of the parameter estimates and the marginal effects of the parameters, the MSE of the predicted latent values, and the MSE of the predicted probabilities. We use the ME and GME estimators developed by Golan, Judge, and Perloff (1996). Golan et al. examine ME estimation in the context of sampling experiments using artificial data. They include one empirical example involving occupational choice data with a high degree of collinearity. Adkins (1997) carries out Monte Carlo experiments with a slightly different GME formulation using real data. We examine ME and GME estimation in the context of discrete choice models using real data as a basis for Monte Carlo experiments.

We carry out binary choice experiments involving the choice of whether or not to attend post-secondary education (PSE) as well as a credit scoring model, which we discuss in section 2.6, for used car loans. In addition, we carry out multinomial choice experiments with the PSE data. We discuss linear discriminant analysis in section 2.2, ME estimation in binary choice models in section 2.3 and GME estimation in binary choice models in section 2.4. We discuss prediction in multinomial choice models in section 2.5. We review the credit scoring literature in section 2.6, describe our sampling experiments in section 2.7 and present results in section 2.8.

2.2 Digression on Discriminant Analysis

In linear discriminant analysis (DA) we divide the sample into two groups, those who default and those who repay their loans. The personal and financial characteristics of the individuals are contained in a matrix, x. We define

$$y = \begin{cases} 1 & \text{for default} \\ 0 & \text{for repayment} \end{cases}$$

and denote the values of x as x_1 for those who default and x_0 for those who repay. Linear DA assumes that x_1 and x_0 are distributed as multivariate normal with different means but equal covariance matrices. Since the true means and covariance matrix are unknown, we estimate them following Maddala (1983)

$$\hat{\mu}_1 = \frac{1}{n_1}\sum_i x_{1i} = \bar{x}_1 \tag{2.1}$$

$$\hat{\mu}_0 = \frac{1}{n_0}\sum_i x_{0i} = \bar{x}_0 \tag{2.2}$$

$$S = \frac{1}{(n_1 + n_0 - 2)}\left[\sum_i (x_{1i} - \bar{x}_1)(x_{1i} - \bar{x}_1)' + \sum_i (x_{0i} - \bar{x}_0)(x_{0i} - \bar{x}_0)'\right], \tag{2.3}$$

where S is the estimated covariance matrix for x, n_1 is the number of observations where $y = 1$, and n_0 is the number of observations where $y = 0$. Fisher (1936) argues that maximizing the variance between the two groups relative to the overall variance of y provides the best discrimination between the two groups. Therefore, we maximize the ratio

$$\phi = \frac{[\beta'(\bar{x}_1 - \bar{x}_0)]^2}{\beta' S \beta}, \tag{2.4}$$

where β is a $K \times 1$ vector of unknown parameters, $(\bar{x}_1 - \bar{x}_0)$ is a $K \times 1$ vector of differences in the means of the independent variables between the two groups and S is the $K \times K$ estimated covariance matrix for x. The numerator is the variance between the two groups and the denominator is the variance of y. The DA parameter estimates are given by

$$\hat{\beta} = S^{-1}(\bar{x}_1 - \bar{x}_0) \tag{2.5}$$

$$\hat{\alpha} = \ln\left(\frac{n_1}{n_0}\right) - \frac{\hat{\beta}'}{2}(\bar{x}_1 + \bar{x}_0), \tag{2.6}$$

where $\hat{\alpha}$ is the intercept and $\hat{\beta}$ is the $K \times 1$ vector of parameter estimates for the other explanatory variables.

Efron (1975) and Maddala (1983) both show that, given the assumption that x_1 and x_0 are distributed multivariate normal, the probability that an individual with a given set of characteristics will choose $y = 1$ is calculated as

$$P(y_i = 1|x_i) = \frac{\exp(\alpha + \beta' x_i)}{1 + \exp(\alpha + \beta' x_i)}, \tag{2.7}$$

which is a logistic cumulative distribution function. Thus, while the coefficient estimates are different the predicted probabilities are calculated in the same manner for DA, logit, and ME.

2.3 Maximum Entropy Estimation in Binary Choice Models

Denzau, Gibbons, and Greenberg (1989), Soofi (1992), and Golan, Judge, and Perloff (1996) discuss ME estimation of the unknown parameters in a discrete multinomial choice problem. The multinomial choice model yields a natural application of the ME methodology since the unknown parameters are in the form of probabilities. The multinomial choice model with only two alternatives is the binary choice model, which we present here. However, we carry out Monte Carlo experiments for both binary and multinomial choice models and compare the ME estimator to traditional estimators. We give the ME solutions to both the binary and multinomial choice models in Appendix A.

In Jaynes' dice problem, we use ME to estimate the unknown probabilities given a set of known support points and the observed average from a large number of prior rolls. In binary choice models, each of N individuals must select one of two alternatives. We observe a set of binary random variables, y_{ij}, which equal one if individual i selects alternative j ($j = 1, 2$), and zero otherwise. Thus, in the binary choice problem we have a known set of support points for y, 0 and 1, and a known prior for the unknown probabilities, the proportion of individuals selecting each alternative. We estimate the

probability that y is equal to each of the support points. Since each individual must select one of the alternatives, $y_{i1} + y_{i2} = 1$ for all i individuals. We are interested in $p_{ij} = \Pr(y_{ij} = 1)$, or the probability that individual i selects alternative j. We know that $p_{i1} + p_{i2} = 1$ for all individuals and assume that the unknown probabilities are related to a set of explanatory variables. We specify this relationship through a data constraint in the entropy maximization problem. The model is written as

$$y_{ij} = G(x_i'\beta_j) + e_{ij} = p_{ij} + e_{ij} \tag{2.8}$$

or in matrix notation as

$$y = p + e, \tag{2.9}$$

where $y = [y_1' \quad y_2']'$ is a $2N \times 1$ vector of known choices, $p = [p_1' \quad p_2']'$ is a $2N \times 1$ vector of unknown probabilities, $e = [e_1' \quad e_2']'$ is a $2N \times 1$ vector of unknown random errors, x_i is a $K \times 1$ vector of known individual specific characteristics or attributes, and β_j is a $K \times 1$ vector of unknown parameters; y_1 is the $N \times 1$ vector of binary variables observed for alternative $j = 1$, p_1 is the $N \times 1$ vector of probabilities for choosing alternative $j = 1$ and e_1 is the associated vector of errors for these probabilities (y_2, p_2, and e_2 are similarly defined for alternative $j = 2$). The probit model results if we assume the errors have a standard normal distribution while the logit model results if we assume the errors are independent and identically distributed with a Weibull distribution. There are $2N$ observations since we estimate the probability that each individual will choose either of the two alternatives. That is, we will estimate p_{i1} and p_{i2}, subject to the constraint $p_{i1} + p_{i2} = 1$. In the ME framework, we impose the prior information that the probabilities must sum to one for each individual.

Denzau, Gibbons, and Greenberg (1989) and Soofi (1992) consider the entropy of the unknown probabilities in the discrete choice model, yielding the traditional ME estimator. The ME model does not include the error term in the constraints of the entropy maximization problem. However, Golan, Judge, and Perloff (1996) argue that there is also uncertainty or entropy associated with the unknown errors in the model. Golan et al. develop the generalized maximum entropy (GME) estimator, which

includes the unknown and unobservable errors in both the entropy function and its constraints. We examine both the ME and the GME estimators.

ME estimation yields the most uniform probability distribution compatible with the prior information. Therefore, in order to solve equation (2.8) for the unknown probabilities using ME, we specify a set of constraints compatible with the observed data. Soofi (1992) and Golan, Judge, and Perloff (1996) argue that the data constraint on the model should be specified to preserve the observed sums of the sample characteristics. If we include a constant term, this data constraint requires that the mean of the predicted probabilities is equal to the percentage of observations for which $y_i = 1$. This is analogous to the estimated probabilities in the dice problem being chosen such that the expected value of a roll is equal to the observed average from a large number of prior rolls. This constraint is written as

$$\sum_{i=1}^{N} \hat{p}_{ij} x_{ik} = \sum_{i=1}^{N} y_{ij} x_{ik}, \qquad \forall\, k, j$$

which is also the first-order condition for the logit model. Finally, the sum of the estimated probabilities must equal one for each individual.

We estimate the ME probabilities by solving the constrained optimization problem

$$\max H(p) = -p' \ln(p) \tag{2.10}$$

subject to

$$(I_2 \otimes X')y = (I_2 \otimes X')p \tag{2.11}$$

$$\begin{bmatrix} I_N & I_N \end{bmatrix} p = i_N \tag{2.12}$$

where $\otimes$ is the Kronecker product, I_N represents an $N \times N$ identity matrix, i_N is an $N \times 1$ vector of ones and X is an $N \times K$ matrix of individual characteristics. The first constraint is the data constraint, which preserves the observed sums of the sample characteristics. The second constraint is the additivity constraint that requires the predicted probabilities to sum to one for each observation. The Lagrangian for the ME binary choice problem is

$$\mathcal{L} = -p' \ln(p) + \lambda' \left[(I_2 \otimes X')p - (I_2 \otimes X')y \right] + \gamma' \left[i_N - \begin{bmatrix} I_N & I_N \end{bmatrix} p \right],$$

where $\lambda = [\lambda_1' \quad \lambda_2']'$ is a $2K \times 1$ vector of parameter estimates and γ is an $N \times 1$ vector of Lagrange multipliers for the additivity constraint. The FOC's for the ME binary choice problem are

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{2N} - \ln(p) + (I_2 \otimes X)\lambda - [I_N \quad I_N]'\gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = (I_2 \otimes X')p - (I_2 \otimes X')y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_N - [I_N \quad I_N]p = 0.$$

Solving the $2N$ FOC's for the unknown p's yields

$$\hat{p} = \exp\left[(I_2 \otimes X)\hat{\lambda}\right]\exp\left[-[I_N \quad I_N]'\gamma - i_{2N}\right],$$

which implies that

$$\hat{p}_{ij} = \exp(x_i'\hat{\lambda}_j)\exp(-\hat{\gamma}_i - 1).$$

Since additivity requires that $p_{i1} + p_{i2} = 1$ and $\exp(-\hat{\gamma}_i - 1)$ is constant for a given individual, the optimal ME probabilities can be rewritten as

$$\hat{p}_{ij} = \frac{\exp(x_i'\hat{\lambda}_j)}{\sum_{j=1}^{2}\exp(x_i'\hat{\lambda}_j)} = \frac{\exp(x_i'\hat{\lambda}_j)}{\Omega_i(\hat{\lambda})},$$

where $\Omega_i(\hat{\lambda})$ represents the partition function for the i^{th} individual. Thus, the optimal ME probabilities are a function of the explanatory variables and the associated vector of parameters, λ. Replacing λ with the more familiar parameter vector β and normalizing the K-vector $\beta_2 = 0$ yields the ME predicted probabilities

$$\hat{p}_{ij} = \frac{\exp(x_i'\hat{\beta}_j)}{1 + \exp(x_i'\hat{\beta}_1)}, \quad j = 1,2 \tag{2.13}$$

which are identical to the predicted probabilities obtained from logit estimation. Note that we do not estimate the parameters directly in ME. We only obtain the parameter estimates, $\hat{\beta}$, indirectly as the Lagrange multipliers for the data constraint.

Soofi (1992) and Golan, Judge, and Perloff (1996) argue that the ME solution does not make any parametric assumptions about the error distribution. However, imposing the data constraint (2.11) assumes that the errors sum to zero, even in small samples. Equation (2.11) requires that the number of individuals choosing each alternative is equal to the sum of the predicted probabilities for choosing that alternative. We know that for each individual and alternative that

$$y_{ij} = p_{ij} + e_{ij},$$

but the probabilities and random errors are unknown. Summing over all individuals and alternatives yields

$$(I_2 \otimes X')y = (I_2 \otimes X')p + (I_2 \otimes X')e, \tag{2.14}$$

which is simply equation (2.11) with the error term included. In the ME formulation we assume that $(I_2 \otimes X')e = 0$, which implies

$$\mathrm{E}[(I_2 \otimes X')y] = (I_2 \otimes X')p. \tag{2.15}$$

Since the true probabilities are unknown, we must assume that the errors sum to zero if we are to impose constraint (2.11). In addition, the data constraint that we impose in the ME formulation is the first-order condition for the logit model (Maddala 1983, p. 26). By imposing the logit first-order condition we are implicitly assuming the same error distribution as the logit model. Finally, we assume independence between X and e.

Taking the second derivative of the Lagrangian with respect to p shows that the Hessian for the ME binary choice problem is negative definite, and thus ensures a unique global solution to the entropy maximization problem. Golan, Judge, and Perloff (1996) obtain the information matrix for the p's and, using a result from Lehmann (1983), obtain a ME information matrix for the β's which is identical to the maximum likelihood (ML) information matrix for the multinomial choice problem. Thus, the maximum likelihood logit and the ME solutions are identical and have the same asymptotic properties.

Both linear and nonlinear restrictions on the parameters can be included as additional constraints in the ME problem. However, we can also impose parameter restrictions in both logit and probit using constrained optimization. In our Monte Carlo experiments in this chapter, we compare the effects of adding restrictions in both ME and traditional maximum likelihood methods (logit and probit).

As discussed in Chapter 1, the ME problem may also be set up as an unconstrained optimization problem by substituting the optimal probabilities into the original Lagrangian. For $\lambda \in \Re$, let $\hat{p}(\lambda)$ represent the functional form of the optimal ME probabilities, equation (2.13). Since the optimal probabilities satisfy the additivity constraint (2.12), this term drops out of the original Langrangian and the concentrated Lagrangian is

$$M(\lambda) = -\hat{p}'(\lambda)\ln(\hat{p}(\lambda)) + \lambda'[(I_2 \otimes X')\hat{p}(\lambda) - (I_2 \otimes X')y]$$

$$= -\hat{p}'(\lambda)\ln\left[\frac{\exp(x_i'\lambda_j)}{\sum_j \exp(x_i'\lambda_j)}\right] + \lambda'[(I_2 \otimes X')\hat{p}(\lambda) - (I_2 \otimes X')y]$$

$$= -\hat{p}'(\lambda)(I_2 \otimes X)\lambda + \sum_i \ln[\Omega_i(\lambda)] + \hat{p}'(\lambda)(I_2 \otimes X)\lambda - y'(I_2 \otimes X)\lambda .$$

Replacing λ with β and normalizing the K-vector $\beta_2 = 0$, yields

$$M(\lambda) = -y'X\lambda + \sum_i \ln[1 + \exp(x_i'\lambda)], \tag{2.16}$$

which we minimize to obtain $\hat{\beta}$. We then substitute $\hat{\beta}$ into (2.13) to obtain the optimal ME probabilities. This is an unconstrained optimization problem that is solved for the K unknown β's rather than the $2N$ unknown probabilities. Thus, the dual problem is computationally simpler in most cases.

2.4 Generalized Maximum Entropy Estimation in Binary Choice Models

Golan, Judge, and Perloff (1996) develop an entropy method to estimate both the unknown probabilities and the unknown errors. They argue that jointly estimating the probabilities and errors is more efficient in small samples than estimating only the probabilities. Golan et al., measure efficiency in terms of the mean squared error of the estimated parameter vector, $\hat{\beta}$. They find that the GME

estimator has lower mean squared error, $MSE(\hat{\beta})$, than logit in sampling experiments with artificial data and in a sampling experiment using real data on occupational choice.

Recall that ME allows us to estimate unknown discrete probability distributions. However, the unknown errors in the binary choice model are not in the form of probabilities. Therefore, Golan, Judge, and Perloff (1996) reparameterize the error term in equation (2.9) so that the random errors have the properties of probabilities. They refer to the reparameterization as the generalized maximum entropy (GME) formulation. In GME, we must assume that the unknown errors may be bounded *a priori*. The error bounds are based on prior information or economic theory. In the binary choice problem, (2.9), we know the errors fall in the interval $[-1,1]$ since y can only take two values, 0 or 1, and the predicted probabilities must be between 0 and 1. Therefore, for each random variable, e_{ij}, there exists $w \in [0,1]$ such that

$$e_{ij} = (-1)w + (1)(1-w) = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} w \\ 1-w \end{bmatrix}.$$

In addition, we know that $y = p + e$ must be either 0 or 1.

Golan, Judge, and Perloff (1996) define a set of $M \geq 2$ support points, symmetric about zero, which bound the errors and w_{ij} is the associated $M \times 1$ vector of weights on these points. The error vector for each choice is expressed as

$$e = Vw = \begin{bmatrix} v'_{1j} & 0 & . & 0 \\ 0 & v'_{2j} & . & 0 \\ . & . & . & . \\ 0 & 0 & . & v'_{Nj} \end{bmatrix} \cdot \begin{bmatrix} w_{1j} \\ w_{2j} \\ . \\ w_{Nj} \end{bmatrix}, \qquad j = 1,2 \tag{2.17}$$

where e is a $2N \times 1$ vector of random errors, V is a $2N \times 2NM$ matrix of support points and w is a $2NM \times 1$ vector of unknown weights such that $w_{ij} > 0$ and $w'_{ij} i_M = 1$ for all i and j. We generally choose $M = 3$ support points, which are symmetric about zero, and use the same set of support points for each e_{ij}. Thus, for the binary choice GME formulation each v'_{ij} in equation (2.17) is the row vector

$\begin{pmatrix}-1 & 0 & 1\end{pmatrix}$. However, the definition of support points may vary depending on the researcher's prior information. By defining $e = Vw$, Golan, Judge, and Perloff (1996) rewrite the binary choice equation (2.9) in matrix notation as

$$y = p + Vw, \tag{2.18}$$

where y and V are known and we use GME to estimate the unknown p and w vectors.

Shannon (1948) and Jaynes (1957a,b) show that entropy is additive for independent sources of uncertainty. Golan, Judge, and Perloff (1996) assume independence between the unknown probabilities and errors in the multinomial choice model and obtain the GME probabilities by maximizing

$$H(p,w) = -p'\ln(p) - w'\ln(w) \tag{2.19}$$

subject to the constraints

$$(I_2 \otimes X')y = (I_2 \otimes X')p + (I_2 \otimes X')Vw \tag{2.20}$$

$$\begin{bmatrix} I_N & I_N \end{bmatrix} p = i_N \tag{2.21}$$

$$(I_{2N} \otimes i'_M)w = i_{2N}. \tag{2.22}$$

Equation (2.20) is the data constraint and equations (2.21) and (2.22) are the additivity constraints, which require that the probabilities must sum to one. We incorporate the prior information that the errors must fall between -1 and 1 to estimate the unknown errors for each individual. The Lagrangian for the GME binary choice problem is

$$\mathcal{L} = -p'\ln(p) - w'\ln(w) + \lambda'\left[(I_2 \otimes X')p + (I_2 \otimes X')Vw - (I_2 \otimes X')y\right]$$
$$+\gamma'\left[i_N - \begin{bmatrix} I_N & I_N \end{bmatrix} p\right] + \tau'\left[i_{2N} - (I_{2N} \otimes i'_M)w\right],$$

where $\lambda = \begin{bmatrix} \lambda'_1 & \lambda'_2 \end{bmatrix}'$ is a $2K \times 1$ vector of parameter estimates, γ is an $N \times 1$ vector of Lagrange multipliers for the additivity constraint on the unknown probabilities, and $\tau = \begin{bmatrix} \tau'_1 & \tau'_2 \end{bmatrix}'$ is a $2N \times 1$ vector of Lagrange multipliers for the addivity constraint on the unknown error weights. The FOC's for the GME binary choice problem are

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{2N} - \ln(p) + (I_2 \otimes X)\lambda - [I_N \quad I_N]'\gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = (I_2 \otimes X')p + (I_2 \otimes X')Vw - (I_2 \otimes X')y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_N - [I_N \quad I_N]p = 0$$

$$\frac{\partial \mathcal{L}}{\partial w} = -i_{2NM} - \ln(w) + V'(I_2 \otimes X)\lambda - (I_{2N} \otimes i_M)\tau = 0$$

$$\frac{\partial \mathcal{L}}{\partial \tau} = i_{2N} - (I_{2N} \otimes i'_M)w = 0.$$

The $2N$ FOC's for the p's are the same as in the ME formulation, so we again obtain the solution (2.13) for the predicted probabilities. Solving the $2NM$ FOC's for the w's yields

$$w = \exp\left[V'(I_2 \otimes X)\hat{\lambda}\right]\exp\left[-(I_{2N} \otimes i_M)\tau - i_{2NM}\right],$$

which implies that

$$w_{ijm} = \exp(x_i'\hat{\lambda}_j v_{ijm})\exp\left[-\tau_{ij} - 1\right].$$

For the GME problem, the partition function varies with each alternative ($j = 1,2$) as well as with each individual ($i = 1,\dots,N$). Because additivity requires that $\sum_{m=1}^{M} w_{ijm} = 1$, and $\exp(-\tau_{ij} - 1)$ is constant for a given i and j as we vary m, the optimal GME error weights may be rewritten as

$$w_{ijm} = \frac{\exp(x_i'\hat{\lambda}_j v_{ijm})}{\sum_{m=1}^{M}\exp(x_i'\hat{\lambda}_j v_{ijm})} = \frac{\exp(x_i'\hat{\lambda}_j v_{ijm})}{\Psi_{ij}(\hat{\lambda}_j)},$$

where Ψ_{ij} is the partition function for a given i and j. We again replace λ with β and normalize $\beta_2 = 0$ to obtain the optimal GME error weights

$$\hat{w}_{ijm} = \frac{\exp(x_i'\hat{\beta}_j v_{ijm})}{\sum_{m=1}^{M}\exp(x_i'\hat{\beta}_j v_{ijm})}. \tag{2.23}$$

The GME error estimates are obtained through the equation

$$\hat{e} = V\hat{w}. \tag{2.24}$$

The normalization $\beta_2 = 0$ implies that $\hat{e}_{i2} = 0$ for all observations, if the error support is symmetric about zero, as we have assumed. [Proof: $\beta_2 = 0 \Rightarrow \hat{w}_{i2m} = 1/M$ for all i, m. Since the support points v_{i2m} are symmetric about zero, $\hat{e}_{i2} = \sum_{m=1}^{M} v_{i2m}\hat{w}_{i2m} = \sum_{m=1}^{M}(v_{i2m}/M) = 0$ for all i observations.] The estimated value of each error term in GME is a function of the personal characteristics, the estimated parameters, and the bounds placed on the errors *a priori*.

The GME solution can also be obtained using the unconstrained dual formulation. For $\lambda \in \Re$, let $\hat{p}(\lambda)$ represent the functional form of the optimal GME probabilities, (2.13), and let $\hat{w}(\lambda)$ represent the functional form of the optimal weights placed on the error support, (2.23). Since the additivity constraints, (2.21) and (2.22), are satisfied by the optimal probabilities these terms drop out of the original Lagrangian and the concentrated Lagrangian is

$$\begin{aligned} M(\lambda) &= -\hat{p}'(\lambda)\ln(\hat{p}'(\lambda)) - \hat{w}'(\lambda)\ln(\hat{w}'(\lambda)) \\ &\quad + \lambda'\left[(I_2 \otimes X')\hat{p}(\lambda) + (I_2 \otimes X')V\hat{w}(\lambda) - (I_2 \otimes X')y\right] \\ &= -\hat{p}'(\lambda)(I_2 \otimes X)\lambda + \sum_i \ln\left[\Omega_i(\lambda)\right] - \hat{w}'(\lambda)V'(I_2 \otimes X)\lambda \\ &\quad + \sum_i\sum_j \ln\left[\Psi_{ij}(\lambda_j)\right] + \hat{p}'(\lambda)(I_2 \otimes X)\lambda + \hat{w}'(\lambda)V'(I_2 \otimes X)\lambda - y'(I_2 \otimes X)\lambda. \end{aligned}$$

Replacing λ with β and normalizing $\beta_2 = 0$ yields

$$M(\lambda) = -y'X\lambda + \sum_i \ln\left[\Omega_i(\lambda)\right] + \sum_i\sum_j \ln\left[\Psi_{ij}(\lambda_j)\right], \tag{2.25}$$

which is minimized to obtain $\hat{\beta}_1$. The dual formulation is particularly useful for the GME problem since we must estimate $2N$ unknown probabilities and $2NM$ unknown error weights. The dual formulation allows us to obtain the same results while only estimating the K unknown parameters.

The GME formulation developed by Golan, Judge, and Perloff (1996) is an innovative way to include the error term in the entropy maximization problem. The GME formulation allows us to

incorporate prior sample information about the error values, as well as the parameter values, in the form of constraints on the model. For example, in a binary choice model, we know that the error values must fall in the interval [-1,1]. Because we specify this as a constraint, the GME errors will fall in this range. However, we also know that $y = p + e$ must equal 0 or 1. In GME, $\hat{y} = \hat{p} + \hat{e}$ and we should constrain $0 \le \hat{y} \le 1$ since the true probability must lie between 0 and 1. For example, if the predicted probability that $y = 1$ is $\hat{p} = 0.75$ we know that $-0.75 \le \hat{e} \le 0.25$. Golan, Judge, and Perloff (1996) do not impose this constraint in their GME formulation. We develop an alternative GME formulation below that imposes $0 \le \hat{y} \le 1$.

To add the constraint that $0 \le \hat{y} \le 1$ in GME, we maximize the entropy function, equation (2.19), subject to (2.20), (2.21), and (2.22) and the additional constraints

$$p_1 + Vw_1 \le 1 \tag{2.26}$$

$$p_1 + Vw_1 \ge 0. \tag{2.27}$$

We only need to place these constraints on $p_1 + e_1$. Since $e_2 = 0$ due to the normalization $\beta_2 = 0$, $p_2 + e_2$ is already constrained to the [0,1] interval. The estimated probabilities for our alternative GME formulation are

$$\hat{p}_{ij} = \frac{\exp(x_i'\hat{\beta}_j - \hat{\theta}_{1ij} + \hat{\theta}_{2ij})}{1 + \exp(x_i'\hat{\beta}_1 - \hat{\theta}_{1i1} + \hat{\theta}_{2i1})} \qquad j = 1,2, \text{and} \tag{2.28}$$

$$\hat{w}_{ijm} = \frac{\exp(x_i'\hat{\beta}_j v_{ijm} - \hat{\theta}_{1ij} v_{ijm} + \hat{\theta}_{2ij} v_{ijm})}{\sum_{m=1}^{M} \exp(x_i'\hat{\beta}_j v_{ijm} - \hat{\theta}_{1ij} v_{ijm} + \hat{\theta}_{2ij} v_{ijm})}, \tag{2.29}$$

where θ_1 is the Lagrange multiplier for constraint (2.26), θ_2 is the Lagrange multiplier for constraint (2.27), $\beta_2 = 0$, $\theta_{12} = 0$, and $\theta_{22} = 0$. The Kuhn-Tucker conditions require that if constraints (2.26) and (2.27) are not binding then $\theta_1 = \theta_2 = 0$. In this case, the optimal GME probabilities and error weights are the same as those obtained by Golan, Judge, and Perloff (1996).

2.5 Predicting Multinomial Choices Using GME

Golan, Judge, and Perloff (1996) assume independence between the unknown and unobservable p and w, so that we may add their entropies. Examination of the GME solutions (2.13) and (2.23) shows that both the unknown probabilities and errors depend on the explanatory variables and the parameter estimates. We assume independence between p and e yet we estimate them jointly using the sample information. The joint estimation of $\hat{p}$ and $\hat{e}$ leads to two possible alternatives for predicting outcomes for multinomial choices based on the GME estimates.

Golan, Judge, and Perloff (1996) use $\Pr(y=1)=\hat{p}_1$ to predict outcomes. In binary choice problems, we generally predict an individual to choose alternative $j=1$ if $\hat{p}_1 > 0.5$. From the GME predicted probabilities, (2.13), we have $\hat{p}_1 > 0.5$ iff $\hat{y}_1^* = x_i'\hat{\beta}_1 > 0$, where y^* is a latent variable that determines whether or not y equals 0 or 1. Additionally, $\hat{p}_1 > 0.5$ implies that $\hat{e}_1 = V\hat{w}_1 > 0$ while $\hat{p}_1 < 0.5$ implies that $\hat{e}_1 < 0$. [Proof: $\hat{e}_{i1} = \sum_{m=1}^{M} v_{i1m}\hat{w}_{i1m}$; if $x_i'\hat{\beta}_1 > 0$ then $\hat{w}_{i11} < \hat{w}_{i12} < \ldots < \hat{w}_{i1M}$ and $\hat{e}_{i1} > 0$. Conversely, if $x_i'\hat{\beta}_1 < 0$ then $\hat{w}_{i11} > \hat{w}_{i12} > \ldots > \hat{w}_{i1M}$ and $\hat{e}_{i1} < 0$.] Finally, $\hat{e}_1$ is larger in absolute value the closer $\hat{p}$ is to an extreme value, 0 or 1, or the wider the error bounds (since $\hat{e}_{i1}$ increases as either v_{i1m} or $\hat{w}_{i1m}$ increases, and $\hat{w}_{i1m}$ increases as $x_i'\hat{\beta}_1$ moves away from zero in either direction). The effect of including the error term in the GME formulation is to make the estimated probabilities more uniform than the ME-logit predicted probabilities (each estimated probability is shrunk towards 0.5 with GME). GME estimation will only differ in predicting choices from ME-logit when the ME-logit probability is close to 0.5.

An alternative prediction method is to examine $\hat{y}$, the estimated value of y itself. In ME-logit, $\hat{y} = \hat{p}$ which is between 0 and 1. Additionally, the proportion of 1's in the sample is equal to the mean of $\hat{p}$. In GME, $\hat{y} = \hat{p} + \hat{e}$ and the proportion of 1's in the sample is equal to the mean of $\hat{y}$. We predict an individual to choose alternative 1 if $\hat{y}_1$ is greater than some threshold (usually 0.5).

Since $\hat{e}_1 > 0$ for $\hat{p}_1 > 0.5$, $\hat{y}_1$ will be greater than $\hat{p}_1$ for $\hat{p}_1 > 0.5$ and less than $\hat{p}_1$ for $\hat{p}_1 < 0.5$. For our Monte Carlo sampling experiments we use $\hat{y} = \hat{p} + \hat{e}$ to predict outcomes and to calculate the MSE of the predicted probabilities. Since GME shrinks $\hat{p}$ toward 0.5 by placing weight on the error term, $\hat{y}$ provides a better comparison to the true probabilities than $\hat{p}$ does. We compare estimation results using logit, probit, DA, ME, and GME.

2.6 Credit Scoring Models

Credit scoring models identify individuals who are most at risk to default on a loan. Individuals will either default or repay a loan, making credit scoring a binary choice problem. Therefore, we estimate the model using the binary choice ME and GME models discussed in this chapter. We compare the ME estimates to traditional binary choice estimates including probit, logit, and linear discriminant analysis (DA), which we describe in section 2.2.

The earliest credit scoring papers use linear DA to classify individuals as either likely to repay or likely to default on a loan. Examples of credit scoring models using DA are found in Myers and Forgy (1963) and Altman (1968). Recent credit scoring models use discrete choice econometric methods, either probit or logit, to predict whether or not an individual will repay a loan. Steenackers and Goovaerts (1989) and Lawrence, Smith, and Rhoades (1992) estimate credit scoring models using logit. Knapp and Seaks (1992) use a probit model to estimate the probability of default on federally guaranteed student loans.

Boyes, Hoffman, and Low (1989) estimate a credit scoring model using a censored probit model developed by Manski and Lerman (1977). The censored model accounts for the fact that some individuals apply for and do not receive loans. In addition to data on individuals receiving loans, Boyes et al. have data on individuals who are denied loans and the censored probit model incorporates this information. Estimation of credit scoring models without accounting for this censoring problem leads to biased and inconsistent estimators. Golan, Judge, and Perloff (1997) discuss the censoring problem in the context of ME estimation. However, a ME estimation procedure has not yet been developed for

truncated data, in which we do not observe characteristics of individuals who are denied loans. We discuss both censoring and truncation issues in Chapter 5.

2.7 Monte Carlo Design

We carry out several Monte Carlo experiments and compare the ME and GME estimators to traditional estimators: probit, logit, and DA. We carry out two types of sampling experiments: the first type only evaluates the percentage of correct predictions while the second type also examines the MSE of the predicted probabilities and latent values as well as the MSE of the parameter estimates and the marginal effects. In this section, we describe the data, describe the experimental design, and then discuss variations in the experimental design.

2.7.1 Data

Used Cars Data. The used cars data set comes from Martin and Hill (1999). The used cars data consist of 151,659 used car loans completed between 1 January 1990 and 31 December 1994. A completed loan is one which was paid off, either on schedule or early, or which went into default. There are no active loans in our sample. In the full sample, 132,705 loans (87.5%) were successfully paid off while 18,954 loans (12.5%) went into default. Thus, the sample is unbalanced between loans that were paid off and loans that went into default. The unbalanced sample affects prediction, which is discussed below. We include the following explanatory variables as well as a constant in the model:

- Cdcmi = combined monthly income,
- Cdfpay = monthly loan payment,
- Cdtdp = down payment percentage,
- Age = age of buyer,
- Amtfin = amount financed,
- Netrate = interest rate,
- Home = a dummy variable equal to one if the individual is a homeowner,
- Cosign = a dummy variable equal to one if there is a cosigner for the loan.

Table 2.1 gives summary statistics for the independent variables for the full used cars data set. The sample coefficient of variation is defined as $CV_x = s(x) / \bar{x}$, where $s(x)$ is the sample standard

deviation of x and $\bar{x}$ is the sample mean of x. The dependent variable is equal to one for defaults and zero for repayments. Belsley, Kuh, and Welsch (1980, pp. 100-104) define the condition number of the $X'X$ matrix, denoted $\kappa(X'X)$, to be $\sqrt{\lambda_1 / \lambda_K}$ where λ_1 is the largest characteristic root and λ_K is the smallest characteristic root and the regressors have been normalized to unit length. They conclude that for condition numbers greater than about 30 multicollinearity is a problem in the model. We find the maximum condition number for the used cars data to be $\kappa(X'X) = 15.5$, which indicates that multicollinearity is not a problem in the used cars data.

Table 2.1 Summary Statistics for Used Cars Data (*N*=151,659 Observations)

Variable	Mean	Min.	Max.	Standard Deviation	Coefficient of Variation
Cdcmi	1376.46	0	9999	1414.83	1.03
Cdfpay	185.62	1.24	2288.30	94.31	0.51
Cdtdp	11.25	0	99	15.60	1.39
Age	36.87	0	99	14.48	0.39
Amtfin	4091.54	17.07	40,000.00	3039.70	0.74
Netrate	17.68	0	90.14	10.01	0.57
Home	0.19	0	1	0.39	2.05
Cosign	0.07	0	1	0.26	3.71

PSE Data. The post-secondary education (PSE) data, from the National Center for Educational Statistics (1996), consist of 9,450 high school graduates. In the data set, 7,317 (77.4%) students attended some type of post-secondary education and 2,133 (22.6%) students did not attend any post-secondary education. We include the following explanatory variables as well as a constant in the model:

- Black = a dummy variable equal to one if the individual is African-American,
- Othrace = a dummy variable equal to one if the individual is neither African-American nor Caucasian,
- Catholic = a dummy variable equal to one if the individual attended a catholic high school,
- Income1 = a dummy variable equal to one if hh income is between $20,000-$50,000,
- Income2 = a dummy variable equal to one if hh income is between $50,000-$75,000,
- Income3 = a dummy variable equal to one if hh income is greater than $75,000,
- GPA = high school GPA,

- Parhs = a dummy variable equal to one if a parent has a high school degree only,
- Parcoll = a dummy variable equal to one if a parent has a college degree.

The PSE data has a maximum condition number $\kappa(X'X) = 11.8$, so we conclude that multicollinearity is not a problem with the PSE data. Table 2.2 gives summary statistics for the independent variables in the PSE data set.

Table 2.2 Summary Statistics for PSE Data (N=9,450 Observations)

Variable	Mean	Min.	Max.	Standard Deviation	Coefficient of Variation
Black	0.10	0	1	0.30	3.02
Othrace	0.19	0	1	0.39	2.07
Catholic	0.07	0	1	0.26	3.55
Income1	0.44	0	1	0.50	1.13
Income2	0.21	0	1	0.40	1.97
Income3	0.16	0	1	0.36	2.32
GPA	1.95	0.02	3.65	0.72	0.37
Parhs	0.60	0	1	0.49	0.82
Parcoll	0.33	0	1	0.47	1.43

We also carry out multinomial choice experiments using the PSE data. In the multinomial choice experiments we define the dependent variable as y=0 for no PSE, y=1 for a 2-year institution, y=2 for a 4-year institution, and y=3 for other PSE. We carry out the multinomial choice experiments in the same manner as the binary choice experiments.

2.7.2 Prediction Experiments

In the prediction experiments, we draw random samples that reflect the true proportion of defaults and repayments (or PSE and non-PSE) in the data as well as samples with an equal number of observations for each alternative. We estimate the model with each random sample and obtain the percentage of correct predictions. We vary the sample size and perform the experiments both with and without inequality restrictions placed on the parameters. Therefore, in the prediction experiments, we have 12 experimental designs:

1) A random sample of 20,000 observations (5,000 for PSE data).
2) A random sample of 5,000 observations (2,000 for PSE data).
3) A random sample of 2,000 observations (1,000 for PSE data).

4) A balanced sample with 10,000 each repayments and defaults (2,500 PSE and 2,500 no PSE).

5) A balanced sample with 2,500 each repayments and defaults (1,000 PSE and 1,000 no PSE).

6) A balanced sample with 1,000 each repayments and defaults (500 PSE and 500 no PSE).

7) A random sample of 20,000 observations with inequality restrictions placed on the parameters (5,000 for PSE data).

8) A random sample of 5,000 observations with inequality restrictions placed on the parameters (2,000 for PSE data).

9) A random sample of 2,000 observations with inequality restrictions placed on the parameters (1,000 for PSE data).

10) A balanced sample of 20,000 observations with inequality restrictions placed on the parameters (5,000 for PSE data).

11) A balanced sample of 5,000 observations with inequality restrictions placed on the parameters (2,000 for PSE data).

12) A balanced sample of 2,000 observations with inequality restrictions placed on the parameters (1,000 for PSE data).

We refer to these experiments as prediction 1-12. Table 2.3 summarizes the dimensions for the binary choice prediction experiments.

Table 2.3 Dimensions of Monte Carlo Prediction Experiments

Experiment	Used Cars Sample Size	PSE Sample Size	Balance	Restrictions
Prediction 1	20,000	5,000	Unbalanced	Unrestricted
Prediction 2	20,000	5,000	Balanced	Unrestricted
Prediction 3	5,000	2,000	Unbalanced	Unrestricted
Prediction 4	5,000	2,000	Balanced	Unrestricted
Prediction 5	2,000	1,000	Unbalanced	Unrestricted
Prediction 6	2,000	1,000	Balanced	Unrestricted
Prediction 7	20,000	5,000	Unbalanced	Restricted
Prediction 8	20,000	5,000	Balanced	Restricted
Prediction 9	5,000	2,000	Unbalanced	Restricted
Prediction 10	5,000	2,000	Balanced	Restricted
Prediction 11	2,000	1,000	Unbalanced	Restricted
Prediction 12	2,000	1,000	Balanced	Restricted

2.7.3 MSE Experiments

For simplicity, we describe the MSE experiments using the used cars data. The experiments using the PSE data are carried out in the same manner with the only change being the sample size. We

draw two random samples for the Monte Carlo MSE experiments. First, we draw a random sample of 20,000 observations from the full sample and use this data to create an estimation sample. In the first MSE experiment, this random sample actually contains 17,458 loans (87.3%) that were successfully paid off and 2,542 loans (12.7%) that went into default. Thus, the estimation sample data contains roughly the same percentage of defaults and repayments as the full sample. A randomly drawn hold-out sample of 5,000 observations is used to verify the predictive ability of the model. We expect the model to predict well in the estimation sample, since the estimates are obtained using these data. Thus, the hold-out sample is used to examine how well the estimators predict in a sample of observations outside those used to develop the model. We carry out several experiments with these data to compare the ME and GME estimation techniques to traditional binary choice estimation techniques of probit, logit, and DA.

We use generated data for the Monte Carlo MSE experiments. We choose the probit estimates from the 20,000 estimation sample observations as the "true" parameters. Probit estimation gives the vector of parameters, β_p, which we use to generate the data. We also carry out experiments with β_p chosen to give a larger percentage of defaults. We generate new estimation and hold-out data sets during each Monte Carlo iteration using the original probit estimates as

$$y^* = X\beta_p + e \qquad e \sim N(0,1)$$

and

$$y_0^* = X_0\beta_p + e_0 \qquad e_0 \sim N(0,1)$$

where y^* is an $N \times 1$ vector of latent values, X is an $N \times K$ matrix of explanatory variables, y_0^* is an $N_0 \times 1$ vector of latent values for the hold-out sample, X_0 is an $N_0 \times K$ matrix of explanatory variables for the hold-out sample, and e and e_0 are vectors of random errors drawn from a standard normal distribution. The latent values are used to generate the observed sample where

$$y = \begin{cases} 1 & \text{if } y^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.30}$$

and y_0 is determined in the same manner. As part of the experimental design, we also carry out experiments with errors drawn from standardized t- and standardized chi-square distributions to examine the robustness of the estimators.

Finally, during each iteration we estimate the model using the generated data with the alternative estimation techniques (probit, logit, DA, ME, or GME). The new estimates, $\hat{\beta}$, are then used to obtain the predicted probabilities as follows:

$$\hat{p} = \Phi(X\hat{\beta}) \tag{2.31}$$

$$\hat{p}_0 = \Phi(X_0\hat{\beta}), \tag{2.32}$$

where X_0 is the $5{,}000 \times K$ matrix of explanatory variables for the hold-out sample and $\Phi(z)$ represents the standard normal cumulative distribution function evaluated at z (we use the logistic cumulative distribution function for logit, DA, ME, and GME). [This Monte Carlo step is repeated NSAM=500 times.] The generated data contain an average of 17,459 repayments and 2,541 defaults, which are nearly identical to the numbers in the original data set.

In general, an individual would be assigned to the category $y = 1$ if the predicted probability were greater than 0.5 and to $y = 0$ otherwise. However, given the unbalanced sample in the used cars data, each of the models predicts one hundred percent repayment using these criteria. Therefore, following Greene (1997, pp. 892-893) we lower the prediction threshold in order to generate an estimation sample percentage of predicted defaults that is closest to the true percentage of defaults. For the original estimation sample 20,000 observations, a threshold value of 0.175 generates 2,666 predicted defaults (13.3%). We use this threshold for prediction in the unbalanced Monte Carlo estimation and hold-out samples. For balanced samples, which are also considered, we use the threshold value of 0.5.

We obtain three separate sets of GME parameter estimates. The first set (GME1) constrains the errors to lie in the interval $[-1,1]$. We know that the errors must fall in this range for binary choice models since y can only equal zero or one. Golan, Judge, and Miller (1996, p. 253) argue that the

boundaries for the errors should be set at $\pm 1/\sqrt{N}$ since we assume that the sum of the errors goes to zero as we increase the sample size. For very large samples, we assume that the errors sum to zero and GME reduces to ME-logit as the error bounds go to zero. However, in small samples GME allows the errors to sum to something between 0 and 1 and we eliminate the restriction that the proportion of 1's in the sample is equal to the average of the predicted probabilities. Golan, Judge, and Perloff (1996) using a sample size of $N = 100$, impose error bounds of $[-.1, .1]$. We use these bounds to obtain the GME2 estimator. For our sample size of 20,000 observations the rule suggested by Golan et al. yields boundaries that are approximately equal to ± 0.007. We estimate the model with error bounds of $[-.01, .01]$ for the GME3 estimator to account for our large sample size.

As discussed in section 2.3, the ME and logit coefficient estimates are identical when the data constraint (2.11) is imposed (Soofi 1992). Additionally, the GME estimates approach the ME estimates as we narrow the error bounds. With the wider error bounds in model GME1, $[-1,1]$, the parameter estimates are much smaller than the ME estimates in absolute value. GME shrinks the predicted probabilities toward 0.5 by placing weight on the error term in the entropy function. Therefore, a change in an independent variable has a smaller effect on the dependent variable and the parameter estimates are smaller. As the bounds narrow, less weight falls on the unobservable errors and more weight falls on the parameters. Golan, Judge, and Perloff (1996, p. 845) note that as N goes to infinity (and the error bounds go to zero) the GME estimates are equivalent to the ME estimates.

In addition to calculating the parameter estimates using the various estimation techniques, we calculate a prediction table, which gives the average percentage of repayments and defaults correctly and incorrectly predicted by the model. We calculate the mean squared error for the predicted latent values (YMSE) as

$$YMSE = \sum_{t=1}^{nsam} \left[\sum_{n=1}^{N} (\hat{y}_n^* - y_n^*)^2 \right] \Big/ nsam \tag{2.33}$$

for the estimation sample observations, and

$$YMSE_0 = \sum_{i=1}^{nsam}\left[\sum_{n=1}^{N_0}(\hat{y}_{0n}^* - y_{0n}^*)^2\right]\Big/ nsam \tag{2.34}$$

for the hold-out observations. Recall that y^* are the actual latent values, which are known in the Monte Carlo setting, and $\hat{y}^*$ are the predicted latent values. The true parameter estimates, probabilities, and marginal effects are also known in our Monte Carlo experiments. The MSE for the predicted probabilities is calculated as

$$PMSE = \sum_{i=1}^{nsam}\left[\sum_{n=1}^{N}(\hat{p}_n - p_n)^2\right]\Big/ nsam \tag{2.35}$$

for the in-sample observations, and

$$PMSE_0 = \sum_{i=1}^{nsam}\left[\sum_{n=1}^{N_0}(\hat{p}_{0n} - p_{0n})^2\right]\Big/ nsam \tag{2.36}$$

for the hold-out observations. (However, as discussed in section 2.5 we use $\hat{y} = \hat{p} + \hat{e}$ rather than $\hat{p}$ in ME and GME). In addition, we calculate the MSE $(\hat{\beta})$, which is given by

$$MSE(\hat{\beta}) = \sum_{i=1}^{nsam}\left[\sum_{k=1}^{K}(\hat{\beta}_k - \beta_k)^2\right]\Big/ nsam . \tag{2.37}$$

The root mean squared error (RMSE) is simply the square root of the MSE. The MSE $(\hat{\beta})$ is equal to the variance plus the bias squared of the parameter estimates, $\hat{\beta}$. These terms are given by

$$VAR(\hat{\beta}) = \sum_{i=1}^{nsam}\left[\sum_{k=1}^{K}(\hat{\beta}_k - \bar{\beta}_k)^2\right]\Big/ nsam \tag{2.38}$$

and

$$Bias^2(\hat{\beta}) = \sum_{i=1}^{nsam}\left[\sum_{k=1}^{K}(\bar{\beta}_k - \beta_k)^2\right]\Big/ nsam , \tag{2.39}$$

respectively. Finally, since the parameter estimates are not equal to the marginal effects in discrete choice models we calculate the MSE of the marginal effects, $\partial p / \partial x$. For the probit model the marginal effects are given by

$$\partial p / \partial x = \phi(x'\beta) \cdot \beta , \tag{2.40}$$

where $\phi(z)$ represents the probability density function of the standard normal distribution evaluated at z. For the logit, DA, ME, and GME models the marginal effects are given by

$$\partial p / \partial x = \Lambda(x'\beta)[1 - \Lambda(x'\beta)] \cdot \beta, \tag{2.41}$$

where $\Lambda(z)$ represents the probability density function of the logistic distribution evaluated at z. The MSE of the estimated marginal effects is given by

$$MSE(\partial p / \partial x) = \sum_{i=1}^{nsam} \left[\sum_{k=1}^{K} (\hat{s}_k - s_k)^2 \right] \Big/ nsam, \tag{2.42}$$

where $\hat{s}_k$ represents the estimated marginal effects and s_k represents the true marginal effects, which are known in our sampling experiments. Since the marginal effects must be evaluated at a particular value of x, we evaluate them at the means of each of the continuous variables and at the value that occurs most frequently for each of the dummy variables.

2.7.4 Dimensions of the Experimental Design

The dimensions of the experimental design include the following:

1) First, we carry out the prediction sampling experiments. From the full set of observations we draw data for an estimation sample (20,000 observations) and a hold-out sample (5,000 observations). Based on these data we employ the alternative predictors (with and without inequality constraints) to predict the estimation and hold-out sample choices. (This Monte Carlo step is repeated $NSAM = 500$ times.) For this exercise, we can only evaluate the percentage of correct predictions. Table 2.3 summarizes the experimental design for the prediction experiments.

For the MSE experiments:

2) We vary the "true" parameter vector. We draw an estimation sample with an equal number of defaults and repayments. The alternative values yield a higher percentage of defaults in the generated data.

3) We vary the size of the estimation sample data set. In addition to the sample size of 20,000 observations we carry out experiments with sample sizes of 5,000 and 2,000 observations.

4) To test whether entropy estimation is robust we conduct experiments with errors drawn from different distributions. Errors are drawn from a t-distribution with 3 degrees of freedom and from a chi-square distribution with 5 degrees of freedom, correcting the mean to zero and the variance to one. (Note that the t-distribution yields a smaller percentage of defaults than the other distributions since it has thicker tails; therefore, we lower the threshold value to 0.125 for the samples generated using a t-distribution).

5) We modify the estimators so as to constrain the estimates to take their expected signs in each sample. Imposing such inequality information is made feasible by readily available modern software. For the used cars data, we constrain income (cdcmi), down payment percentage (cdtdp), age, home, and cosign to be negative; as these variables increase we expect the probability of default to decrease. We constrain monthly payment amount (cdfpay), amount financed (amtfin), and interest rate (netrate) to be positive. For the PSE data, we constrain all of the variables except the intercept to be positive.

As in the prediction experiments, we have 12 experimental designs. However, in the MSE experiments we also draw errors from three separate distributions. Therefore, we term the MSE experiments MSE11-MSE13, MSE21-MSE23, ..., MSE121-MSE123. Table 2.4 summarizes the dimensions for the MSE experiments.

2.8 Monte Carlo Results

In this section, we present Monte Carlo results which examine risk measures (MSE of prediction and of the parameter estimates) and the percentage of correct predictions for the alternative estimators. We present results for both binary and multinomial choice experiments. For the binary choice experiments we have data on used car loans and on the decision whether or not to attend post-secondary education.

2.8.1 Binary Choice Prediction Experiments with the Used Cars Data

In the prediction experiments we draw random samples which reflect the true proportion of defaults in the data, as well as samples drawn with an equal number of repayments and defaults. In addition, we vary the estimation sample size and perform experiments both with and without sign restrictions placed on the parameters. Table 2.3 summarizes the experimental design for the prediction experiments.

Table 2.5 reports the average parameter estimates for the alternative estimators over $NSAM = 500$ Monte Carlo iterations as well as the results from the Monte Carlo Prediction 1 experiment with the used cars data. The average parameter estimates have the expected signs for each alternative estimator (the expected signs are summarized in the experimental design in section 2.7.4). The t-statistics calculated from the sample standard errors are given in parentheses. These sample t-statistics are similar for each of the estimation techniques. The sample standard errors are calculated as

Table 2.4 Dimensions of Monte Carlo MSE Experiments

Experiment	Used Cars Sample Size	PSE Sample Size	Type of Data Set	Restrictions	Errors
MSE 11	20,000	5,000	Unbalanced	Unrestricted	Normal
MSE 12	20,000	5,000	Unbalanced	Unrestricted	Std.-t
MSE 13	20,000	5,000	Unbalanced	Unrestricted	Chi-square
MSE 21	20,000	5,000	Balanced	Unrestricted	Normal
MSE 22	20,000	5,000	Balanced	Unrestricted	Std.-t
MSE 23	20,000	5,000	Balanced	Unrestricted	Chi-square
MSE 31	5,000	2,000	Unbalanced	Unrestricted	Normal
MSE 32	5,000	2,000	Unbalanced	Unrestricted	Std.-t
MSE 33	5,000	2,000	Unbalanced	Unrestricted	Chi-square
MSE 41	5,000	2,000	Balanced	Unrestricted	Normal
MSE 42	5,000	2,000	Balanced	Unrestricted	Std.-t
MSE 43	5,000	2,000	Balanced	Unrestricted	Chi-square
MSE 51	2,000	1,000	Unbalanced	Unrestricted	Normal
MSE 52	2,000	1,000	Unbalanced	Unrestricted	Std.-t
MSE 53	2,000	1,000	Unbalanced	Unrestricted	Chi-square
MSE 61	2,000	1,000	Balanced	Unrestricted	Normal
MSE 62	2,000	1,000	Balanced	Unrestricted	Std.-t
MSE 63	2,000	1,000	Balanced	Unrestricted	Chi-square
MSE 71	20,000	5,000	Unbalanced	Restricted	Normal
MSE 72	20,000	5,000	Unbalanced	Restricted	Std.-t
MSE 73	20,000	5,000	Unbalanced	Restricted	Chi-square
MSE 81	20,000	5,000	Balanced	Restricted	Normal
MSE 82	20,000	5,000	Balanced	Restricted	Std.-t
MSE 83	20,000	5,000	Balanced	Restricted	Chi-square
MSE 91	5,000	2,000	Unbalanced	Restricted	Normal
MSE 92	5,000	2,000	Unbalanced	Restricted	Std.-t
MSE 93	5,000	2,000	Unbalanced	Restricted	Chi-square
MSE 101	5,000	2,000	Balanced	Restricted	Normal
MSE 102	5,000	2,000	Balanced	Restricted	Std.-t
MSE 103	5,000	2,000	Balanced	Restricted	Chi-square
MSE 111	2,000	1,000	Unbalanced	Restricted	Normal
MSE 112	2,000	1,000	Unbalanced	Restricted	Std.-t
MSE 113	2,000	1,000	Unbalanced	Restricted	Chi-square
MSE 121	2,000	1,000	Balanced	Restricted	Normal
MSE 122	2,000	1,000	Balanced	Restricted	Std.-t
MSE 123	2,000	1,000	Balanced	Restricted	Chi-square

Table 2.5 Used Cars Data - Monte Carlo Prediction 1 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.3112	-2.2385	-2.2989	-2.2385	-0.4569	-2.0938	-2.2369
	(-30.43)	(-26.90)	(-29.10)	(-26.90)	(-47.36)	(-29.50)	(-26.93)
Cdcmi	-.00004	-.00008	-.00007	-.00008	-.00001	-.00007	-.00008
	(-3.61)	(-3.63)	(-3.89)	(-3.63)	(-3.88)	(-3.69)	(-3.63)
Cdfpay	.00163	.00293	.00328	.00293	.00041	.00260	.00292
	(5.36)	(5.26)	(5.26)	(5.26)	(5.25)	(5.26)	(5.26)
Cdtdp	-.00205	-.00332	-.00334	-.00332	-.00042	-.00290	-.00332
	(-2.71)	(-2.40)	(-2.54)	(-2.40)	(-2.54)	(-2.43)	(-2.40)
Age	-.00390	-.00777	-.00718	-.00777	-.00091	-.00666	-.00775
	(-4.26)	(-4.29)	(-4.49)	(-4.29)	(-4.50)	(-4.35)	(-4.30)
Amtfin	.00001	.00003	.00003	.00003	.00000	.00002	.00003
	(1.63)	(1.78)	(1.52)	(1.78)	(1.53)	(1.73)	(1.78)
Netrate	.00248	.00558	.00342	.00558	.00045	.00450	.00556
	(1.59)	(1.85)	(1.20)	(1.85)	(1.25)	(1.73)	(1.84)
Home	-.19754	-.37652	-.36012	-.37652	-.04570	-.32459	-.37588
	(-5.64)	(-5.58)	(-6.05)	(-5.58)	(-6.01)	(-5.69)	(-5.59)
Cosign	-.08344	-.15468	-.16096	-.15468	-.02031	-.13545	-.15444
	(-1.70)	(-1.66)	(-1.87)	(-1.66)	(-1.86)	(-1.70)	(-1.66)
Estimation Sample							
Percent Correct	80.92	80.97	80.50	80.97	80.90	80.96	80.97
Standard Deviation	.0108	.0105	.0102	.0105	.0125	.0109	.0105
% Repay. Correct	89.62	89.67	88.99	89.67	89.59	89.65	89.67
% Defaults Correct	19.98	20.08	21.05	20.08	20.07	20.09	20.08
Hold-out Sample							
Percent Correct	80.91	80.96	80.49	80.96	80.89	80.94	80.96
Standard Deviation	.0118	.0114	.0113	.0114	.0133	.0118	.0115
% Repay. Correct	89.61	89.66	88.99	89.66	89.58	89.64	89.66
% Defaults Correct	19.75	19.84	20.83	19.84	19.82	19.85	19.85

* N=20,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

$$s.e.(\beta) = \sqrt{\left(\sum_{i=1}^{nsam} \left[\sum_{k=1}^{K} (\hat{\beta}_k - \beta_k)^2 \right] \Big/ nsam \right)}. \tag{2.43}$$

Table 2.6 compares the sample standard errors to the asymptotic standard errors for logit and probit for the used cars Prediction 1 experiment. The asymptotic standard errors given in Table 2.6 are the average of the asymptotic standard errors over $NSAM = 500$ Monte Carlo iterations. The sample and asymptotic standard errors are very close to each other for both probit and logit in the Prediction 1 experiment, as well as in all of our other experiments.

Table 2.6 Comparison of Aymptotic and Sample Standard Errors for Logit and Probit

Variable	Probit asy. s.e.	Probit sample s.e.	Logit asy. s.e.	Logit Sample s.e.
Intercept	.04144	.04309	.07932	.08321
Cdcmi	.00001	.00001	.00002	.00002
Cdfpay	.00028	.00030	.00051	.00056
Cdtdp	.00087	.00076	.00164	.00138
Age	.00084	.00091	.00161	.00181
Amtfin	.00001	.00001	.00001	.00002
Netrate	.00153	.00157	.00290	.00302
Home	.03214	.03502	.06192	.06743
Cosign	.04829	.04898	.09200	.09331

The results from the Prediction 1 experiment indicate very little difference in predictive ability between the alternative estimation techniques. This is expected since, as Greene (1997, p. 875) reports, the normal and logistic distributions are very similar except in the tails. In addition, we showed in sections 2.3 and 2.4 that ME and logit are identical and that GME shrinks each predicted probability toward 0.5, but will only change the predicted outcome of y for probabilities very close to the prediction threshold value. The standard error of prediction is higher for the GME1 estimator than for the alternative estimators.

We find that the percentage of correct predictions in the hold-out sample is very close to the prediction percentage in the estimation sample for each alternative estimator, although the standard deviation of prediction is higher for the hold-out sample than for the estimation sample. Each model correctly predicts 89-90% of the repayments, but only predicts 20-21% of the defaults correctly. This is due to the unbalanced sample. The highly negative coefficient for the intercept suggests that individuals

have a strong predisposition toward repayment and that a number of events must occur to accurately predict default. In addition, since we have a truncated sample, those individuals with the highest default risk were either denied a loan or self-selected not to apply for a loan since they would have been denied.

Note that the GME1 parameter estimates are much smaller in magnitude than those of the alternative estimators. Despite this the predictive performance of GME1 and the alternative estimators is very close. Recall that GME shrinks the predicted probabilities toward 0.5 and only differs from ME-logit prediction for ME-logit probabilities near the prediction threshold. The GME1 parameter estimates are the closer to zero since GME1 has the widest error bounds, -1 and 1, of the GME estimators considered. The ME and logit estimates are identical as noted by Soofi (1992) and Golan, Judge, and Perloff (1996). Finally, DA correctly predicts a greater percentage of defaults but fewer repayments than the alternative estimators do. This is due to the threshold we chose and the fact that DA has a larger number of predicted probabilities that are below the threshold. This difference between DA and the alternative estimators disappears when we use a balanced sample.

Table 2.7 gives the results for the used cars Prediction 2 experiment, which uses a sample with an equal number of defaults and repayments. The results from the balanced sample show virtually no difference in prediction between the alternative estimation techniques. The tables for the used cars Prediction 3-6 experiments are given in Appendix C. The results are consistent as we vary the sample size. The only change is that we see a larger difference in prediction percentage between the estimation and hold-out samples the smaller the sample size. In the used cars Prediction 7-12 experiments, we estimate the models for the used cars data imposing inequality restrictions that constrain each parameter to take its expected sign. We obtain probit and logit estimates using the GAUSS Constrained Optimization module which allows us to place inequality constraints on the parameters. We find very small differences in the parameter estimates and prediction percentages between the restricted and unrestricted experiments because the parameter estimates already take the expected signs in most of the Monte Carlo iterations. We report all of the results for the used cars prediction experiments in Appendix C, Tables C.1-C.12.

Table 2.7 Used Cars Data – Monte Carlo Prediction 2 (See Table 2.3 for design)

Variable	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-.20741	-.33803	-.33436	-.33803	-.08870	-.32818	-.33793
	(-6.39)	(-6.51)	(-6.42)	(-6.51)	(-6.47)	(-6.51)	(-6.51)
Cdcmi	-.00004	-.00007	-.00007	-.00007	-.00002	-.00007	-.00007
	(-5.69)	(-5.70)	(-5.75)	(-5.70)	(-5.76)	(-5.71)	(-5.70)
Cdfpay	.00203	.00336	.00322	.00336	.00086	.00326	.00336
	(8.05)	(8.84)	(8.59)	(8.84)	(8.74)	(8.84)	(8.84)
Cdtdp	-.00308	-.00497	-.00492	-.00497	-.00130	-.00483	-.00497
	(-4.40)	(-4.42)	(-4.40)	(-4.42)	(-4.41)	(-4.42)	(-4.42)
Age	-.00426	-.00687	-.00688	-.00687	-.00182	-.00667	-.00687
	(-6.95)	(-6.90)	(-6.94)	(-6.90)	(-6.96)	(-6.90)	(-6.90)
Amtfin	.00001	.00002	.00002	.00002	.00001	.00002	.00002
	(1.93)	(1.89)	(2.19)	(1.89)	(2.14)	(1.91)	(1.89)
Netrate	.00263	.00399	.00449	.00399	.00116	.00391	.00399
	(2.15)	(2.04)	(2.29)	(2.04)	(2.25)	(2.06)	(2.04)
Home	-.22829	-.36668	-.36717	-.36668	-.09711	-.35620	-.36657
	(-9.22)	(-9.18)	(-9.18)	(-9.18)	(-9.26)	(-9.19)	(-9.18)
Cosign	-.09541	-.15273	-.15301	-.15273	-.04048	-.14838	-.15268
	(-2.45)	(-2.45)	(-2.46)	(-2.45)	(-2.46)	(-2.45)	(-2.45)
Estimation Sample							
Percent Correct	58.11	58.12	58.11	58.12	58.11	58.12	58.12
Standard Deviation	.0035	.0035	.0035	.0035	.0035	.0035	.0035
% Repay. Correct	59.36	59.37	59.39	59.37	59.39	59.37	59.37
% Defaults Correct	56.86	56.86	56.82	56.86	56.83	56.86	56.86
Hold-out Sample							
Percent Correct	58.05	58.07	58.06	58.07	58.06	58.06	58.07
Standard Deviation	.0071	.0070	.0071	.0070	.0071	.0071	.0070
% Repay. Correct	59.33	59.34	59.37	59.34	59.37	59.35	59.34
% Defaults Correct	56.77	56.79	56.74	56.79	56.75	56.78	56.79

* N=20,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

2.8.2 Binary Choice MSE Experiments with the Used Cars Data

In the MSE experiments we draw random samples which reflect the true proportion of defaults in the data, as well as samples drawn with an equal number of repayments and defaults. We generate new data sets using these random samples as a basis. In addition, we vary the estimation sample size, vary the error distribution, and perform experiments both with and without sign restrictions placed on the parameters. Table 2.4 summarizes the experimental design for the MSE experiments.

Tables 2.8-2.10 give the results for the Monte Carlo MSE 11-13 experiments with the used cars data. As in the prediction experiments, we see little variation in the percentage of correct prediction between the alternative estimators. The probit estimator has the lowest MSE of the predicted latent values and of the predicted probabilities as well as the lowest risk measures for $\hat{\beta}$. However, in these experiments we generated the data during each Monte Carlo iteration using the original probit estimates as the true parameter values. The GME estimators generally have lower risk measures than the logit, DA, and ME estimators. The variance is lowest for GME1 since it places more weight on the error term, which shrinks the parameter estimates toward zero. Because of the shrinkage properties of GME we expected GME to have lower variance than the alternative estimators, however we also find that the GME2 and GME3 estimators have lower bias than logit, DA, and ME.

Each of the alternative estimators has higher risk when we draw errors from a *t*- distribution with 3 degrees of freedom, correcting the mean to zero and the variance to one, except for the GME1 estimator which performs slightly better. The alternative estimators generally perform better when we draw errors from a standardized chi-square distribution except for probit, which performs best when we draw errors from the standard normal distribution.

Tables 2.11-2.13 give the results for the Monte Carlo MSE 21-23 experiments, which use a random sample chosen to generate a relatively equal number of defaults and repayments. The results show very little difference in the percentage of correct predictions, the MSE of the predicted latent values, and the MSE of the predicted probabilities. Probit again has the lowest risk measures for $\hat{\beta}$, but the difference between the probit risk measures and the risk measures for the alternative estimators is smaller in the balanced sample. We note that the risk measures for the GME estimators are very close to those for logit, DA, and ME in the balanced sample. This is because GME shrinks the predicted probabilities toward 0.5; with the balanced sample, many of the predicted probabilities are already close to 0.5 so the degree of shrinkage is less. The GME2 estimator, which has error bounds based on the rule suggested the Golan, Judge, and Perloff (1996, p. 845), has the lowest MSE for the marginal effects of

Table 2.8 Used Cars Data – Monte Carlo MSE 11 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.41	79.56	79.07	79.56	79.15	79.49	79.56
Repayments	87.75	87.96	87.25	87.96	87.37	87.87	87.96
Defaults	22.06	21.76	22.78	21.76	22.58	21.90	21.76
Hold-Out	79.61	79.76	79.25	79.76	79.33	79.69	79.76
Repayments	88.04	88.26	87.52	88.26	87.64	88.16	88.26
Defaults	21.29	20.97	22.05	20.97	21.88	21.12	20.97
YRMSE	141.44	184.77	185.56	184.77	177.93	173.59	184.64
$YRMSE_0$	70.71	92.43	92.81	92.43	89.00	86.82	92.36
PRMSE	1.02	1.06	1.32	1.06	1.36	1.02	1.06
$PRMSE_0$	0.51	0.53	0.67	0.53	0.67	0.51	0.53
$MSE(\hat{\beta})$	.00518	1.05177	1.14663	1.05177	.84861	.72540	1.04764
$RMSE(\hat{\beta})$	.07195	1.02556	1.07081	1.02556	.92120	.85170	1.02354
$VAR(\hat{\beta})$	.00518	.01895	.01592	.01895	.00025	.01373	.01888
$Bias^2(\hat{\beta})$	.00000	1.03282	1.13071	1.03282	.84836	.71167	1.02876
MSE ($\partial p / \partial x$)	.00023	.00086	.00063	.00086	.03316	.00072	.00086
RMSE ($\partial p / \partial x$)	.01525	.02936	.02511	.02936	.18209	.02684	.02932

* N=20,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table 2.9 Used Cars Data – Monte Carlo MSE 12 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	89.29	89.21	88.15	89.21	89.69	89.35	89.21
Repayments	95.56	95.46	94.14	95.46	96.07	95.63	95.46
Defaults	10.46	10.65	12.88	10.65	9.47	10.33	10.64
Hold-Out	89.50	89.43	88.37	89.43	89.84	89.55	89.43
Repayments	95.80	95.71	94.38	95.71	96.23	95.86	95.71
Defaults	9.60	9.77	12.15	9.77	8.79	9.49	9.77
YRMSE	148.32	248.97	250.66	248.97	173.21	221.34	248.61
$YRMSE_0$	74.08	124.55	125.37	124.55	86.52	110.68	124.37
PRMSE	1.02	0.93	1.17	0.93	1.63	1.05	0.93
$PRMSE_0$	0.53	0.47	0.59	0.47	0.82	0.54	0.47
$MSE(\hat{\beta})$	.09771	2.70755	2.97512	2.70755	.77347	1.73941	2.69359
$RMSE(\hat{\beta})$	.31259	1.64546	1.72485	1.64546	.87947	1.31887	1.64122
$VAR(\hat{\beta})$	.00717	.03110	.02593	.03110	.00016	.01702	.03085
$Bias^2(\hat{\beta})$	.09054	2.67645	2.94918	2.67645	.77330	1.72239	2.66274
MSE ($\partial p / \partial x$)	.00352	.00794	.00734	.00794	.03024	.00556	.00791
RMSE ($\partial p / \partial x$)	.05929	.08910	.08569	.08910	.17391	.07454	.08891

* N=20,000; unbalanced; unrestricted; errors drawn from *t*- distribution.

Table 2.10 Used Cars Data – Monte Carlo MSE 13 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	83.47	83.44	82.87	83.44	83.69	83.49	83.44
Repayments	93.96	93.91	93.09	93.91	94.28	93.98	93.91
Defaults	10.45	10.54	11.72	10.54	9.99	10.43	10.54
Hold-Out	83.61	83.58	83.00	83.58	83.85	83.63	83.58
Repayments	94.13	94.08	93.24	94.08	94.47	94.16	94.08
Defaults	9.87	9.96	11.17	9.96	9.39	9.83	9.96
YRMSE	141.72	181.98	182.29	181.98	178.11	172.01	181.86
$YRMSE_0$	70.84	91.02	91.18	91.02	89.06	86.02	90.96
PRMSE	1.06	1.03	1.12	1.03	1.31	1.05	1.03
$PRMSE_0$	0.52	0.51	0.56	0.51	0.61	0.51	0.51
$MSE(\hat{\beta})$	.01287	.79170	.84622	.79170	.87912	.54670	.78864
$RMSE(\hat{\beta})$	.11344	.88977	.91990	.88977	.93761	.73939	.88805
$VAR(\hat{\beta})$	.00474	.01712	.01534	.01712	.00025	.01266	.01706
$Bias^2(\hat{\beta})$	.00813	.77457	.83089	.77457	.87888	.53404	.77157
MSE ($\partial p / \partial x$)	.00066	.00202	.00178	.00202	.03465	.00181	.00202
RMSE ($\partial p / \partial x$)	.02560	.04499	.04220	.04499	.18614	.04259	.04496

* N=20,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table 2.11 Used Cars Data – Monte Carlo MSE 21 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	58.22	58.22	58.21	58.22	58.21	58.22	58.22
Repayments	59.02	59.02	59.03	59.02	59.03	59.02	59.02
Defaults	57.39	57.40	57.38	57.40	57.38	57.40	57.40
Hold-Out	58.15	58.15	58.15	58.15	58.15	58.15	58.15
Repayments	58.58	58.57	58.61	58.57	58.61	58.58	58.57
Defaults	57.72	57.73	57.69	57.73	57.69	57.75	57.72
YRMSE	141.42	143.27	143.18	143.27	143.01	143.00	143.27
$YRMSE_0$	70.73	71.68	71.63	71.68	71.50	71.54	71.68
PRMSE	1.46	1.46	1.47	1.46	1.60	1.46	1.46
$PRMSE_0$	0.73	0.73	0.73	0.73	0.76	0.73	0.73
$MSE(\hat{\beta})$	.00316	.06463	.06306	.06463	.05103	.05574	.06453
$RMSE(\hat{\beta})$	.05624	.25422	.25112	.25422	.22590	.23608	.25404
$VAR(\hat{\beta})$	.00316	.00819	.00816	.00819	.00056	.00771	.00819
$Bias^2(\hat{\beta})$	.00000	.05644	.05490	.05644	.05047	.04802	.05635
MSE ($\partial p / \partial x$)	.00050	.00051	.00051	.00051	.01300	.00049	.00051
RMSE ($\partial p / \partial x$)	.02241	.02264	.02261	.02264	.11404	.02216	.02263

* N=20,000; balanced; unrestricted; errors drawn from normal distribution.

Table 2.12 Used Cars Data – Monte Carlo MSE 22 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	62.40	62.40	62.39	62.40	62.39	62.40	62.40
Repayments	63.26	63.25	63.26	63.25	63.27	63.25	63.25
Defaults	61.53	61.54	61.51	61.54	61.51	61.53	61.54
Hold-Out	62.42	62.41	62.41	62.41	62.42	62.42	62.41
Repayments	62.99	62.98	63.03	62.98	63.03	62.98	62.98
Defaults	61.86	61.86	61.80	61.86	61.82	61.86	61.86
YRMSE	142.64	151.76	151.31	151.76	141.92	150.64	151.75
$YRMSE_0$	72.05	76.60	76.37	76.60	71.66	76.04	76.59
PRMSE	1.51	1.45	1.62	1.45	2.45	1.48	1.45
$PRMSE_0$	0.75	0.72	0.81	0.72	1.13	0.74	0.72
$MSE(\hat{\beta})$	.04712	.34878	.34491	.34878	.02129	.31142	.34838
$RMSE(\hat{\beta})$	.21707	.59057	.58729	.59057	.14591	.55805	.59023
$VAR(\hat{\beta})$	.00307	.00811	.00806	.00811	.00050	.00756	.00810
$Bias^2(\hat{\beta})$	.04405	.34067	.33685	.34067	.02079	.30386	.34027
MSE ($\partial p / \partial x$)	.00742	.00812	.00799	.00812	.00882	.00675	.00811
RMSE ($\partial p / \partial x$)	.08611	.09012	.08938	.09012	.09394	.08218	.09004

* N=20,000; balanced; unrestricted; errors drawn from *t*- distribution.

Table 2.13 Used Cars Data – Monte Carlo MSE 23 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	60.64	60.64	60.64	60.64	60.64	60.64	60.64
Repayments	82.95	82.92	83.03	82.92	83.28	82.95	82.92
Defaults	30.28	30.33	30.17	30.33	29.84	30.28	30.33
Hold-Out	60.55	60.55	60.55	60.55	60.55	60.55	60.55
Repayments	82.09	82.04	82.20	82.04	82.45	82.08	82.04
Defaults	31.52	31.57	31.36	31.57	31.02	31.52	31.57
YRMSE	144.21	150.49	150.56	150.49	143.52	149.79	150.48
$YRMSE_0$	72.14	75.22	75.26	75.22	71.79	74.88	75.22
PRMSE	4.33	4.30	4.28	4.30	4.77	4.33	4.30
$PRMSE_0$	2.13	2.11	2.10	2.11	2.24	2.12	2.11
$MSE(\hat{\beta})$	.04824	.30345	.30808	.30345	.02975	.27417	.30314
$RMSE(\hat{\beta})$	.21963	.55086	.55505	.55086	.17249	.52361	.55058
$VAR(\hat{\beta})$	.00310	.00818	.00794	.00818	.00052	.00765	.00817
$Bias^2(\hat{\beta})$	.04514	.29527	.30014	.29527	.02923	.26652	.29497
MSE ($\partial p / \partial x$)	.00731	.00753	.00771	.00753	.00982	.00656	.00752
RMSE ($\partial p / \partial x$)	.08553	.08679	.08782	.08679	.09910	.08102	.08673

* N=20,000; balanced; unrestricted; errors drawn from chi-square distribution.

any of the estimators. We observe similar results between the alternative estimators as we vary the estimation sample size. We give the MSE results for experiments MSE 31-63, which are based on smaller estimation sample sizes, in Appendix C.

In the used cars experiments MSE 71-123, we estimate the models with sign restrictions placed on the parameters. We see virtually no difference in results with an estimation sample size N =20,000 observations and only small differences as we decrease the estimation sample size. We report all of the results for the used cars MSE experiments in Appendix C, Tables C.13-C.48.

2.8.3 Binary Choice Prediction Experiments with the PSE Data

Following the same methodology as with the used cars data, we estimate the probability that a high school graduate chooses to attend post-secondary education. Table 2.14 reports the average

Table 2.14 PSE Data - Monte Carlo Prediction 1 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.5266	-2.7307	-2.9629	-2.7307	-0.4169	-2.4158	-2.7267
	(-14.80)	(-15.05)	(-13.86)	(-15.05)	(-12.80)	(-15.18)	(-15.05)
Black	0.3366	0.6024	0.6110	0.6024	0.0965	0.5367	0.6016
	(4.46)	(4.63)	(4.07)	(4.63)	(4.12)	(4.54)	(4.63)
Other Race	0.1981	0.3878	0.3693	0.3878	0.0589	0.3437	0.3872
	(3.30)	(3.68)	(3.69)	(3.68)	(3.74)	(3.70)	(3.68)
Catholic HS	0.5466	1.0180	0.7114	1.0180	0.1156	0.8438	1.0156
	(4.70)	(4.58)	(6.19)	(4.58)	(6.13)	(5.00)	(4.59)
Income1	0.2634	0.4496	0.6250	0.4496	0.0959	0.4203	0.4493
	(4.54)	(4.50)	(4.74)	(4.50)	(4.78)	(4.58)	(4.50)
Income2	0.5857	1.0144	1.1458	1.0144	0.1792	0.9209	1.0132
	(7.76)	(7.51)	(7.98)	(7.51)	(8.20)	(7.77)	(7.52)
Income3	0.6770	1.2526	1.1065	1.2526	0.1743	1.0530	1.2497
	(6.76)	(6.81)	(7.61)	(6.81)	(7.76)	(7.31)	(6.82)
GPA	0.8602	1.5250	1.4910	1.5250	0.2360	1.3502	1.5227
	(22.03)	(21.74)	(23.01)	(21.74)	(25.99)	(23.48)	(21.77)
Parent HS	0.2777	0.4775	0.7305	0.4775	0.1108	0.4517	0.4772
	(3.53)	(3.56)	(3.78)	(3.56)	(3.81)	(3.63)	(3.56)
Parent College	0.7257	1.3106	1.2666	1.3106	0.1984	1.1343	1.3082
	(7.81)	(7.96)	(6.38)	(7.96)	(6.62)	(7.86)	(7.96)
Estimation Sample							
Percent Correct	80.52	80.55	80.44	80.55	80.29	80.55	80.55
Standard Deviation	.0056	.0056	.0056	.0056	.0053	.0056	.0055
% Repay. Correct	93.96	93.54	93.73	93.54	96.46	94.03	93.55
% Defaults Correct	34.34	35.92	34.77	35.92	24.75	34.25	35.90
Hold-out Sample							
Percent Correct	80.38	80.41	80.25	80.41	80.15	80.40	80.41
Standard Deviation	.0140	.0137	.0138	.0137	.0137	.0138	.0137
% Repay. Correct	93.92	93.51	93.65	93.51	96.41	93.98	93.52
% Defaults Correct	33.95	35.49	34.31	35.49	24.37	33.83	35.47

* N=5,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

parameter estimates over $NSAM = 500$ Monte Carlo iterations as well as the results from the Monte Carlo Prediction 1 experiment with the PSE data. As with the used cars data, we observe little difference in the asymptotic t-statistics, percentage of correct predictions and standard deviation of prediction between the alternative estimators in our prediction experiments with the PSE data. Each estimator correctly predicts 93-96% of those who attend post-secondary education, but only 25-35% of those who do not attend post-secondary education, using a prediction threshold equal to 0.5.

Table 2.15 reports the results for the Monte Carlo Prediction 2 experiment, which uses a sample with an equal number of individuals who attend post-secondary education and individuals who do not attend post-secondary education. As with the used cars data we observe very little difference in prediction between the alternative estimators in a balanced sample. We report all of the results for the PSE prediction experiments in Appendix C, Tables C.49-C.60.

2.8.4 Binary Choice MSE Experiments with the PSE Data

Tables 2.16-2.18 give the results for the Monte Carlo MSE 11-13 experiments with the PSE data. The probit estimator has the lowest risk measures, but again we use the original probit estimates as the true parameters when generating the data. As in the used cars experiments, the GME estimators have lower risk measures than the logit, DA, and ME estimators. The estimator GME1 has the lowest variance and also has a lower bias than GME3. However, it has the highest MSE for the marginal effects and the highest MSE of the predicted probabilities. The GME2 substantially outperforms logit, DA, ME, and GME3 in every risk measure for the Monte Carlo experiments with the unbalanced PSE data.

Tables 2.19-2.21 give the results for the Monte Carlo MSE 21-23 experiments with the PSE data, which use a random sample designed to generate a relatively equal number of students choosing post-secondary education as not. We again notice very little difference in prediction between the alternative estimators when we have a balanced sample. In addition, we notice very little difference in results between the experiments with and without inequality constraints placed on the parameters. We report all of the results for the Monte Carlo MSE experiments with the PSE data in Appendix C, Tables C.61-C.96.

Table 2.15 PSE Data - Monte Carlo Prediction 2 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-2.3216	-3.9639	-4.1999	-3.9639	-0.8266	-3.7298	-3.9613
	(-25.06)	(-23.85)	(-24.60)	(-23.85)	(-31.16)	(-24.87)	(-23.86)
Black	0.3400	0.5847	0.5913	0.5847	0.1169	0.5471	0.5842
	(5.27)	(5.31)	(5.20)	(5.31)	(5.22)	(5.30)	(5.31)
Other Race	0.1731	0.3210	0.3100	0.3210	0.0618	0.2994	0.3208
	(3.14)	(3.40)	(3.36)	(3.40)	(3.38)	(3.41)	(3.40)
Catholic HS	0.5492	0.9424	0.8615	0.9424	0.1728	0.8729	0.9416
	(5.57)	(5.56)	(6.18)	(5.56)	(6.16)	(5.65)	(5.56)
Income1	0.2704	0.4637	0.5053	0.4637	0.0991	0.4378	0.4634
	(5.36)	(5.43)	(5.64)	(5.43)	(5.70)	(5.47)	(5.43)
Income2	0.6134	1.0380	1.1251	1.0380	0.2210	0.9798	1.0374
	(9.30)	(9.35)	(9.44)	(9.35)	(9.64)	(9.41)	(9.35)
Income3	0.7220	1.2699	1.2241	1.2699	0.2429	1.1787	1.2689
	(7.76)	(7.95)	(8.46)	(7.95)	(8.64)	(8.09)	(7.95)
GPA	0.8979	1.5314	1.5957	1.5314	0.3146	1.4379	1.5304
	(29.46)	(28.07)	(27.44)	(28.07)	(35.72)	(29.41)	(28.09)
Parent HS	0.2703	0.4684	0.5018	0.4684	0.0987	0.4420	0.4682
	(3.75)	(3.76)	(4.24)	(3.76)	(4.23)	(3.82)	(3.76)
Parent College	0.7608	1.3010	1.3237	1.3010	0.2615	1.2177	1.3001
	(8.75)	(8.70)	(8.91)	(8.70)	(9.16)	(8.78)	(8.70)
Estimation Sample							
Percent Correct	75.56	75.58	75.51	75.58	75.53	75.58	75.58
Standard Deviation	.0060	.0060	.0059	.0060	.0059	.0059	.0060
% Repay. Correct	71.88	72.11	70.74	72.11	70.89	71.96	72.11
% Defaults Correct	79.24	79.05	80.28	79.05	80.17	79.21	79.05
Hold-out Sample							
Percent Correct	75.50	75.50	75.47	75.50	75.48	75.51	75.50
Standard Deviation	.0138	.0137	.0142	.0137	.0141	.0138	.0137
% Repay. Correct	71.72	71.93	70.59	71.93	70.73	71.79	71.93
% Defaults Correct	79.28	79.08	80.35	79.08	80.23	79.23	79.08

* N=5,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table 2.16 PSE Data - Monte Carlo MSE 11 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.78	79.79	79.72	79.79	79.55	79.79	79.79
PSE	95.53	93.30	93.09	93.30	95.74	93.65	93.30
Non-PSE	34.64	35.42	35.80	35.42	26.36	34.29	35.41
Hold-Out	80.27	80.27	80.14	80.27	79.99	80.28	80.27
PSE	93.31	93.03	92.85	93.03	95.69	93.44	93.04
Non-PSE	36.16	37.09	37.17	37.09	26.87	35.75	37.07
YRMSE	70.78	97.51	93.31	97.51	94.30	87.03	97.36
$YRMSE_0$	31.68	44.07	41.95	44.07	42.63	39.19	44.00
PRMSE	1.28	1.43	2.16	1.43	3.99	1.30	1.43
$PRMSE_0$	0.58	0.65	0.97	0.65	1.79	0.58	0.65
$MSE(\hat{\beta})$	.06564	2.89661	3.79238	2.89661	2.43038	1.65363	2.87848
$RMSE(\hat{\beta})$	.25621	1.70194	1.94740	1.70194	1.55897	1.28593	1.69661
$VAR(\hat{\beta})$	.06527	.20491	.21316	.20491	.00535	.15217	.20410
$Bias^2(\hat{\beta})$	.00037	2.69170	3.57922	2.69170	2.42503	1.50146	2.67437
MSE ($\partial p / \partial x$)	.00973	.01434	.03662	.01434	.42161	.00974	.01419
RMSE ($\partial p / \partial x$)	.09863	.11973	.19136	.11973	.64931	.09869	.11911

*N=5,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table 2.17 PSE Data - Monte Carlo MSE 12 (See Table 2.4 for design)

	Probit	Logit	D.A.	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	84.17	84.20	83.89	84.20	83.38	84.16	84.20
PSE	95.31	94.92	94.72	94.92	97.48	95.40	94.93
Non-PSE	40.18	41.88	41.11	41.88	27.70	39.78	41.86
Hold-Out	84.44	84.47	84.13	84.47	83.64	84.42	84.47
PSE	95.23	94.76	94.47	94.76	97.49	95.33	94.77
Non-PSE	41.07	43.13	42.61	43.13	27.98	40.58	43.09
YRMSE	72.95	128.35	117.68	128.35	92.63	105.17	127.99
$YRMSE_0$	32.69	58.21	52.92	58.21	41.88	47.48	58.04
PRMSE	2.19	1.57	3.17	1.57	6.46	2.39	1.58
$PRMSE_0$	0.99	0.71	1.43	0.71	2.92	1.08	0.71
$MSE(\hat{\beta})$	.27597	7.05055	8.70149	7.05055	2.32254	3.67324	6.99342
$RMSE(\hat{\beta})$	.52533	2.65529	2.94983	2.65529	1.52399	1.91657	2.64451
$VAR(\hat{\beta})$	.08010	.26376	.25734	.26376	.00473	.16427	.26183
$Bias^2(\hat{\beta})$	.19586	6.78678	8.44414	6.78678	2.31780	3.50897	6.73160
MSE ($\partial p / \partial x$)	.02487	.04698	.12674	.04698	.41052	.01449	.04632
RMSE ($\partial p / \partial x$)	.15770	.21676	.35600	.21676	.64072	.12037	.21523

*N=5,000; unbalanced; unrestricted; errors drawn from *t*- distribution.

Table 2.18 PSE Data - Monte Carlo MSE 13 (See Table 2.4 for design)

	Probit	Logit	D.A.	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	81.53	81.55	81.44	81.55	81.18	81.54	81.55
PSE	91.06	90.93	91.29	90.93	93.44	91.34	90.94
Non-PSE	51.28	51.74	50.16	51.74	42.27	50.41	51.73
Hold-Out	82.01	82.02	81.88	82.02	81.67	82.01	82.01
PSE	91.18	91.06	91.07	91.06	93.15	91.38	91.06
Non-PSE	52.09	52.49	51.86	52.49	44.13	51.39	52.47
YRMSE	76.90	134.27	117.46	134.27	93.05	108.07	133.86
$YRMSE_0$	34.64	61.25	53.18	61.25	42.06	49.10	61.05
PRMSE	69.00	68.82	68.60	68.82	71.39	69.30	68.83
$PRMSE_0$	31.74	31.66	31.55	31.66	32.86	31.89	31.67
$MSE(\hat{\beta})$	1.19691	11.7649	12.4019	11.7649	1.92572	6.60128	11.6775
$RMSE(\hat{\beta})$	1.09404	3.43000	3.52163	3.43000	1.38770	2.56930	3.41723
$VAR(\hat{\beta})$	.07910	.25722	.24808	.25722	.00492	.16276	.25537
$Bias^2(\hat{\beta})$	1.11781	11.5077	12.1538	11.5077	1.92080	6.43852	11.4221
MSE ($\partial p / \partial x$)	.17803	.23877	.31962	.23877	.36739	.09986	.23630
RMSE ($\partial p / \partial x$)	.42193	.48864	.56535	.48864	.60613	.31600	.48611

*N=5,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table 2.19 PSE Data - Monte Carlo MSE 21 (See Table 2.4 for design)

	Probit	Logit	D.A.	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	73.53	73.54	73.52	73.54	73.52	73.53	73.54
PSE	70.27	70.43	69.22	70.43	69.35	70.29	70.43
Non-PSE	76.77	76.63	77.80	76.63	77.68	76.75	76.63
Hold-Out	72.73	72.73	72.73	72.73	72.73	72.74	72.73
PSE	68.55	68.73	67.37	68.73	67.50	68.58	68.73
Non-PSE	76.86	76.69	78.03	76.69	77.89	76.85	76.69
YRMSE	70.71	82.66	84.13	82.66	81.18	79.57	82.62
$YRMSE_0$	31.72	36.74	37.41	36.74	36.05	35.45	36.73
PRMSE	1.38	1.46	1.74	1.46	2.79	1.39	1.46
$PRMSE_0$	0.63	0.66	0.79	0.66	1.19	0.63	0.66
$MSE(\hat{\beta})$	.05349	4.46490	5.32587	4.46490	3.79778	3.29101	4.45082
$RMSE(\hat{\beta})$	.23129	2.11303	2.30779	2.11303	1.94879	1.81412	2.10970
$VAR(\hat{\beta})$	.05324	.15664	.14616	.15664	.00520	.13313	.15636
$Bias^2(\hat{\beta})$	.00026	4.30826	5.17972	4.30826	3.79258	3.15789	4.29446
MSE ($\partial p / \partial x$)	.00669	.00759	.00728	.00759	.64692	.00752	.00756
RMSE ($\partial p / \partial x$)	.08182	.08713	.08534	.08713	.80431	.08673	.08698

*N=5,000; balanced; unrestricted; errors drawn from normal distribution.

Table 2.20 PSE Data - Monte Carlo MSE 22 (See Table 2.4 for design)

	Probit	Logit	D.A.	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.26	79.26	79.21	79.26	79.22	79.26	79.26
PSE	75.96	76.20	74.41	76.20	74.61	75.98	76.20
Non-PSE	82.44	82.22	83.84	82.22	83.67	82.43	82.22
Hold-Out	78.44	78.44	78.36	78.44	78.37	78.43	78.44
PSE	74.18	74.48	72.41	74.48	72.62	74.20	74.48
Non-PSE	82.50	82.23	84.03	82.23	83.85	82.48	82.23
YRMSE	73.18	107.51	109.82	107.51	79.08	98.14	107.39
$YRMSE_0$	32.66	47.20	48.31	47.20	35.06	43.23	47.15
PRMSE	1.88	1.50	2.03	1.50	5.36	1.86	1.50
$PRMSE_0$	0.84	0.68	0.91	0.68	2.29	0.82	0.68
$MSE(\hat{\beta})$	1.04610	15.9629	17.8376	15.9629	3.09457	11.4317	15.9039
$RMSE(\hat{\beta})$	1.02279	3.99536	4.22346	3.99536	1.75914	3.38107	3.98796
$VAR(\hat{\beta})$	.06646	.20404	.17101	.20404	.00432	.15561	.20337
$Bias^2(\hat{\beta})$	.97964	15.7589	17.6666	15.7589	3.09025	11.2760	15.7005
MSE ($\partial p / \partial x$)	.05477	.05759	.03267	.05759	.57529	.03196	.05725
RMSE ($\partial p / \partial x$)	.23402	.23997	.18075	.23997	.75848	.17878	.23927

*N=5,000; balanced; unrestricted; errors drawn from *t*- distribution.

Table 2.21 PSE Data - Monte Carlo MSE 23 (See Table 2.4 for design)

	Probit	Logit	D.A.	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	75.45	75.47	75.57	75.47	75.56	75.48	75.47
PSE	69.71	69.54	68.23	69.54	68.38	69.39	69.54
Non-PSE	80.49	80.69	82.02	80.69	81.88	80.83	80.69
Hold-Out	74.68	74.70	74.81	74.70	74.80	74.72	74.70
PSE	67.81	67.60	66.31	67.60	66.46	67.47	67.60
Non-PSE	80.64	80.86	82.17	80.86	82.03	81.01	80.87
YRMSE	71.41	89.98	93.92	89.98	80.60	85.42	89.93
$YRMSE_0$	32.01	39.89	41.66	39.89	35.74	37.97	39.87
PRMSE	33.18	33.00	33.04	33.00	34.71	33.20	33.00
$PRMSE_0$	14.15	14.07	14.09	14.07	14.75	14.15	14.07
$MSE(\hat{\beta})$	.14525	6.80643	8.32870	6.80643	3.61119	5.00014	6.78433
$RMSE(\hat{\beta})$	.38112	2.60891	2.88595	2.60891	1.90031	2.23610	2.60467
$VAR(\hat{\beta})$	.05470	.16849	.15567	.16849	.00484	.13920	.16813
$Bias^2(\hat{\beta})$	.09055	6.63794	8.17303	6.63794	3.60635	4.86094	6.61620
MSE ($\partial p / \partial x$)	.00754	.00788	.00821	.00788	.62967	.00900	.00788
RMSE ($\partial p / \partial x$)	.08685	.08878	.09062	.08878	.79352	.09487	.08874

*N=5,000; balanced; unrestricted; errors drawn from chi-square distribution.

2.8.5 Multinomial Choice Experiments with the PSE Data

In this section, we present Monte Carlo results for multinomial choice experiments using the PSE data. We compare the ME and GME estimators to the multinomial logit estimator and present results for both the prediction and MSE experiments. For the multinomial choice experiments we define the dependent variable as $y=0$ for no PSE, $y=1$ for a 2-year institution, $y=2$ for a 4-year institution, and $y=3$ for other PSE. The data include 2,133 (22.6%) individuals who did not attend PSE, 2,463 (26.1%) individuals who attended a 2-year institution, 4,608 (48.8%) individuals who attended a 4-year institution, and 246 (2.6%) individuals who attended other types of PSE. We carry out the multinomial choice experiments in the same manner as the binary choice experiments.

For the prediction experiments we only evaluate the percentage of correct predictions. We carry out experiments using sample sizes of $N = 5{,}000$ and $N = 2{,}000$ individuals. We do not impose parameter sign restrictions in the multinomial choice experiments since the parameters and the marginal effects do not necessarily take the same signs in the multinomial logit model (Greene 1997, p. 916). In addition, we found that parameter sign restrictions had little impact in the binary choice experiments. Thus, we have only two experimental designs for the multinomial choice prediction experiments. We assign individuals based on the highest predicted probability between the four alternatives.

Table 2.22 reports results for the multinomial choice prediction 1 experiment and Table 2.23 reports results for the multinomial choice prediction 2 experiment. We find very little difference in prediction between the alternative estimators. The only difference is that GME1 predicts slightly better for the 4-year and non-PSE groups and predicts much worse for the 2-year group. This is due to the shrinkage properties of the GME estimator. In a multinomial choice problem GME shrinks the predicted probabilities toward $1/J$, where J is the number of alternatives. Intuitively, when we specify wider error bounds, we place more weight on the errors and less on the data. Therefore, GME assigns more weight to the alternatives that occur most frequently. When we narrow the error bounds we place more weight on the data and the predicted probabilities are closer to logit. As in the binary choice problem, ME and logit are equivalent.

Table 2.22 PSE Data - Monte Carlo Multinomial Prediction 1 (N=5,000)

Estimation Sample	Mult. Logit	ME	GME1	GME2	GME3
Percent Correct	59.76	59.76	59.39	59.74	59.76
Standard Deviation	.0068	.0068	.0064	.0068	.0068
% Non-PSE Correct	59.78	59.78	62.68	59.77	59.77
% 2-Year Correct	17.87	17.87	8.56	17.38	17.87
% 4-Year Correct	85.29	85.29	88.18	85.54	85.30
% Other Correct	0.00	0.00	0.00	0.00	0.00
Hold-Out Sample					
Percent Correct	59.48	59.48	59.20	59.48	59.48
Standard Deviation	.0158	.0158	.0163	.0158	.0158
% Non-PSE Correct	59.17	59.17	62.17	59.15	59.16
% 2-Year Correct	17.66	17.66	8.41	17.17	17.66
% 4-Year Correct	85.17	85.17	88.11	85.43	85.17
% Other Correct	0.00	0.00	0.00	0.00	0.00

Table 2.23 PSE Data - Monte Carlo Multinomial Prediction 2 (N=2,000)

Estimation Sample	Mult. Logit	ME	GME1	GME2	GME3
Percent Correct	59.90	59.90	59.56	59.89	59.90
Standard Deviation	.0106	.0106	.0103	.0106	.0106
% Non-PSE Correct	59.68	59.68	61.74	59.63	59.67
% 2-Year Correct	18.64	18.64	10.13	18.19	18.64
% 4-Year Correct	85.16	85.16	88.04	85.39	85.17
% Other Correct	0.00	0.00	0.00	0.00	0.00
Hold-Out Sample					
Percent Correct	59.25	59.25	59.05	59.24	59.25
Standard Deviation	.0166	.0166	.0165	.0165	.0166
% Non-PSE Correct	59.01	59.01	61.13	58.97	59.00
% 2-Year Correct	17.65	17.65	9.48	17.20	17.65
% 4-Year Correct	84.83	84.83	87.82	85.07	84.83
% Other Correct	0.00	0.00	0.00	0.00	0.00

For the MSE experiments, we evaluate several empirical risk functions for the alternative estimators. The experiments are conducted in the same manner as the binary choice MSE experiments, described in section 2.7.3. We generate data during each Monte Carlo iteration using the multinomial logit parameter estimates from the original estimation sample as the true parameters. We carry out experiments using sample sizes of $N = 5{,}000$ and $N = 2{,}000$ individuals. As part of the experimental design, we carry out experiments with random errors drawn from standard normal, standardized t- and standardized chi-square distributions to examine the robustness of the estimators. We refer to the experiments as MSE11-MSE23.

Tables 2.24-2.26 give results for the Monte Carlo multinomial choice MSE 11-13 experiments with the PSE data. Tables 2.27-2.29 give results for the Monte Carlo multinomial choice MSE 21-23 experiments with the PSE data. Consistent with the binary choice results, we find that the GME2 and

Table 2.24 PSE Data - Monte Carlo Multinomial MSE 11

	Mult.Logit	ME	GME1	GME2	GME3
% Correctly Predicted					
In-Sample	65.22	65.22	65.00	65.22	65.22
Non-PSE	67.56	67.56	70.82	67.85	67.55
2-Year	19.42	19.42	9.06	18.46	19.42
4-Year	86.66	86.66	89.71	87.00	86.67
Other	0.01	0.01	0.00	0.00	0.00
Hold-Out	64.67	64.67	64.56	64.70	64.68
Non-PSE	64.80	64.80	68.93	65.17	64.78
2-Year	20.12	20.12	9.24	19.20	20.12
4-Year	86.54	86.54	89.69	86.87	86.55
Other	0.02	0.02	0.00	0.00	0.00
YRMSE	642.82	468.74	312.36	172.27	267.66
$YRMSE_0$	287.76	209.87	139.84	77.23	119.70
$MSE(\hat{\beta})$	329.64	134.08	16.85	7.16	18.39
$VAR(\hat{\beta})$	269.47	105.03	0.02	0.50	10.48
$Bias^2(\hat{\beta})$	60.17	29.04	16.83	6.66	7.91
$RMSE(\hat{\beta})$	18.16	11.58	4.10	2.68	4.29

*N=5,000; errors drawn from normal distribution.

Table 2.25 PSE Data - Monte Carlo Multinomial MSE 12

	Mult.Logit	ME	GME1	GME2	GME3
% Correctly Predicted					
In-Sample	70.09	70.09	69.64	70.09	70.09
Non-PSE	74.11	74.11	77.40	74.70	74.10
2-Year	20.64	20.64	8.89	19.08	20.64
4-Year	89.41	89.41	91.91	89.78	89.41
Other	0.00	0.00	0.00	0.00	0.00
Hold-Out	69.58	69.58	69.16	69.59	69.57
Non-PSE	71.88	71.88	75.94	72.58	71.87
2-Year	21.56	21.56	9.15	19.98	21.55
4-Year	89.23	89.23	91.80	89.61	89.23
Other	0.00	0.00	0.00	0.00	0.00
YRMSE	309.15	267.38	311.76	187.00	215.59
$YRMSE_0$	137.89	119.28	139.43	83.65	96.16
$MSE(\hat{\beta})$	84.27	49.78	16.09	12.24	23.42
$VAR(\hat{\beta})$	69.86	36.22	0.01	0.54	9.53
$Bias^2(\hat{\beta})$	14.41	13.57	16.08	11.70	13.89
$RMSE(\hat{\beta})$	9.18	7.06	4.01	3.50	4.84

*N=5,000; errors drawn from *t*- distribution.

Table 2.26 PSE Data - Monte Carlo Multinomial MSE 13

	Mult.Logit	ME	GME1	GME2	GME3
% Correctly Predicted					
In-Sample	65.01	65.01	64.70	65.01	65.02
Non-PSE	68.03	68.03	70.30	68.22	68.02
2-Year	16.45	16.45	7.56	15.63	16.45
4-Year	87.45	87.45	89.83	87.72	87.46
Other	0.00	0.00	0.00	0.00	0.00
Hold-Out	64.68	64.68	64.42	64.69	64.68
Non-PSE	65.95	65.95	68.85	66.16	65.94
2-Year	17.16	17.16	7.79	16.35	17.15
4-Year	87.34	87.34	89.76	87.62	87.34
Other	0.00	0.00	0.00	0.00	0.00
YRMSE	188.80	181.19	312.99	176.81	169.25
$YRMSE_0$	84.48	81.06	140.01	79.17	75.70
$MSE(\hat{\beta})$	15.38	11.36	16.98	4.89	7.41
$VAR(\hat{\beta})$	13.16	9.17	0.02	0.64	5.32
$Bias^2(\hat{\beta})$	2.23	2.19	16.97	4.25	2.09
$RMSE(\hat{\beta})$	3.92	3.37	4.12	2.21	2.72

*N=5,000; errors drawn from chi-square distribution.

Table 2.27 PSE Data - Monte Carlo Multinomial MSE 21

	Mult.Logit	ME	GME1	GME2	GME3
% Correctly Predicted					
In-Sample	64.98	64.98	64.69	65.00	64.99
Non-PSE	68.18	68.18	72.02	68.46	68.16
2-Year	25.27	25.27	14.63	24.45	25.27
4-Year	84.85	84.85	88.00	85.17	84.86
Other	0.03	0.03	0.00	0.00	0.00
Hold-Out	64.77	64.77	64.55	64.80	64.77
Non-PSE	67.54	67.54	71.30	67.87	67.53
2-Year	24.98	24.98	14.63	24.20	24.99
4-Year	84.90	84.90	88.11	85.21	84.91
Other	0.00	0.00	0.00	0.00	0.00
YRMSE	1006.21	627.75	205.86	116.36	192.33
$YRMSE_0$	719.08	448.33	145.98	82.78	136.26
$MSE(\hat{\beta})$	2713.69	971.84	24.83	8.38	52.49
$VAR(\hat{\beta})$	1631.19	587.89	0.03	1.05	33.03
$Bias^2(\hat{\beta})$	1082.50	383.95	24.80	7.33	19.46
$RMSE(\hat{\beta})$	52.09	31.17	4.98	2.89	7.24

*N=2,000; errors drawn from normal distribution.

Table 2.28 PSE Data - Monte Carlo Multinomial MSE 22

	Mult.Logit	ME	GME1	GME2	GME3
% Correctly Predicted					
In-Sample	69.99	69.99	69.42	69.99	69.98
Non-PSE	74.65	74.65	78.20	75.33	74.63
2-Year	27.70	27.70	16.04	26.26	27.69
4-Year	87.83	87.83	90.34	88.16	87.83
Other	0.02	0.02	0.00	0.00	0.01
Hold-Out	69.62	69.62	69.13	69.61	69.62
Non-PSE	74.04	74.04	77.63	74.68	74.03
2-Year	27.36	27.36	15.87	25.96	27.35
4-Year	87.67	87.67	90.30	88.00	87.68
Other	0.00	0.00	0.00	0.00	0.00
YRMSE	589.00	397.61	205.21	125.36	158.87
$YRMSE_0$	421.98	284.94	145.72	89.53	113.12
$MSE(\hat{\beta})$	1419.40	612.16	24.05	12.97	56.10
$VAR(\hat{\beta})$	1082.85	462.78	0.03	1.23	36.25
$Bias^2(\hat{\beta})$	336.55	149.38	24.02	11.74	19.85
$RMSE(\hat{\beta})$	37.67	24.74	4.90	3.60	7.49

*N=2,000; errors drawn from t- distribution.

Table 2.29 PSE Data - Monte Carlo Multinomial MSE 23

	Mult.Logit	ME	GME1	GME2	GME3
% Correctly Predicted					
In-Sample	64.95	64.95	64.54	64.95	64.95
Non-PSE	68.85	68.85	71.61	69.05	68.84
2-Year	22.61	22.61	13.24	21.81	22.60
4-Year	85.80	85.80	88.29	86.08	85.80
Other	0.00	0.00	0.00	0.00	0.00
Hold-Out	64.64	64.64	64.37	64.64	64.64
Non-PSE	67.84	67.84	70.64	68.08	67.82
2-Year	22.34	22.34	13.25	21.59	22.35
4-Year	85.73	85.73	88.33	85.99	85.74
Other	0.00	0.00	0.00	0.00	0.00
YRMSE	349.38	260.62	206.17	118.63	125.40
$YRMSE_0$	253.56	188.84	146.19	84.36	89.36
$MSE(\hat{\beta})$	486.88	248.09	25.01	6.77	26.98
$VAR(\hat{\beta})$	375.14	191.73	0.04	1.42	22.46
$Bias^2(\hat{\beta})$	111.74	56.36	24.97	5.36	4.52
$RMSE(\hat{\beta})$	22.07	15.75	5.00	2.60	5.19

*N=2,000; errors drawn from chi-square distribution.

GME3 estimators dominate the ME and logit estimators in terms of our empirical risk measures while the GME1 estimator outperforms ME and logit for many of the risk measures. The GME estimators have considerably lower variance and also have less bias than ME and logit, particularly for the smaller sample size. In addition, GME has lower MSE for predicting the latent values. Thus, in this set of experiments using real data we find that GME dominates multinomial logit estimation in a multinomial choice problem.

2.8.6 Response Surfaces for Monte Carlo Experiments

In this section, we estimate response surfaces for the binary choice Monte Carlo experiments with the used cars data. Hendry (1984) and Davidson and MacKinnon (1993) discuss response surfaces as a means of summarizing the results from a set of Monte Carlo experiments. A response surface is a

regression model with a measure of the outcome from the experiment as the dependent variable and dimensions of the experimental design as the independent variables. We estimate the following response surface regression

$$MSE(\hat{\beta}) = \alpha_1 + \alpha_2 N + \alpha_3 \text{stdt} + \alpha_4 \text{stdchi} + \alpha_5 \text{restrict} + \alpha_6 \text{unbal} + \mu_i, \qquad i = 1, \ldots, M \quad (2.44)$$

where N is the sample size for the estimation sample, stdt is a dummy variable that equals 1 when the errors are drawn from a standardized t- distribution, stdchi is a dummy variable that equals 1 when the errors are drawn from a standardized chi-square distribution, restrict is a dummy variable that equals 1 when we place inequality restrictions on the parameter estimates, unbal is a dummy variable that equals 1 when we have an unbalanced sample, M is the number of Monte Carlo experiments, and $MSE(\hat{\beta})$ is the MSE of the estimator. We also estimate a response surface with $MSE(\hat{\beta}) / MSE(probit)$ as the dependent variable. Finally, we estimate response surfaces using the MSE of the marginal effects as the dependent variable.

For each of the response surface regressions we have $M = 36$ observations corresponding to the dimensions of our Monte Carlo experiments, which are summarized in Table 2.4. We estimate the response surface regressions using OLS. Table 2.30 summarizes the response surface estimates with $MSE(\hat{\beta})$ as the dependent variable and Table 2.31 summarizes the response surface estimates with $MSE(\hat{\beta}) / MSE(probit)$ as the dependent variable. Tables 2.32 and 2.33 summarize the results for response surfaces using the MSE of the marginal effects rather than the MSE of the parameter estimates.

The response surface estimates show that the probit, ME, GME2, and GME3 risk measures increase when we draw errors from a standardized t- or standardized chi-squared distribution. The sample size and parameter inequality restrictions do not have a significant impact on the risk measures of the alternative estimators. Unbalanced samples increase the MSE of the parameter estimates, but decrease the MSE of the marginal effects. This is not surprising though since the marginal effects are maximized for probabilities near 0.5.

Table 2.30 Response Surfaces for MSE ($\hat{\beta}$)

Variable	Probit	ME	GME1	GME2	GME3
Intercept	.0294**	-.0587	.0210	.0182	-.0574
	(.0118)	(.1907)	(.0194)	(.1167)	(.1895)
Sample Size	-.0000	-.0000	.0000**	-.0000	-.0000
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Stdt	.0732**	.9626**	-.0490**	.6277**	.9574**
	(.0116)	(.1860)	(.0189)	(.1139)	(.1848)
Stdchi	.0293**	-.0034	.0090	.0231	-.0030
	(.0116)	(.1860)	(.0189)	(.1139)	(.1848)
Restrict	-.0043	-.0230	.0004	-.0158	-.0229
	(.0094)	(.1518)	(.0154)	(.0930)	(.1509)
Unbal	.0182	1.3062**	.7494**	.8192**	1.2993**
	(.0094)	(.1518)	(.0154)	(.0930)	(.1509)
R^2	0.64	0.79	0.99	0.80	0.79
$\hat{\sigma}^2$	0.008	0.208	0.002	0.078	0.205

* M =36; standard errors in parentheses.

Table 2.31 Response Surfaces for MSE ($\hat{\beta}$) / MSE(Probit)

Variable	ME	GME1	GME2	GME3
Intercept	7.66	2.81	6.95	7.66
	(14.39)	(12.16)	(9.86)	(14.33)
Sample Size	.0024**	.0021**	.0017**	.0024**
	(.0007)	(.0006)	(.0005)	(.0007)
Stdt	-37.29**	-36.74**	-26.62**	-37.16**
	(14.03)	(11.86)	(9.62)	(13.98)
Stdchi	-31.26**	-21.40	-21.96**	-31.14**
	(14.03)	(11.86)	(9.62)	(13.98)
Restrict	1.35	1.22	0.93	1.34
	(11.46)	(9.68)	(7.85)	(11.41)
Unbal	42.44**	35.68**	27.74**	42.25**
	(11.46)	(9.68)	(7.85)	(11.41)
R^2	0.52	0.53	0.52	0.52
$\hat{\sigma}^2$	1181.69	844.15	555.29	1172.37

* M =36; standard errors in parentheses.

Table 2.32 Response Surfaces for MSE ($\partial p / \partial x$)

Variable	Probit	ME	GME1	GME2	GME3
Intercept	.0056**	.0050**	.0093**	.0048**	.0050**
	(.0008)	(.0009)	(.0022)	(.0008)	(.0009)
Sample Size	-.0000**	-.0000**	-.0000	-.0000**	-.0000**
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Stdt	.0043**	.0062**	-.0032	.0048**	.0062**
	(.0008)	(.0009)	(.0022)	(.0008)	(.0009)
Stdchi	.0040**	.0044**	-.0002	.0040**	.0044**
	(.0008)	(.0009)	(.0022)	(.0008)	(.0009)
Restrict	-.0001	-.0001	-.0000	-.0001	-.0001
	(.0007)	(.0007)	(.0018)	(.0006)	(.0007)
Unbal	-.0047**	-.0029**	.0260**	-.0031**	-.0029**
	(.0007)	(.0007)	(.0018)	(.0006)	(.0007)
R^2	0.76	0.73	0.88	0.74	0.73
$\hat{\sigma}^2$	0.00	0.00	0.00	0.00	0.00

* M =36; standard errors in parentheses.

Table 2.33 Response Surfaces for MSE ($\partial p / \partial x$) / MSE(Probit)

Variable	ME	GME1	GME2	GME3
Intercept	.6985**	5.74	.6944**	.6997**
	(.1915)	(10.18)	(.1709)	(.1915)
Sample Size	.0000**	.0018**	.0000**	.0000**
	(.0000)	(.0005)	(.0000)	(.0000)
Stdt	-.0615	-32.85**	-.2079	-.0650
	(.1867)	(9.93)	(.1667)	(.1868)
Stdchi	-.0740	-22.30**	-.0021	-.0749
	(.1867)	(9.93)	(.1667)	(.1868)
Restrict	.0314	0.98	.0283	.0308
	(.1525)	(8.11)	(.1361)	(.1525)
Unbal	1.1198**	29.43**	.8164**	1.1151**
	(.1525)	(8.11)	(.1361)	(.1525)
R^2	0.71	0.55	0.64	0.70
$\hat{\sigma}^2$	0.21	591.94	0.17	0.21

* M =36; standard errors in parentheses.

Relative to the probit estimator, we find that risk measures for the GME estimators increase as we increase the sample size. Thus, GME performs relatively better when we have a small sample size. Additionally, we find that GME risk measures decrease relative to probit when we draw errors from standardized t- or standardized chi-squared distributions. This makes sense because the probit model assumes normal errors. Finally, GME is more sensitive to the unbalanced sample than probit. This is because shrinkage is greater for GME the farther the predicted probabilities are from 0.5. Thus, the parameter estimates become closer to zero the more unbalanced the sample since we place more weight on the errors.

2.9 *Conclusions*

We examine both ME and GME estimation in binary and multinomial choice models using actual data as a basis for Monte Carlo sampling experiments. We find that ME is equivalent to logit when we specify the ME data constraint following Soofi (1992) and Golan, Judge, and Perloff (1996). We show that GME shrinks the parameter estimates toward zero and the predicted probabilities toward 0.5 and that the degree of shrinkage is greater the wider we specify the error bounds. In addition, shrinkage is greater when we have an unbalanced sample with many predicted probabilities far from 0.5. For balanced samples, we observe very little difference in prediction between the alternative estimators.

GME performs well within Monte Carlo sampling experiments in terms of risk measures for the parameter estimates, the predicted latent variables, the predicted probabilities, and the marginal effects. The GME3 estimator dominates ME-logit over all of the risk measures for our Monte Carlo experiments. The GME2 estimator also dominates ME-logit in our experiments and the risk gains are greater than for GME3. We expected the GME estimators to have lower variance than ME-logit, however the we find that the bias is also lower for GME. In addition, the MSE for the marginal effects is lower for GME2 and GME3 than for ME-logit. Therefore, based on our sampling experiments we conclude that GME estimation is a viable alternative to logit estimation, particularly when we have a small or an unbalanced sample where the shrinkage properties of GME are greater.

2.10 References

Adkins, L. (1997). A monte carlo study of a generalized maximum entropy estimator of the binary choice model. In Fomby, T. B., R. C. Hill (Eds.) *Advances in Econometrics*, Vol. 12. Greenwich, CT: JAI Press, 189-197.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance* 23(4): 589-609.

Belsley, D., E. Kuh, R. E. Welsch. (1980). *Regression diagnostics*. New York: John Wiley and Sons.

Boyes, W. J., D. L. Hoffman, S. A. Low. (1989). An econometric analysis of the bank credit scoring Problem, *Journal of Econometrics* 40: 3-14.

Davidson, R., J.G. MacKinnon. (1993). *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Denzau, A. T., P. C. Gibbons, E. Greenberg. (1989). Bayesian estimation of proportions with a cross-entropy prior, *Communications in Statistics-Theory and Methods* 18: 1843-1861.

Efron, B. (1975). The efficiency of logistic regression compared to normal discriminant analysis, *Journal of the American Statistical Association* 70: 892-898.

Fisher, R.A. (1936). The use of multiple measurement in taxonomic problems, *Annals of Eugenics* 7: 179-188.

Golan, A., G. Judge, D. Miller. (1996). *Maximum entropy econometrics: Robust estimation with limited data*. New York: John Wiley and Sons.

Golan, A., G. Judge, J. M. Perloff. (1996). A maximum entropy approach to recovering information from multinomial response data, *Journal of the American Statistical Association* 91 (June): 841-853.

________. (1997). Estimation and inference with censored and ordered multinomial response data, *Journal of Econometrics* 79: 23-51.

Greene, W.H. (1997). *Econometric Analysis*. New Jersey: Prentice-Hall.

Hendry, D.F. (1984). Monte Carlo experimentation in econometrics. In Z. Griliches, M.D. Intriligator (Eds.) *Handbook of Econometrics*, Vol. II. Amsterdam, North-Holland: Elsevier Science Publishers, 937-976.

Jaynes, E. T. (1957a). Information theory and statistical mechanics, *Physical Review* 106: 620-630.

________. (1957b). Information theory and statistical mechanics II, *Physical Review* 108: 171-190.

Knapp, L. G., T. G. Seaks. (1992). An analysis of the probability of default on federally guaranteed student loans, *Review of Economics and Statistics* 74: 404-411.

Lawrence, E. C., L. D. Smith, M. Rhoades. (1992). An analysis of default risk in mobile home credit, *Journal of Banking and Finance* 16: 299-312.

Lehmann, E. (1983). *Theory of Point Estimation.* New York: John Wiley and Sons.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics.* New York: Cambridge University Press.

Manski, C. F., S. R. Lerman. (1977). The estimation of choice probabilities from choice based samples, *Econometrica* 45 (8): 1977-1988.

Martin, R. E., R. C. Hill. (1999). *Economic Inquiry*, forthcoming.

Myers, J. H., E. W. Forgy. (1963). The development of numerical credit evaluation systems, *Journal of the American Statistical Association* 58: 799-806.

U.S. Department of Education, Office of Educational Research and Improvement. (1996). National education longitudinal study: 1988-94 data files and electronic codebook system, base year through third followup ECB/CD-ROM, 1996. National Center for Educational Statistics.

Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Technical Journal* 27: 379-423.

Soofi, E. S. (1992). A generalizable formulation of conditional logit with diagnostics, *Journal of the American Statistical Association* 87(Sept.): 812-816.

Steenackers, A., M. J. Goovaerts. (1989). A credit scoring model for personal loans, *Mathematics and Economics* 8: 31-34.

CHAPTER 3

MONTE CARLO EXPERIMENTS COMPARING THE GME AND OLS ESTIMATORS IN A LINEAR REGRESSION MODEL

3.1 Introduction

In this chapter we use the generalized maximum entropy (GME) estimator developed by Golan, Judge, and Miller (1996) to estimate linear regression models. We carry out several Monte Carlo experiments comparing the GME estimator to the ordinary least squares (OLS) estimator. We compare the estimators on the basis of the mean squared error (MSE) of the parameter estimates and the MSE of the predicted values. Golan et al. discuss both asymptotic and finite sample properties of the GME estimator and show that the GME estimator is consistent and asymptotically normal. Additionally, in Monte Carlo sampling experiments using artificial data, they find that although the GME estimator is biased it has lower variance and MSE than the OLS estimator over most of the parameter space. We conduct several Monte Carlo experiments using actual data as a basis. We discuss ME estimation in the general linear model (GLM) in section 3.2, GME estimation in the GLM and shrinkage properties of the GME estimator in section 3.3, describe our sampling experiments in section 3.4, and present results in section 3.5.

3.2 Maximum Entropy Estimation in Linear Regression Models

Golan, Judge, and Miller (1996) develop a maximum entropy formulation to estimate the unknown parameters in a linear regression model. Since the unknown parameters in the linear regression model are not in the form of probabilities, Golan et al. reparameterize the model in a manner similar to the error transformation of the GME model for the binary choice problem described in Chapter 2. Thus, in using ME to solve linear regression problems, we must specify a matrix of support points that bounds the unknown parameters and errors *a priori*. This allows us to combine both sample and nonsample information as in the restricted least squares (RLS), inequality restricted least squares

(IRLS), and Stein-rule estimators. However, the ME estimator is more similar to Bayesian inference since we must specify a prior distribution for the unknown parameters.

We carry out Monte Carlo experiments specifying different support matrices for the unknown parameters and errors. If we do not have good prior information, we can specify very wide parameter bounds to ensure that the true parameters lie within these bounds. This is analogous to the noninformative prior in Bayesian inference. In Chapter 4, we examine the cost of specifying a parameter support that does not contain the true parameters, which are known in our sampling experiments.

Golan, Judge, and Miller (1996, p. 88) specify four different entropy formulations to estimate the unknown parameters in the GLM. The first of these, the ME formulation, does not include the error term in the entropy function or its constraints. The GME formulations, which we discuss in section 3.3, do include the error term in the entropy function and its constraints. Excluding the error term the GLM is written as

$$y_t = x_t'\beta \tag{3.1}$$

or, in matrix notation as

$$y = X\beta, \tag{3.2}$$

where y is an $N \times 1$ vector of sample observations, X is an $N \times K$ matrix of explanatory variables, and β is a $K \times 1$ vector of unknown parameters. Jaynes (1957a, 1957b) shows that ME allows us to estimate the unknown probabilities in a discrete probability distribution. Therefore, Golan, Judge, and Miller (1996) reparameterize the GLM such that the unknown parameters are in the form of probabilities, which we can estimate through ME. To do so they assume that the parameters may be bounded *a priori*. We must specify a parameter support matrix in order to obtain entropy estimates for the unknown parameters in the GLM. Let z_{k1} be the smallest possible value of β_k and z_{k2} be the largest possible value of β_k. Then, for each parameter, β_k, there exists $p_k \in [0,1]$ such that

$$\beta_k = p_k z_{k1} + (1-p_k) z_{k2} = \begin{bmatrix} z_{k1} & z_{k2} \end{bmatrix} \begin{bmatrix} p_k \\ 1-p_k \end{bmatrix}. \tag{3.3}$$

The parameter support is based on prior information or economic theory. For example, we would specify boundaries of $z_{k1} = 0$ and $z_{k2} = 1$ when estimating the marginal propensity to consume. However, specifying the largest and smallest possible values for each variable is not an easy task since economic theory does not usually provide this information.

Golan, Judge, and Miller (1996) define a set of $M \geq 2$ support points, which may or may not be symmetric about zero and which bound the unknown parameters. Let z_k represent the $M \times 1$ support vector for the k^{th} parameter and let p_k represent the associated $M \times 1$ vector of weights on these support points. The parameter vector, β, is expressed as

$$\beta = Zp = \begin{bmatrix} z_1' & 0 & \cdots & 0 \\ 0 & z_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_K' \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_K \end{bmatrix}, \qquad k = 1,2,\ldots,K \tag{3.4}$$

where β is a $K \times 1$ vector of unknown parameters, Z is a $K \times KM$ matrix of support points, and p is a $KM \times 1$ vector of unknown weights such that $p_{km} > 0$ and $p_k' i_M = 1$ for all k. Golan et al. (1996, p. 140) suggest that we do not gain much precision in our estimates when using more than $M = 5$ support points. The vector of support points may vary for each parameter. We write the reparameterized model as

$$y = XZp, \tag{3.5}$$

where y, X, and Z are known and we obtain the ME parameter estimates through the equation

$$\hat{\beta} = Z\hat{p}, \tag{3.6}$$

where $\hat{p}$ is the estimated vector of weights for the parameter support. Golan et al. assume that Z is block diagonal. That is, the support points for any one parameter do not directly affect any of the other parameter estimates. We develop a support matrix that is not block diagonal, which allows us to specify sign and other inequality restrictions on the parameters.

The ME probabilities for the linear regression problem are estimated by solving the constrained optimization problem

$$\max H(p) = -p' \ln(p) \tag{3.7}$$

subject to

$$y = XZp \tag{3.8}$$

$$(I_K \otimes i'_M)p = i_K. \tag{3.9}$$

Equation (3.8) is the data constraint and equation (3.9) is the additivity constraint, which requires that the probabilities must sum to one for each of the K parameters. The Lagrangian for the ME linear regression problem is

$$\mathcal{L} = -p' \ln(p) + \lambda'(XZp - y) + \gamma'[i_K - (I_K \otimes i'_M)p],$$

where λ is an $N \times 1$ vector of Lagrange multipliers used to obtain the optimal probabilities and γ is a $K \times 1$ vector of Lagrange multipliers for the additivity constraint on the unknown probabilities. The FOC's for the ME problem are

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{KM} - \ln(p) + Z'X'\lambda - (I_K \otimes i_M)\gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = XZp - y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_K - (I_K \otimes i'_M)p = 0.$$

Solving the KM FOC's for the unknown p's yields

$$\hat{p} = \exp(Z'X'\hat{\lambda})\exp[-i_{KM} - (I_K \otimes i_M)\gamma],$$

which implies that

$$\hat{p}_{km} = \exp(z_{km}x'_k\hat{\lambda})\exp(-1 - \gamma_k),$$

where x_k is the $N \times 1$ vector of observations for the k^{th} explanatory variable. Since the additivity constraint requires that $\sum_{m=1}^{M} p_{km} = 1$ and $\exp(-1 - \gamma_k)$ is constant for a given parameter, the optimal probabilities may be rewritten as

$$\hat{p}_{km} = \frac{\exp(z_{km} x_k' \hat{\lambda})}{\sum_{m=1}^{M} \exp(z_{km} x_k' \hat{\lambda})} = \frac{\exp(z_{km} x_k' \hat{\lambda})}{\Omega_k(\hat{\lambda})}, \tag{3.10}$$

where $\Omega_k(\hat{\lambda})$ represents the partition function for the k^{th} parameter. Taking the second derivative of the Lagrangian with respect to p shows that the Hessian for the ME linear regression problem is negative definite, and thus ensures a unique global solution to the entropy maximization problem.

In the dice problem, described in Chapter 1, we use ME to estimate the unknown probabilities given a set of known support points and a known prior (the observed average from previous rolls). In discrete choice problems, described in Chapter 2, we use ME to estimate the probability of each individual choosing each of J alternatives. We have a known support in discrete choice models and specify a data constraint that links the unknown parameters to the observed data. In the linear regression model, we use ME to estimate the probability that a parameter is equal to each of the M support points for that parameter. In the linear regression problem we do not observe a prior for the unknown parameters nor do we know the parameter support points. Instead, we specify a support matrix based on prior information or economic theory and we specify a data constraint involving the unknown probabilities and the observed dependent and independent variables. Thus, in the linear regression model, the estimated probabilities are a function of the explanatory variables, the vector of Lagrange multipliers for the data constraint, and the support points placed on the parameters *a priori.*

As discussed in Chapter 1, we can solve the ME problem as an unconstrained optimization problem by substituting the optimal probabilities into the original Lagrangian. For $\lambda \in \Re$, let $\hat{p}(\lambda)$ represent the functional form of the optimal ME probabilities, (3.10). Since the optimal probabilities satisfy the additivity constraint, (3.9), this term drops out of the concentrated Lagrangian, which is written as

$$\begin{aligned} M(\lambda) &= -\hat{p}'(\lambda)\ln(\hat{p}(\lambda)) + \lambda'[XZ\hat{p}(\lambda) - y] \\ &= -\hat{p}'(\lambda)\ln\left[\frac{\exp(z_{km} x_k' \lambda)}{\sum_{m} \exp(z_{km} x_k' \lambda)}\right] + \lambda'[XZ\hat{p}(\lambda) - y] \end{aligned}$$

$$= -\hat{p}'(\lambda)Z'X'\lambda + \sum_k \ln[\Omega_k(\lambda)] + \hat{p}'(\lambda)Z'X'\lambda - y'\lambda$$

$$= -y'\lambda + \sum_k \ln[\Omega_k(\lambda)].$$

We minimize the concentrated Lagrangian to obtain $\hat{\lambda}$, which we substitute into (3.10) to obtain the optimal ME probabilities. We then substitute the optimal probabilities into (3.6) to obtain the ME parameter estimates. The dual formulation allows us to obtain the ME parameter estimates by estimating the *N* unknown $\lambda's$ rather than the *KM* unknown probabilities.

In practice, there is no vector of parameters, β, that satisfies equation (3.2) and therefore the entropy function does not converge under the ME formulation. Equation (3.2) requires that the errors equal zero for all observations, which implies that the sum of squared errors (SSE) equals zero. Since ME cannot satisfy this constraint, it gets as close as possible by minimizing the SSE subject to the constraints. Thus, the ME and OLS parameter estimates are identical as long as the OLS estimates are contained in the parameter support. Therefore, we use the GME formulation in our Monte Carlo experiments.

3.3 Generalized Maximum Entropy Estimation in Linear Regression Models

Golan, Judge, and Miller (1996, Ch. 6) jointly estimate the unknown parameters and the unknown errors in the GLM using generalized maximum entropy (GME). They specify three different GME formulations; we use the GME-D formulation (Golan et al., p. 88), which estimates β using the traditional GLM as the data constraint. We write the GLM in matrix form as

$$y = X\beta + e, \tag{3.11}$$

where e is an $N \times 1$ vector of unknown errors. The alternative GME formulations are transformations of the GME-D formulation, which we focus on. As they do with the parameters, Golan et al. assume that the errors may be bounded *a priori*. Let v_{i1} be the smallest possible value of e_i and v_{i2} be the largest possible value of e_i. Then for each random error, e_i, there exists $w \in [0,1]$ such that

$$e_i = w_i v_{i1} + (1 - w_i)v_{i2} = \begin{bmatrix} v_{i1} & v_{i2} \end{bmatrix} \begin{bmatrix} w_i \\ 1 - w_i \end{bmatrix}. \tag{3.12}$$

However we do not know the error bounds in the linear regression model as we do in discrete choice models and placing boundaries on the unknown errors may be difficult in practice.

Following Pukelsheim (1994), Golan, Judge, and Miller (1996) suggest setting the error bounds as $v_{i1} = -3\sigma$ and $v_{i2} = 3\sigma$, where σ is the standard deviation of e. However, to use this rule we must know or estimate σ. We further discuss the error support in section 3.4, where we describe our Monte Carlo experiments. In our Monte Carlo experiments we define $\sigma = 1$. However, following Golan, Judge, and Perloff (1997) we calculate the variance of y assuming it follows a uniform distribution between its minimum and maximum values, $s_y^2 = (y_{\max} - y_{\min})^2 / 12$, and use s_y in setting our error bounds. We use this method to generate a series of error bounds and examine which set performs best in our experiments.

Golan, Judge, and Miller (1996) define a set of $J \geq 2$ support points, symmetric about zero, which bound the random errors, and w_t is the associated $J \times 1$ vector of weights on these points. The error vector is expressed as

$$e = Vw = \begin{bmatrix} v_1' & 0 & \cdots & 0 \\ 0 & v_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_N' \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}, \qquad n = 1,2,\ldots,N \tag{3.13}$$

where e is an $N \times 1$ vector of random errors, V is an $N \times NJ$ matrix of support points, and w is an $NJ \times 1$ vector of unknown weights such that $w_{ij} > 0$ and $w_i' i_J = 1$ for all i. Following Golan et al., we choose $J = 3$ support points, which are symmetric about zero, and use the same set of support points for each e_t. We write the reparameterized model in matrix notation as

$$y = XZp + Vw, \tag{3.14}$$

where y, X, Z, and V are known and we estimate the unknown p and w vectors through GME. We obtain the GME error estimates through the equation

$$\hat{e} = V\hat{w}, \tag{3.15}$$

where $\hat{w}$ is the estimated vector of weights for the unknown errors.

We assume that the unknown weights on the parameter and the error supports for the GLM are independent and estimate them jointly by solving the constrained optimization problem

$$\max H(p,w) = -p'\ln(p) - w'\ln(w) \tag{3.16}$$

subject to

$$y = XZp + Vw \tag{3.17}$$

$$(I_K \otimes i'_M)p = i_K \tag{3.18}$$

$$(I_N \otimes i'_J)w = i_N\,. \tag{3.19}$$

Equation (3.17) is the data constraint and equations (3.18) and (3.19) are the additivity constraints, which require that the probabilities sum to one. The Lagrangian for the GME problem is

$$\mathcal{L} = -p'\ln(p) - w'\ln(w) + \lambda'(XZp + Vw - y) + \gamma'\left[i_K - (I_K \otimes i'_M)p\right] + \delta'\left[i_N - (I_N \otimes i'_J)w\right],$$

where λ is an $N \times 1$ vector of Lagrange multipliers for the data constraint, γ is a $K \times 1$ vector of Lagrange multipliers for the additivity constraint on the unknown probabilities for the parameter support, and δ is an $N \times 1$ vector of Lagrange multipliers for the additivity constraint on the unknown error weights. The FOC's for the GME problem are

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{km} - \ln(p) + Z'X'\lambda - (I_K \otimes i_M)\gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial w} = -i_{NJ} - \ln(w) + V'w - (I_N \otimes i_J)\delta = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = XZp + Vw - y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_K - (I_K \otimes i'_M)p = 0$$

$$\frac{\partial \mathcal{L}}{\partial \delta} = i_N - (I_N \otimes i'_J)w = 0.$$

The KM FOC's for the p's are the same as in the ME formulation, so we again obtain the solution (3.10) for the optimal weights on the parameter support. Solving the NJ FOC's for the w's yields

which implies that

$$\hat{w}_{nj} = \exp(v_{nj}\hat{\lambda}_n)\exp(-1-\delta_n).$$

Since the addivity constraint requires that $\sum_{j=1}^{J} w_{nj} = 1$ and $\exp(-1-\delta_n)$ is constant for a given error, we rewrite the optimal error probabilities as

$$\hat{w}_{nj} = \frac{\exp(v_{nj}\hat{\lambda}_n)}{\sum_{j=1}^{J}\exp(v_{nj}\hat{\lambda}_n)} = \frac{\exp(v_{nj}\hat{\lambda}_n)}{\Psi_n(\hat{\lambda}_n)}, \tag{3.20}$$

where $\Psi_n(\hat{\lambda}_n)$ represents the partition function for the n^{th} error. Thus, the unknown error weights are a function of the Lagrange multipliers for the data constraint and the support points placed on the errors *a priori*.

To obtain the unconstrained GME dual solution to the linear regression problem, we let $\hat{p}(\lambda)$ represent the functional form of the optimal GME probabilities, (3.10), and let $\hat{w}(\lambda)$ represent the functional form of the optimal error probabilities, (3.20). Since the additivity constraints, (3.18) and (3.19), are satisfied by the optimal probabilities these terms drop out of the original Lagrangian and the concentrated Lagrangian is written as

$$M(\lambda) = -\hat{p}'(\lambda)\ln(\hat{p}(\lambda)) - \hat{w}'(\lambda)\ln(\hat{w}(\lambda)) + \lambda'[XZ\hat{p}(\lambda) + V\hat{w}(\lambda) - y]$$

$$= -\hat{p}'(\lambda)\ln\left[\frac{\exp(z_{km}x_k'\lambda)}{\sum_m \exp(z_{km}x_k'\lambda)}\right] - \hat{w}'(\lambda)\ln\left[\frac{\exp(v_{nj}\lambda_n)}{\sum_j \exp(v_{nj}\lambda_n)}\right]$$

$$+\lambda'[XZ\hat{p}(\lambda) + V\hat{w}(\lambda) - y]$$

$$= -y'\lambda + \sum_k \ln[\Omega_k(\lambda)] + \sum_n \ln[\Psi_n(\lambda_n)],$$

which we minimize to obtain $\hat{\lambda}$. This vector is then substituted into (3.10) and (3.20) to obtain the optimal GME probabilities. The optimal probabilities are then substituted into (3.6) to obtain the GME parameter estimates. The GME error estimates are obtained through equation (3.15). The dual

formulation allows us to obtain GME parameter estimates by estimating the N unknown λ's rather than the KM unknown parameter weights and the NJ unknown error weights.

3.3.1 Shrinkage Properties of the GME Estimator

In Chapter 2, we show that the GME estimator in binary choice models shrinks the predicted probabilities towards 0.5. GME selects the most uniform probabilities compatible with the observed data, and by estimating the error term along with the unknown probabilities we allow the predicted probabilities to be more uniform. The GME estimator in linear regression problems is also a shrinkage estimator, although it works differently than in the binary choice problem. In the linear regression problem, the GME estimator is similar to the Stein-like and empirical Bayes estimators described by Hill, Cartwright, and Arbaugh (1991) since GME allows us to introduce nonsample information into the estimation problem.

GME selects the most uniform probability distribution compatible with the constraints, which are based on prior information. GME shrinks the parameter estimates towards the expected value of the parameter support, which is specified *a priori*. Since the expected value of the parameter support is equal to the sum of the support points multiplied by the associated prior distribution it is known as the prior mean of the unknown parameters. For example, suppose we specify a parameter support that is symmetric about zero. If the prior probability distribution is uniform, then the prior mean of the parameter is equal to zero (since $\hat{\beta}_k = z_k' \hat{p}_k$). Since GME selects the probability distribution that is most uniform yet compatible with the prior information, GME shrinks each parameter estimate toward its prior mean, in this case zero. The GME parameter estimates are a function of the parameter and error support matrices, which we specify, based on nonsample information. In our Monte Carlo experiments, described in section 3.4, we vary the parameter bounds to shrink the parameters towards different values representing different nonsample information.

3.4 Monte Carlo Design

We carry out several Monte Carlo experiments to compare the GME estimator to the traditional least squares estimator. We evaluate the estimators in terms of the MSE of the parameter estimates and

the MSE of the predicted values. In this section we describe the data, describe the experimental design, and then discuss variations in the experimental design.

3.4.1 Data

For our linear regression experiments, we use sales data for a particular brand of tuna fish. Hill, Cartwright, and Arbaugh (1991) use Stein-like rules to estimate a 'price-promotion' model with the tuna data. The data consist of observations on sales, prices, and promotions for four brands of canned tuna fish. We have 52 weeks of data for three supermarket chains, or 156 weekly observations. Following Hill et al., we specify a semi-log functional form and our model is

$$\begin{aligned} S_{1tl} = \exp\{&\beta_1 + \beta_2 R_{1tl} + \beta_3 (R_{1tl} - P_{1tl}) / R_{1tl} + \beta_4 DIS_{1tl} + \beta_5 DISMAD_{1tl} \\ &+\beta_6 R_{2tl} + \beta_7 (R_{2tl} - P_{2tl}) / R_{2tl} + \beta_8 DIS_{2tl} + \beta_9 DISMAD_{2tl} \\ &+\beta_{10} R_{3tl} + \beta_{11} (R_{3tl} - P_{3tl}) / R_{3tl} + \beta_{12} DIS_{3tl} + \beta_{13} DISMAD_{3tl} \\ &+\beta_{14} R_{4tl} + \beta_{15} (R_{4tl} - P_{4tl}) / R_{4tl} + \beta_{16} DIS_{4tl} + \beta_{17} DISMAD_{4tl}\}. \end{aligned} \tag{3.21}$$

The independent variables are defined as

S_{itl} = unit sales (cans of tuna), $i = 1$ for target brand,

R_{itl} = regular price, $i = 2,3,4$ for competitive brands,

P_{itl} = actual price, $t = 1,\ldots,52$ weeks, $l = 1,2,3$ chain,

$(R_{itl} - P_{itl}) / R_{itl}$ = deal discount (always ≥ 0),

DIS_{itl} = display only indicator,

$DISMAD_{itl}$ = display and major ad indicator.

The regular price (R_{itl}) is an imputed value designed to reflect the product price in the absence of any discounts or promotions. The deal discount $((R_{itl} - P_{itl}) / R_{itl})$ is the non-negative percentage off the regular price. Displays are retailer promotions in the store and major ads are retailer promotions in newspapers.

It is well known that multicollinearity is a problem in marketing models. We test for multicollinearity using all 156 weekly observations. Belsley, Kuh, and Welsch (1980, pp. 100-104) define the condition number of the $X'X$ matrix, denoted $\kappa(X'X)$, to be $\sqrt{\lambda_1 / \lambda_K}$ where λ_1 is the largest characteristic root and λ_K is the smallest characteristic root and the regressors have been normalized to unit length. They conclude that for condition numbers greater than about 30 multicollinearity is a problem in the model. We find the maximum condition number for the tuna data to be $\kappa(X'X) = 132.7$, which indicates that multicollinearity is a problem in our model. Examination of the characteristic vectors shows that the collinearity is among the regular prices of the different brands, particularly brands 2 and 3. Golan, Judge, and Miller (1996) find that GME has lower risk in terms of squared error loss than OLS does in sampling experiments using artificial data. The GME estimator performs especially well when the data are collinear. Therefore, the marketing model using the tuna data provides a good empirical test of the GME estimator.

3.4.2 Monte Carlo Experiments

In our Monte Carlo experiments, we are interested in both the MSE of the parameter estimates and the MSE of prediction. We use generated data for the Monte Carlo iterations. We draw a random sample, X, of 150 observations from the 156 total observations. We use this data to create an estimation sample. We also draw a random sample, X_0, of 50 observations which we use to create a hold-out sample. We choose β_{true} as the true parameter vector. We generate new data during each Monte Carlo iteration using the true parameter values as

$$y = X\beta_{true} + e \qquad e \sim N(0,1)$$

and

$$y_0 = X_0\beta_{true} + e_0 \qquad e_0 \sim N(0,1),$$

where y is an $N \times 1$ vector of generated observations, X is an $N \times K$ matrix of explanatory variables, y_0 is an $N_0 \times 1$ vector of generated observations for a hold-out sample, X_0 is an $N_0 \times K$ matrix of explanatory variables for the hold-out sample, and e and e_0 are vectors of random errors drawn from a standard normal distribution. As part of the experimental design, we also carry out

experiments with errors drawn from standardized t- and standardized chi-square distributions to examine the robustness of the estimators.

During each Monte Carlo iteration, we estimate the model using the generated data with the alternative estimation techniques (OLS and GME). We use the parameter estimates to predict the dependent variable in-sample as

$$\hat{y} = X\hat{B} \tag{3.22}$$

and out-of-sample as

$$\hat{y}_0 = X_0\hat{\beta}, \tag{3.23}$$

where $\hat{\beta}$ is the estimated parameter vector using either OLS or GME. [This Monte Carlo step is repeated $NSAM = 500$ times.] For GME, we also calculate $\hat{y} = X\hat{\beta} + \hat{e}$ for the estimation sample observations (we cannot estimate $\hat{e}$ out-of-sample in the linear regression problem). Since $y = XZp + Vw$ is our data constraint, $\hat{y}$ will equal y unless the optimal GME probabilities cannot satisfy the data constraint.

We estimate the model using three different parameter support matrices:

1) In Parameter Support 1 (PS1), we specify $\begin{bmatrix}-20 & -10 & 0 & 10 & 20\end{bmatrix}'$ as the support for the intercept, price and discount variables and $\begin{bmatrix}-10 & -5 & 0 & 5 & 10\end{bmatrix}'$ as the support for the advertising dummy variables. We expect price and discount percentage to have a larger impact on sales than advertising so we allow for wider bounds. All parameters are shrunk toward zero under this specification.

2) In Parameter Support 2 (PS2), we specify $\begin{bmatrix}-200 & -100 & 0 & 100 & 200\end{bmatrix}'$ as the support for the intercept, price and discount variables and $\begin{bmatrix}-100 & -50 & 0 & 50 & 100\end{bmatrix}'$ as the support for the advertising dummy variables. This experiment allows us to test the performance of GME under a very wide parameter support, which would be used when we have very little prior information about the expected parameter estimates. Under this

specification, the parameters are again shrunk toward zero. However, the degree of shrinkage should be less since we rely more on the data than on prior information.

3) In Parameter Support 3 (PS3), we specify the parameter support to be symmetric about the "true" parameters. Thus, the estimates will be closest to the true parameters yet compatible with the observed data. We specify the parameter support as $[\beta-20 \quad \beta-10 \quad \beta \quad \beta+10 \quad \beta+20]'$ for the intercept, price and discount variables and $[\beta-10 \quad \beta-5 \quad \beta \quad \beta+5 \quad \beta+10]'$ for the advertising dummy variables, where the $\beta's$ above are the "true" parameters used to generate the data.

Within each parameter support, we vary the sample size, the random errors, and the error support. In addition, we perform the experiments both with and without inequality restrictions placed on the parameters for parameter supports 1 and 2. We term the restricted parameter supports PSR1 and PSR2. We discuss the inequality restrictions in section 3.4.4 below.

We use estimation samples of size $N = 150$, $N = 100$, and $N = 50$. Golan, Judge, and Miller (1996) suggest using the 3σ-rule (Pukelsheim 1994) for setting the error bounds, where σ represents the standard deviation of e. In our sampling experiments, we normalize the standard deviation of the errors to be equal to 1. Following Golan, Judge, and Perloff (1997) we calculate the sample standard deviation of y assuming it follows a uniform distribution between its minimum and maximum values; the sample standard deviation of y is equal to $s_y = 1.5$ for the full sample. We construct four GME estimators using error bounds of $(-s_y \quad 0 \quad s_y)$ for GME1, $(-2s_y \quad 0 \quad 2s_y)$ for GME2, $(-3s_y \quad 0 \quad 3s_y)$ for GME3, and $(-4s_y \quad 0 \quad 4s_y)$ for GME4. The GME2 estimator coincides with the 3σ-rule based on our known standard deviation ($\sigma = 1$) while the GME3 estimator follows the 3σ-rule based on the sample standard deviation of y.

We calculate the mean squared error of the predicted values (YMSE) as

$$YMSE = \sum_{i=1}^{nsam}\left[\sum_{n=1}^{N}(\hat{y}_n - y_n)^2\right]\Big/nsam \qquad (3.24)$$

for the estimation sample observations, and

$$YMSE_0 = \sum_{i=1}^{nsam}\left[\sum_{n=1}^{N_0}(\hat{y}_{0n} - y_{0n})^2\right] \Big/ nsam \tag{3.25}$$

for the hold-out observations. In addition, we calculate the $MSE(\hat{\beta})$, which is given by

$$MSE(\hat{\beta}) = \sum_{i=1}^{nsam}\left[\sum_{k=1}^{K}(\hat{\beta}_k - \beta_k)^2\right] \Big/ nsam. \tag{3.26}$$

The root mean squared error (RMSE) is simply the square root of the MSE. The $MSE(\hat{\beta})$ is equal to the variance plus the bias squared of the parameter estimates, $\hat{\beta}$. These terms are given by

$$VAR(\hat{\beta}) = \sum_{i=1}^{nsam}\left[\sum_{k=1}^{K}(\hat{\beta}_k - \bar{\beta}_k)^2\right] \Big/ nsam, \tag{3.27}$$

and

$$Bias^2(\hat{\beta}) = \sum_{i=1}^{nsam}\left[\sum_{k=1}^{K}(\bar{\beta}_k - \beta_k)^2\right] \Big/ nsam, \tag{3.28}$$

respectively.

3.4.3 Dimensions of the Experimental Design

The dimensions of the experimental design include the following:

1) We vary the size of the estimation sample data set. In addition to the sample size of 150 observations we carry out experiments with sample sizes of 100 and 50 observations.

2) To test whether GME estimation is robust we conduct experiments with errors drawn from different distributions. Errors are drawn from a *t*-distribution with 3 degrees of freedom and from a chi-square distribution with 5 degrees of freedom, correcting the mean to zero and the variance to one.

3) We vary the parameter and error support matrices for the GME estimator to reflect different nonsample information.

4) We modify the estimators so as to constrain the estimates to take their expected signs in each sample. Imposing such inequality information is made feasible by readily available modern software. In OLS, we simply minimize the SSE subject to the inequality restrictions, which are imposed through the GAUSS Constrained Optimization module. In GME, we impose the inequality restrictions through the parameter support. For the tuna data, we constrain the coefficient for the regular price of the target brand (β_2) to be negative and the coefficients for the regular prices of all other brands to be positive. We expect that a higher price for the target brand will reduce sales of the target brand while higher prices of competitive brands will increase sales of the target brand. Similarly, we

constrain the coefficient for discount percentage of the target brand (β_3) to be positive and the coefficients for discount percentage of all other brands to be negative. Finally, we constrain both advertising coefficients for the target brand to be positive and the coefficient for display plus major ad to be greater than the coefficient for display ad only ($\beta_5 > \beta_4$). We impose all of these constraints through the parameter support matrix.

We refer to the experiments as LR11-LR13,..., LR61-63. Table 3.1 summarizes the dimensions for the linear regression experiments.

Table 3.1 Dimensions of Monte Carlo Linear Regression Experiments

Experiment	Sample Size	Restrictions	Errors
LR 11	150	Unrestricted	Normal
LR 12	150	Unrestricted	Std.-t
LR 13	150	Unrestricted	Chi-square
LR 21	100	Unrestricted	Normal
LR 22	100	Unrestricted	Std.-t
LR 23	100	Unrestricted	Chi-square
LR 31	50	Unrestricted	Normal
LR 32	50	Unrestricted	Std.-t
LR 33	50	Unrestricted	Chi-square
LR 41	150	Restricted	Normal
LR 42	150	Restricted	Std.-t
LR 43	150	Restricted	Chi-square
LR 51	100	Restricted	Normal
LR 52	100	Restricted	Std.-t
LR 53	100	Restricted	Chi-square
LR 61	50	Restricted	Normal
LR 62	50	Restricted	Std.-t
LR 63	50	Restricted	Chi-square

3.4.4 Parameter Restrictions in GME

In GME, we impose all of our parameter restrictions through the parameter support matrix, Z. We modify Z such that it is not block diagonal, which allows us to impose parameter inequality restrictions of the type $\beta_1 > \beta_2$. Using the block diagonal support matrix defined by Golan, Judge, and Miller (1996) we can easily impose sign and other inequality restrictions on individual parameters through the support matrix. For example, in the tuna data we constrain the coefficient for the regular price of the target brand, β_2, to be negative. If we specify the support for β_2 as

$[-20 \quad -15 \quad -10 \quad -5 \quad 0]'$, our GME estimate is

$$\hat{\beta}_2 = -20\hat{p}_{21} - 15\hat{p}_{22} - 10\hat{p}_{23} - 5\hat{p}_{24} - 0\hat{p}_{25} < 0, \tag{3.29}$$

since $\hat{p}_{2m} > 0$ for all M support points.

To impose cross-equation inequality restrictions using the parameter support matrix we specify a parameter support matrix that is not block diagonal. For example, in the tuna data we constrain $\beta_5 > \beta_4$ (the coefficient for display and major ad is greater than the coefficient for display only for the target brand). To do this we specify these unknown parameters as

$$\begin{bmatrix}\beta_4\\ \beta_5\end{bmatrix} = Z^{*}\begin{bmatrix}p_4\\ p_5\end{bmatrix} = \begin{bmatrix}z_4' & 0\\ z_4' & z_5'\end{bmatrix}\cdot\begin{bmatrix}p_4\\ p_5\end{bmatrix}, \tag{3.30}$$

where Z^{*} is the $2\times 2M$ sub-matrix of support points for β_4 and β_5, and p_4 and p_5 represent the unknown probabilities associated with the support points for these parameters. In our experiments with inequality restrictions, we specify both z_4' and z_5' to be $\begin{bmatrix}0 & 2.5 & 5 & 7.5 & 10\end{bmatrix}'$. In this case, the GME estimates are

$$\hat{\beta}_4 = 0\hat{p}_{41} + 2.5\hat{p}_{42} + 5\hat{p}_{43} + 7.5\hat{p}_{44} + 10\hat{p}_{45} > 0, \text{ and}$$

$$\begin{aligned}\hat{\beta}_5 &= 0(\hat{p}_{41} + \hat{p}_{51}) + 2.5(\hat{p}_{42} + \hat{p}_{52}) + 5(\hat{p}_{43} + \hat{p}_{53}) + 7.5(\hat{p}_{44} + \hat{p}_{54}) + 10(\hat{p}_{45} + \hat{p}_{55})\\ &= \hat{\beta}_4 + 0\hat{p}_{51} + 2.5\hat{p}_{52} + 5\hat{p}_{53} + 7.5\hat{p}_{54} + 10\hat{p}_{55} > \hat{\beta}_4 .\end{aligned} \tag{3.31}$$

Using this method, we may specify inequality restrictions between any set of parameters.

The solution to the GME linear regression problem when the parameter support matrix is not block diagonal is obtained in the same manner as the traditional GME solution. The Lagrangian is formed the same way and solving the KM FOC's for the unknown p's we again obtain

$$\hat{p} = \exp(Z'X'\hat{\lambda})\exp[-i_{KM} - (I_K \otimes i_M)\gamma].$$

However, since Z is not block diagonal we now obtain

$$\hat{p}_{km} = \exp\left(\sum_{r=1}^{K} z_{rkm}x_r'\lambda\right)\exp(-1-\gamma_k).$$

Since the additivity constraint requires that $\sum_{m=1}^{M} p_{km} = 1$ and $\exp(-1-\gamma_k)$ is constant for a given parameter, the optimal probabilities may be rewritten as

$$\hat{p}_{km} = \frac{\exp\left(\sum_{r=1}^{K} z_{rkm} x_r' \lambda\right)}{\sum_{m=1}^{M} \exp\left(\sum_{r=1}^{K} z_{rkm} x_r' \lambda\right)} = \frac{\exp\left(\sum_{r=1}^{K} z_{rkm} x_r' \lambda\right)}{\Omega_k(\lambda)}. \tag{3.32}$$

3.5 Monte Carlo Results

In this section we present Monte Carlo results which examine risk measures for the OLS and GME estimators in a linear regression model. Section 3.5.1 gives results for experiments using PS1 as the GME parameter support. Section 3.5.2 gives results for experiments using PSR1, which includes parameter inequality restrictions, as the GME parameter support. Section 3.5.3 gives results for PS2, section 3.5.4 gives results for PSR2, and section 3.5.5 gives results for PS3.

3.5.1 Parameter Support 1 (PS1)

In this section, we present results for Monte Carlo experiments using PS1 as the GME parameter support. Table 3.2 gives summary statistics for the independent variables in the tuna data set over all 156 observations. The sample coefficient of variation is defined as $CV_x = s(x)/\bar{x}$, where $s(x)$ is the sample standard deviation of x and $\bar{x}$ is the sample mean of x. Table 3.3 summarizes PS1 and the prior means of the GME parameter estimates, Table 3.4 gives parameter estimates using the alternative estimators on all 156 observations in the tuna data set while Tables 3.5-3.7 give risk measures of the alternative estimators in sampling experiments LR11-LR13.

We set the "true" parameter values fairly close to the OLS estimates with a few exceptions. The OLS estimates for the coefficients on the regular price of brands 3 and 4 and the discount percentage of brand 4 do not take the expected signs. Therefore, we set the "true" parameter values to 1 for the regular price of brands 3 and 4 and to -0.5 for the discount percentage of brand 4. The GME estimates for Price3 are positive while the rest of the GME estimates take the same sign as the OLS estimates. However, the GME estimates for DISMAD2 are negative for GME1 and GME2 and positive for GME3 and GME4.

Table 3.2 Summary Statistics for Tuna Fish Data (*N*=156 Observations)

Variable	Mean	Min.	Max.	Standard Deviation	Coefficient of Variation
Price1	0.89	0.66	1.12	0.13	0.15
Discount1	0.08	0	0.52	0.11	1.45
DIS1	0.33	0	1	0.47	1.42
DISMAD1	0.12	0	1	0.33	2.69
Price2	0.90	0.65	1.07	0.12	0.14
Discount2	0.08	0	0.55	0.12	1.44
DIS2	0.28	0	1	0.45	1.60
DISMAD2	0.14	0	1	0.35	2.48
Price3	0.88	0.59	1.02	0.11	0.12
Discount3	0.06	0	0.58	0.11	1.73
DIS3	0.38	0	1	0.49	1.29
DISMAD3	0.11	0	1	0.31	2.87
Price4	0.72	0.55	0.89	0.09	0.13
Discount4	0.04	0	0.25	0.07	1.80
DIS4	0.12	0	1	0.32	2.78
DISMAD4	0.05	0	1	0.22	4.32

Table 3.3 PS1 Support Points and Prior Means

Variable	z_1	z_2	z_3	z_4	z_5	Prior Mean
Intercept	-20	-10	0	10	20	0
Price1	-20	-10	0	10	20	0
Discount1	-20	-10	0	10	20	0
DIS1	-10	-5	0	5	10	0
DISMAD1	-10	-5	0	5	10	0
Price2	-20	-10	0	10	20	0
Discount2	-20	-10	0	10	20	0
DIS2	-10	-5	0	5	10	0
DISMAD2	-10	-5	0	5	10	0
Price3	-20	-10	0	10	20	0
Discount3	-20	-10	0	10	20	0
DIS3	-10	-5	0	5	10	0
DISMAD3	-10	-5	0	5	10	0
Price4	-20	-10	0	10	20	0
Discount4	-20	-10	0	10	20	0
DIS4	-10	-5	0	5	10	0
DISMAD4	-10	-5	0	5	10	0

Table 3.4 Parameter Estimates for Tuna Fish Data using Alternative Estimators

Variable	OLS	GME1	GME2	GME3	GME4	"True" MC Parameters
Intercept	4.08	4.06	4.02	3.97	3.92	4.00
	(8.02)					
Price1	-4.60	-4.75	-3.50	-2.58	-1.83	-4.50
	(-3.43)					
Discount1	3.61	3.20	3.37	3.23	3.05	3.50
	(5.56)					
DIS1	0.53	0.51	0.51	0.51	0.50	0.50
	(4.53)					
DISMAD1	1.05	1.09	1.05	1.05	1.05	1.00
	(4.68)					
Price2	10.21	8.72	7.38	5.72	4.48	9.00
	(5.07)					
Discount2	-0.39	-0.10	-0.31	-0.30	-0.28	-0.40
	(-0.50)					
DIS2	0.10	0.09	0.14	0.18	0.20	0.10
	(0.64)					
DISMAD2	-0.07	-0.11	-0.02	0.02	0.05	-0.05
	(-0.29)					
Price3	-1.31	0.54	0.13	0.59	0.86	1.00
	(-0.69)					
Discount3	-2.19	-2.88	-2.41	-2.32	-2.20	-2.00
	(-2.68)					
DIS3	0.36	0.40	0.39	0.40	0.41	0.40
	(2.95)					
DISMAD3	0.33	0.38	0.35	0.34	0.31	0.30
	(1.25)					
Price4	-1.54	-1.69	-1.04	-0.62	-0.26	1.00
	(-1.16)					
Discount4	0.14	0.65	0.42	0.45	0.42	-0.50
	(0.13)					
DIS4	0.29	0.26	0.28	0.29	0.29	0.30
	(1.57)					
DISMAD4	0.19	0.14	0.15	0.14	0.13	0.20
	(0.65)					
R^2	0.72					
$\hat{\sigma}^2$	0.39					

* N =156; Parameter support = PS1.

Table 3.5 Monte Carlo Linear Regression 11 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	49.49	247.17	38.25	40.15	48.10
RMSE ($\hat{\beta}$)	7.04	15.72	6.18	6.34	6.94
Var ($\hat{\beta}$)	49.44	244.38	30.19	17.92	11.84
$Bias^2(\hat{\beta})$	0.05	2.79	8.07	22.23	36.26
PRMSE	11.58	44.90	11.65	11.68	11.75
PRMSE ($\hat{p}+\hat{e}$)		42.92	0.00	0.00	0.00
$PRMSE_0$	7.54	26.45	7.52	7.49	7.50
R^2	0.77				

* N=150; unrestricted; normal errors.

Table 3.6 Monte Carlo Linear Regression 12 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	46.42	507.60	285.70	133.66	91.45
RMSE ($\hat{\beta}$)	6.81	22.53	16.90	11.56	9.56
Var ($\hat{\beta}$)	46.35	500.20	277.14	113.30	57.18
$Bias^2(\hat{\beta})$	0.07	7.40	8.57	20.36	34.26
PRMSE	11.10	86.69	53.01	30.34	16.64
PRMSE ($\hat{p}+\hat{e}$)		85.46	50.77	27.46	11.29
$PRMSE_0$	7.42	50.74	31.36	18.45	10.51
R^2	0.77				

* N=150; unrestricted; standardized *t*- errors.

Table 3.7 Monte Carlo Linear Regression 13 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	48.82	273.75	90.61	44.77	48.99
RMSE ($\hat{\beta}$)	6.99	16.55	9.52	6.69	7.00
Var ($\hat{\beta}$)	48.70	263.34	79.36	22.24	12.48
Bias^2($\hat{\beta}$)	0.11	10.41	11.25	22.53	36.51
PRMSE	11.51	45.64	14.14	11.65	11.68
PRMSE ($\hat{p}+\hat{e}$)		44.85	7.04	0.01	0.00
PRMSE$_0$	7.40	26.91	8.89	7.44	7.39
R^2	0.77				

* N=150; unrestricted; standardized chi-square errors.

We find that the shrinkage in GME is greater as we widen the error bounds, thus allowing more weight to be placed on the estimated error term. In addition, the GME estimates are shrunk most for the components with the lowest coefficients of variation. These are also the components with the smallest variances and the largest characteristic roots in the Belsley et al. (1980) singular value decomposition. Thus, these are the components responsible for the collinearity. This type of shrinkage is consistent with Stein-like estimators, which also shrink components with small variances more than components with large variances (Judge et al. 1985, p. 923). In contrast, certain ridge estimators shrink components with large variances more than components with small variance (Judge et al. 1985, p. 921).

Tables 3.5 and 3.7 show that the GME estimator does have a lower MSE than OLS under certain sets of error bounds when the errors are drawn from a standard normal distribution or a standardized chi-square distribution with 5 degrees of freedom. However, OLS always outperforms GME when the errors are drawn from a standardized *t*- distribution with 3 degrees of freedom, which has thicker tails than the other error distributions we consider. Golan, Judge, and Miller (1996, p. 142) note that the standardized *t*- distribution has thicker tails than the normal or standardized chi-square distributions and they specify a wider error support for the *t*- distribution in their experiments. They find that GME has a lower mean square error loss (MSEL) than OLS even when the errors are drawn from a standardized *t*- distribution. However, in our Monte Carlo experiments with errors drawn from a

standardized t- distribution OLS has a lower MSE than GME even when we specify very wide error bounds, as in the GME4 estimator.

The GME1 estimator, which has the narrowest error bounds, does not perform well under any of the error distributions. We find that the variance of the GME estimator decreases as we widen the error bounds; as we place more weight on the errors the parameter estimates do not vary as much in each Monte Carlo iteration. However, the bias increases as we widen the error bounds. Thus, it appears there is a unique error support that minimizes the MSE $(\hat{\beta})$ for a given GME parameter support.

We find that the estimation sample prediction MSE is lower for OLS than for GME in every case. However, the out-of-sample prediction MSE is lower for GME than for OLS for both the normal and chi-square experiments when we specify wider error bounds. We find that our estimate $\hat{y} = \hat{p} + \hat{e}$ is equal to the true value of y when we allow wide enough error bounds, which means that the data constraint is satisfied by the optimal probabilities. Recall that $y = p + e$ is a constraint in the GME model. Therefore, GME is unable to find a solution that satisfies both the data and additivity constraints when the PRMSE using $\hat{p} + \hat{e}$ is not equal to zero.

Tables 3.8-3.10 give risk measures of the alternative estimators in sampling experiments LR21-LR23, which have estimation sample size $N = 100$. The bias for each estimator, except GME1, increases as we decrease the sample size. The variance also increases as the sample size decreases except for experiment LR22, which has errors drawn from a standardized t- distribution, where the variance of the GME estimators decreases as the sample size decreases. We again find that GME outperforms OLS in terms of MSE $(\hat{\beta})$ and out-of-sample prediction when the errors are drawn from a standard normal or standardized chi-square distribution. The amount of risk gain by GME compared to OLS is greater the smaller the sample size. For the GME estimator the out-of-sample prediction MSE declines as we widen the error bounds even in cases where the in-sample prediction MSE begins to increase.

Tables 3.11-3.13 give risk measures of the alternative estimators in sampling experiments LR31-LR33, which have estimation sample size $N = 50$. We again observe that the risk measures increase at a much greater rate for the OLS estimator than for the GME estimator as we decrease the

Table 3.8 Monte Carlo Linear Regression 21 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	69.38	267.80	47.51	50.10	58.21
RMSE ($\hat{\beta}$)	8.33	16.36	6.89	7.08	7.63
Var ($\hat{\beta}$)	69.29	265.99	35.57	20.46	12.82
Bias^2($\hat{\beta}$)	0.08	1.82	11.94	29.64	45.39
PRMSE	9.12	33.17	9.20	9.26	9.35
PRMSE ($\hat{p}+\hat{e}$)		31.24	0.00	0.00	0.00
PRMSE$_0$	7.81	24.25	7.73	7.68	7.65
R^2	0.70				

* N=100; unrestricted; normal errors.

Table 3.9 Monte Carlo Linear Regression 22 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	60.96	532.49	215.62	103.09	85.30
RMSE ($\hat{\beta}$)	7.81	23.08	14.68	10.15	9.24
Var ($\hat{\beta}$)	60.84	528.64	196.54	70.54	38.99
Bias^2($\hat{\beta}$)	0.12	3.85	19.08	32.55	46.31
PRMSE	8.75	69.45	34.04	16.68	15.37
PRMSE ($\hat{p}+\hat{e}$)		68.45	32.39	13.75	12.64
PRMSE$_0$	7.65	52.28	24.96	12.91	11.84
R^2	0.70				

* N=100; unrestricted; standardized *t*- errors.

Table 3.10 Monte Carlo Linear Regression 23 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	67.69	306.45	82.25	53.99	58.61
RMSE ($\hat{\beta}$)	8.23	17.51	9.07	7.35	7.66
Var ($\hat{\beta}$)	67.59	298.22	69.48	24.56	13.63
Bias^2 ($\hat{\beta}$)	0.10	8.23	12.77	29.43	44.98
PRMSE	9.13	50.74	10.82	9.33	9.37
PRMSE ($\hat{p}+\hat{e}$)		50.18	5.45	0.00	0.00
PRMSE$_0$	7.81	36.86	8.91	7.77	7.67
R^2	0.70				

* N=100; unrestricted; standardized chi-square errors.

Table 3.11 Monte Carlo Linear Regression 31 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	211.38	284.54	79.54	76.97	81.41
RMSE ($\hat{\beta}$)	14.54	16.87	8.92	8.77	9.02
Var ($\hat{\beta}$)	211.31	279.43	47.32	22.37	12.56
Bias^2 ($\hat{\beta}$)	0.07	5.11	32.21	54.60	68.85
PRMSE	5.73	18.33	5.91	6.06	6.19
PRMSE ($\hat{p}+\hat{e}$)		17.17	0.00	0.00	0.00
PRMSE$_0$	9.10	19.33	8.45	8.25	8.15
R^2	0.84				

* N=50; unrestricted; normal errors.

Table 3.12 Monte Carlo Linear Regression 32 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	204.29	433.90	178.84	113.22	99.39
RMSE ($\hat{\beta}$)	14.29	20.83	13.37	10.64	9.97
Var ($\hat{\beta}$)	204.06	421.99	143.43	55.38	29.87
Bias^2($\hat{\beta}$)	0.23	11.91	35.41	57.84	69.52
PRMSE	5.69	37.90	19.21	8.58	8.13
PRMSE ($\hat{p}+\hat{e}$)		37.33	18.26	5.39	4.61
PRMSE$_0$	9.01	38.50	20.32	10.88	10.37
R^2	0.84				

* N=50; unrestricted; standardized *t*- errors.

Table 3.13 Monte Carlo Linear Regression 33 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	202.24	358.57	99.90	80.70	83.08
RMSE ($\hat{\beta}$)	14.22	18.94	10.00	8.98	9.11
Var ($\hat{\beta}$)	202.05	348.99	64.97	24.91	13.46
Bias^2($\hat{\beta}$)	0.18	9.58	34.93	55.79	69.62
PRMSE	5.74	32.80	6.22	6.07	6.18
PRMSE ($\hat{p}+\hat{e}$)		32.22	1.86	0.00	0.00
PRMSE$_0$	9.08	32.91	8.83	8.31	8.16
R^2	0.84				

* N=50; unrestricted; standardized chi-square errors.

sample size. The variance for the GME estimator decreases as the sample size decreases when the errors are drawn from a standardized *t*- distribution. In fact, the $MSE(\hat{\beta})$ is lower for GME2, GME3, and GME4 than for OLS in this experiment.

3.5.2 Restricted Parameter Support 1 (PSR1)

In this section, we present Monte Carlo results using PSR1, which places inequality restrictions on the parameters, as the GME parameter support. We examine risk measures for inequality restricted least squares (IRLS) and GME estimators with inequality restrictions placed on the parameters. Table

3.14 summarizes PSR1 and the prior means of the GME parameter estimates. Each parameter is shrunk toward its prior mean, which no longer equals zero for all parameters in the restricted model. Table 3.15 gives parameter estimates using the alternative estimators on all 156 observations in the tuna data set while Tables 3.16-3.18 give risk measures of the alternative estimators in sampling experiments LR41-LR43.

We observe that the IRLS estimator has lower risk than the OLS estimator from experiment LR11 does, although the IRLS estimator is biased. The risk measures increase for our GME estimators under this set of restrictions. However, this is largely because the prior means change when we impose the inequality restrictions since each estimator is shrunk toward its prior mean. In a sense, every GME estimator has inequality restrictions since we must specify a parameter support that bounds each parameter. In Chapter 4, we further discuss the impact of the prior mean when imposing inequality restrictions through the parameter support matrix. We obtain similar results for sampling experiments LR51-LR63, whose results are given in Appendix D, Tables D.13-D.18.

3.5.3 Parameter Support 2 (PS2)

In this section, we present results for Monte Carlo experiments using PS2, which places very wide bounds on the unknown parameters, as the GME parameter support. This type of parameter support would be used if we do not have good nonsample information with which to specify the parameter support. This is analogous to the noninformative prior in Bayesian inference. Table 3.19 summarizes PS2 and the prior means of the GME parameter estimates, Table 3.20 gives parameter estimates using the alternative estimators on all 156 observations in the tuna data set while Tables 3.21-3.23 give risk measures of the alternative estimators in sampling experiments LR11-LR13.

The results for PS2 show that the GME estimates are very close to the OLS estimates when we specify wide support vectors for each parameter. In addition, OLS has lower risk than GME for most of the experiments. GME3 and GME4 have lower risk than OLS when the errors are standard normal, but the difference is very small. These results indicate that there is not much reason to use GME unless we have nonsample information that we can use to restrict the parameter space. The results for experiments LR21-LR33 are given in Appendix D, Tables D.21-D.27.

Table 3.14 PSR1 Support Points and Prior Means

Variable	z_1	z_2	z_3	z_4	z_5	Prior Mean
Intercept	-20	-10	0	10	20	0
Price1	-20	-15	-10	-5	0	-10
Discount1	0	5	10	15	20	10
DIS1	0	2.5	5	7.5	10	5
DISMAD1	0	2.5	5	7.5	10	5
Price2	0	5	10	15	20	10
Discount2	-20	-15	-10	-5	0	-10
DIS2	-10	-5	0	5	10	0
DISMAD2	-10	-5	0	5	10	0
Price3	0	5	10	15	20	10
Discount3	-20	-15	-10	-5	0	-10
DIS3	-10	-5	0	5	10	0
DISMAD3	-10	-5	0	5	10	0
Price4	0	5	10	15	20	10
Discount4	-20	-15	-10	-5	0	-10
DIS4	-10	-5	0	5	10	0
DISMAD4	-10	-5	0	5	10	0

Table 3.15 Parameter Estimates for Tuna Fish Data using Alternative Estimators

Variable	OLS	GME1	GME2	GME3	GME4	"True" MC Parameters
Intercept	3.66	3.23	1.99	1.04	0.29	4.00
Price1	-5.05	-6.95	-9.19	-10.88	-11.76	-4.50
Discount1	3.50	3.16	3.79	4.17	4.58	3.50
DIS1	0.52	0.54	0.56	0.63	0.73	0.50
DISMAD1	1.08	1.17	1.33	1.57	1.84	1.00
Price2	8.65	8.33	7.87	7.77	7.78	9.00
Discount2	-0.56	-0.89	-2.38	-3.56	-4.57	-0.40
DIS2	0.14	0.19	0.35	0.41	0.43	0.10
DISMAD2	0.00	0.08	0.47	0.75	0.98	-0.05
Price3	0.00	2.28	4.26	5.68	6.49	1.00
Discount3	-2.50	-3.32	-3.87	-4.82	-5.68	-2.00
DIS3	0.33	0.30	0.18	0.10	0.05	0.40
DISMAD3	0.32	0.33	0.27	0.32	0.39	0.30
Price4	0.00	0.68	3.46	5.36	6.56	1.00
Discount4	0.00	-0.61	-2.42	-3.78	-4.87	-0.50
DIS4	0.30	0.33	0.49	0.58	0.64	0.30
DISMAD4	0.18	0.31	0.48	0.63	0.73	0.20

* N =156; Parameter support = PSR1.

Table 3.16 Monte Carlo Linear Regression 41 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE $(\hat{\beta})$	32.29	183.23	84.49	150.89	205.01
RMSE $(\hat{\beta})$	5.68	13.54	9.19	12.28	14.32
Var $(\hat{\beta})$	30.22	157.61	11.37	5.12	3.14
Bias$^2(\hat{\beta})$	2.07	25.62	73.12	145.77	201.87
PRMSE	11.64	72.33	12.20	12.94	13.85
PRMSE $(\hat{p}+\hat{e})$		71.28	0.00	0.00	0.00
PRMSE$_0$	7.47	42.40	7.72	8.06	8.48

* N=150; restricted; normal errors.

Table 3.17 Monte Carlo Linear Regression 42 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE $(\hat{\beta})$	31.30	314.07	200.90	194.73	222.66
RMSE $(\hat{\beta})$	5.59	17.72	14.17	13.95	14.92
Var $(\hat{\beta})$	29.12	257.62	119.53	50.37	24.38
$\text{Bias}^2(\hat{\beta})$	2.18	56.45	81.37	144.36	198.27
PRMSE	11.16	82.45	36.56	23.86	16.71
PRMSE $(\hat{p}+\hat{e})$		81.22	33.38	19.44	8.18
PRMSE_0	7.37	48.51	21.81	14.56	10.37

* N=150; restricted; standardized *t*- errors.

Table 3.18 Monte Carlo Linear Regression 43 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE $(\hat{\beta})$	32.54	189.50	113.85	153.58	206.14
RMSE $(\hat{\beta})$	5.70	13.77	10.67	12.39	14.36
Var $(\hat{\beta})$	29.64	153.92	36.20	7.33	3.44
$\text{Bias}^2(\hat{\beta})$	2.90	35.57	77.65	146.25	202.69
PRMSE	11.56	63.21	16.43	12.91	13.77
PRMSE $(\hat{p}+\hat{e})$		59.00	10.25	0.02	0.00
PRMSE_0	7.34	37.15	10.11	8.03	8.38

* N=150; restricted; standardized chi-square errors.

Table 3.19 PS2 Support Points and Prior Means

Variable	z1	z2	z3	z4	z5	Prior Mean
Intercept	-200	-100	0	100	200	0
Price1	-200	-100	0	100	200	0
Discount1	-200	-100	0	100	200	0
DIS1	-100	-50	0	50	100	0
DISMAD1	-100	-50	0	50	100	0
Price2	-200	-100	0	100	200	0
Discount2	-200	-100	0	100	200	0
DIS2	-100	-50	0	50	100	0
DISMAD2	-100	-50	0	50	100	0
Price3	-200	-100	0	100	200	0
Discount3	-200	-100	0	100	200	0
DIS3	-100	-50	0	50	100	0
DISMAD3	-100	-50	0	50	100	0
Price4	-200	-100	0	100	200	0
Discount4	-200	-100	0	100	200	0
DIS4	-100	-50	0	50	100	0
DISMAD4	-100	-50	0	50	100	0

Table 3.20 Parameter Estimates for Tuna Fish Data using Alternative Estimators

Variable	OLS	GME1	GME2	GME3	GME4	"True" MC Parameters
Intercept	4.08 (8.02)	4.06	4.07	4.08	4.08	4.00
Price1	-4.60 (-3.43)	-4.91	-4.67	-4.60	-4.56	-4.50
Discount1	3.61 (5.56)	3.38	3.56	3.58	3.59	3.50
DIS1	0.53 (4.53)	0.52	0.53	0.53	0.53	0.50
DISMAD1	1.05 (4.68)	1.07	1.05	1.05	1.05	1.00
Price2	10.21 (5.07)	9.57	10.05	10.07	10.03	9.00
Discount2	-0.39 (-0.50)	-0.17	-0.34	-0.37	-0.38	-0.40
DIS2	0.10 (0.64)	0.08	0.09	0.10	0.10	0.10
DISMAD2	-0.07 (-0.29)	-0.11	-0.08	-0.07	-0.07	-0.05
Price3	-1.31 (-0.69)	-0.22	-1.05	-1.17	-1.18	1.00
Discount3	-2.19 (-2.68)	-2.68	-2.32	-2.25	-2.23	-2.00
DIS3	0.36 (2.95)	0.38	0.37	0.37	0.36	0.40
DISMAD3	0.33 (1.25)	0.36	0.34	0.33	0.33	0.30
Price4	-1.54 (-1.16)	-1.62	-1.56	-1.53	-1.52	1.00
Discount4	0.14 (0.13)	0.42	0.21	0.18	0.17	-0.50
DIS4	0.29 (1.57)	0.27	0.29	0.29	0.29	0.30
DISMAD4	0.19 (0.65)	0.17	0.18	0.18	0.18	0.20
R^2	0.72					
$\hat{\sigma}^2$	0.39					

* N =156; Parameter support = PS2.

Table 3.21 Monte Carlo Linear Regression 11 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	49.49	64.44	51.47	48.76	48.05
RMSE ($\hat{\beta}$)	7.04	8.03	7.17	6.98	6.93
Var ($\hat{\beta}$)	49.44	64.32	51.41	48.68	47.93
$Bias^2$ ($\hat{\beta}$)	0.05	0.12	0.06	0.08	0.11
PRMSE	11.58	12.07	11.61	11.58	11.58
PRMSE ($\hat{p}+\hat{e}$)		3.63	0.10	0.00	0.00
$PRMSE_0$	7.54	7.80	7.55	7.54	7.53
R^2	0.77				

* N=150; unrestricted; normal errors.

Table 3.22 Monte Carlo Linear Regression 12 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	46.42	75.99	352.57	259.73	246.67
RMSE ($\hat{\beta}$)	6.81	8.72	18.78	16.12	15.71
Var ($\hat{\beta}$)	46.35	75.92	351.74	259.44	246.44
$Bias^2$ ($\hat{\beta}$)	0.07	0.08	0.82	0.29	0.23
PRMSE	11.10	12.11	17.15	13.65	13.66
PRMSE ($\hat{p}+\hat{e}$)		6.71	10.92	5.05	4.80
$PRMSE_0$	7.42	8.02	10.86	9.02	8.86
R^2	0.77				

* N=150; unrestricted; standardized *t*- errors.

Table 3.23 Monte Carlo Linear Regression 13 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	48.82	56.90	106.61	57.61	50.01
RMSE ($\hat{\beta}$)	6.99	7.54	10.33	7.59	7.07
Var ($\hat{\beta}$)	48.70	56.71	105.66	57.36	49.75
Bias^2($\hat{\beta}$)	0.11	0.20	0.95	0.26	0.26
PRMSE	11.51	12.34	11.99	11.56	11.51
PRMSE ($\hat{p}+\hat{e}$)		5.97	0.67	0.04	0.00
PRMSE$_0$	7.40	7.86	7.80	7.48	7.42
R^2	0.77				

* N=150; unrestricted; standardized chi-square errors.

3.5.4 Restricted Parameter Support 2 (PSR2)

In this section, we present results for Monte Carlo experiments using PSR2 as the GME parameter support. Table 3.24 summarizes PSR2 and the prior means of the GME parameter estimates. Each parameter is shrunk toward its prior mean, which in the restricted model does not equal zero for all parameters. Table 3.25 gives parameter estimates using the alternative estimators on all 156 observations in the tuna data set while Tables 3.26-3.28 give risk measures of the alternative estimators in sampling experiments LR41-LR43. The results show that GME performs very poorly when we specify a wide parameter support and also place inequality restrictions on the parameters. This is because the parameters are shrunk towards their prior means which do not equal the true parameters. We discuss GME estimation with inequality restrictions in Chapter 4. The results for experiments LR51-LR63 are given in Appendix D, Tables D.31-D.36.

3.5.5 Parameter Support 3 (PS3)

In this section, we present results for Monte Carlo experiments using PS3, which includes relatively narrow bounds that are symmetric about the true parameters, as the GME parameter support. In contrast to PS2, which represents a situation where we have very little nonsample information, PS3 represents a situation where we have very good nonsample information about the true parameters. We

Table 3.24 PSR2 Support Points and Prior Means

Variable	z1	z2	z3	z4	z5	Prior Mean
Intercept	-200	-100	0	100	200	0
Price1	-200	-150	-100	-50	0	-100
Discount1	0	50	100	150	200	100
DIS1	0	25	50	75	100	50
DISMAD1	0	25	50	75	100	50
Price2	0	50	100	150	200	100
Discount2	-200	-150	-100	-50	0	-100
DIS2	-100	-50	0	50	100	0
DISMAD2	-100	-50	0	50	100	0
Price3	0	50	100	150	200	100
Discount3	-200	-150	-100	-50	0	-100
DIS3	-100	-50	0	50	100	0
DISMAD3	-100	-50	0	50	100	0
Price4	0	50	100	150	200	100
Discount4	-200	-150	-100	-50	0	-100
DIS4	-100	-50	0	50	100	0
DISMAD4	-100	-50	0	50	100	0

Table 3.25 Parameter Estimates for Tuna Fish Data using Alternative Estimators

Variable	OLS	GME1	GME2	GME3	GME4	"True" MC Parameters
Intercept	3.66	3.61	3.03	2.17	1.17	4.00
Price1	-5.05	-6.12	-7.83	-10.84	-14.42	-4.50
Discount1	3.50	3.15	3.72	4.01	4.36	3.50
DIS1	0.52	0.53	0.53	0.54	0.55	0.50
DISMAD1	1.08	1.13	1.15	1.24	1.35	1.00
Price2	8.65	9.05	10.13	11.22	12.66	9.00
Discount2	-0.56	-0.42	-1.19	-2.01	-2.94	-0.40
DIS2	0.14	0.12	0.20	0.28	0.36	0.10
DISMAD2	0.00	-0.05	0.12	0.29	0.48	-0.05
Price3	0.00	0.78	1.52	3.25	5.17	1.00
Discount3	-2.50	-3.09	-2.98	-3.55	-4.30	-2.00
DIS3	0.33	0.34	0.26	0.16	0.04	0.40
DISMAD3	0.32	0.35	0.27	0.23	0.19	0.30
Price4	0.00	0.00	0.72	2.30	4.10	1.00
Discount4	0.00	-0.02	-1.00	-2.07	-3.28	-0.50
DIS4	0.30	0.28	0.38	0.46	0.54	0.30
DISMAD4	0.18	0.22	0.32	0.44	0.58	0.20

* N =156; Parameter support = PSR2.

Table 3.26 Monte Carlo Linear Regression 41 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	32.29	520.38	55.13	124.30	268.01
RMSE ($\hat{\beta}$)	5.68	22.81	7.42	11.15	16.37
Var ($\hat{\beta}$)	30.22	153.00	29.81	24.47	21.47
$\text{Bias}^2(\hat{\beta})$	2.07	367.39	25.32	99.83	246.55
PRMSE	11.64	157.57	11.81	12.19	13.01
PRMSE ($\hat{p}+\hat{e}$)		157.18	0.74	0.00	0.01
PRMSE_0	7.47	95.99	7.59	7.82	8.32

* N=150; restricted; normal errors.

Table 3.27 Monte Carlo Linear Regression 42 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	31.30	1697.01	357.11	307.48	378.60
RMSE ($\hat{\beta}$)	5.59	41.19	18.90	17.54	19.46
Var ($\hat{\beta}$)	29.12	686.75	274.69	186.80	120.83
$\text{Bias}^2(\hat{\beta})$	2.18	1010.26	82.42	120.68	257.77
PRMSE	11.16	293.28	103.15	75.83	53.30
PRMSE ($\hat{p}+\hat{e}$)		293.12	102.65	75.27	51.70
PRMSE_0	7.37	178.30	61.68	44.96	31.63

* N=150; restricted; standardized *t*- errors.

Table 3.28 Monte Carlo Linear Regression 43 – PSR2 (See Table 3.1 for design)

	OLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	32.54	901.76	103.96	132.67	272.72
RMSE ($\hat{\beta}$)	5.70	30.03	10.20	11.52	16.51
Var ($\hat{\beta}$)	29.64	262.08	67.25	30.03	22.44
Bias^2($\hat{\beta}$)	2.90	639.68	36.72	102.64	250.28
PRMSE	11.56	210.04	30.03	12.17	12.94
PRMSE ($\hat{p}+\hat{e}$)		208.10	26.60	0.14	0.01
PRMSE$_0$	7.34	127.97	18.24	7.78	8.23

* N=150; restricted; standardized chi-square errors.

do not specify a restricted version of PS3. Table 3.29 summarizes PS3 and the prior means of the GME parameter estimates, Table 3.30 gives parameter estimates using the alternative estimators on all 156 observations in the tuna data set while Tables 3.31-3.33 give risk measures of the alternative estimators in sampling experiments LR11-LR13.

When we specify the parameter bounds as symmetric about the true parameters, GME has considerably lower risk than OLS. In fact, the GME estimator has both a lower variance and bias than OLS in our sampling experiments with normal and chi-square errors. GME performs especially well when we specify wide error bounds since this leads to greater shrinkage toward the prior means, which we specify to be the true parameter values. Thus, if we have good nonsample information we should specify wide error bounds to shrink to parameter estimates toward the prior means. The MSE of the GME4 estimator is more than four times lower than the MSE of the OLS estimator when we have normal errors and a sample size of $N=150$. With a sample size of $N=50$, the MSE of the GME4 estimator is nearly twenty times lower than the MSE of the OLS estimator. The results for experiments LR21-LR33 are given in Appendix D, Tables D.39-D.45.

3.5.6 Response Surfaces for Monte Carlo Experiments

In this section, we estimate response surfaces for our Monte Carlo experiments. Hendry (1984) and Davidson and MacKinnon (1993) discuss response surfaces as a means of summarizing the results from a set of Monte Carlo experiments. A response surface is a regression model with a measure of the

Table 3.29 PS3 Support Points and Prior Means

Variable	$z1$	$z2$	$z3$	$z4$	$z5$	Prior Mean
Intercept	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
Price1	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
Discount1	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
DIS1	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β
DISMAD1	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β
Price2	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
Discount2	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
DIS2	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β
DISMAD2	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β
Price3	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
Discount3	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
DIS3	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β
DISMAD3	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β
Price4	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
Discount4	$\beta-20$	$\beta-10$	β	$\beta+10$	$\beta+20$	β
DIS4	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β
DISMAD4	$\beta-10$	$\beta-5$	β	$\beta+5$	$\beta+10$	β

Table 3.30 Parameter Estimates for Tuna Fish Data using Alternative Estimators

Variable	OLS	GME1	GME2	GME3	GME4	"True" MC Parameters
Intercept	4.08 (8.02)	4.04	3.98	3.91	3.84	4.00
Price1	-4.60 (-3.43)	-5.07	-4.71	-4.74	-4.80	-4.50
Discount1	3.61 (5.56)	3.23	3.53	3.54	3.53	3.50
DIS1	0.53 (4.53)	0.52	0.52	0.52	0.52	0.50
DISMAD1	1.05 (4.68)	1.09	1.06	1.06	1.06	1.00
Price2	10.21 (5.07)	9.32	9.54	9.23	8.98	9.00
Discount2	-0.39 (-0.50)	-0.11	-0.36	-0.41	-0.44	-0.40
DIS2	0.10 (0.64)	0.08	0.11	0.12	0.13	0.10
DISMAD2	-0.07 (-0.29)	-0.12	-0.06	-0.04	-0.03	-0.05
Price3	-1.31 (-0.69)	0.31	-0.61	-0.41	-0.22	1.00
Discount3	-2.19 (-2.68)	-2.87	-2.37	-2.34	-2.33	-2.00
DIS3	0.36 (2.95)	0.40	0.36	0.36	0.35	0.40
DISMAD3	0.33 (1.25)	0.38	0.33	0.32	0.31	0.30
Price4	-1.54 (-1.16)	-1.74	-1.27	-0.99	-0.73	1.00
Discount4	0.14 (0.13)	0.56	0.17	0.07	-0.01	-0.50
DIS4	0.29 (1.57)	0.26	0.29	0.30	0.31	0.30
DISMAD4	0.19 (0.65)	0.16	0.18	0.19	0.20	0.20
R^2	0.72					
$\hat{\sigma}^2$	0.39					

* N =156; Parameter support = PS1.

Table 3.31 Monte Carlo Linear Regression 11 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	49.49	241.13	30.99	18.30	12.03
RMSE ($\hat{\beta}$)	7.04	15.53	5.57	4.28	3.47
Var ($\hat{\beta}$)	49.44	240.32	30.95	18.27	12.01
Bias^2 ($\hat{\beta}$)	0.05	0.81	0.04	0.03	0.02
PRMSE	11.58	42.96	11.63	11.61	11.64
PRMSE ($\hat{p}+\hat{e}$)		41.05	0.00	0.00	0.00
PRMSE$_0$	7.54	25.28	7.51	7.45	7.42
R^2	0.77				

* N=150; unrestricted; normal errors.

Table 3.32 Monte Carlo Linear Regression 12 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	46.42	510.84	282.68	117.81	57.89
RMSE ($\hat{\beta}$)	6.81	22.60	16.81	10.85	7.61
Var ($\hat{\beta}$)	46.35	509.93	281.92	117.53	57.73
Bias^2 ($\hat{\beta}$)	0.07	0.90	0.76	0.28	0.16
PRMSE	11.10	64.76	51.54	24.48	15.41
PRMSE ($\hat{p}+\hat{e}$)		62.88	49.19	21.05	9.39
PRMSE$_0$	7.42	37.93	30.44	15.14	9.82
R^2	0.77				

* N=150; unrestricted; standardized *t*- errors.

Table 3.33 Monte Carlo Linear Regression 13 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE $(\hat{\beta})$	48.82	270.24	82.95	22.66	12.63
RMSE $(\hat{\beta})$	6.99	16.44	9.11	4.76	3.55
Var $(\hat{\beta})$	48.70	265.62	82.19	22.62	12.61
$\text{Bias}^2(\hat{\beta})$	0.11	4.61	0.76	0.04	0.01
PRMSE	11.51	33.86	16.53	11.59	11.57
PRMSE $(\hat{p}+\hat{e})$		30.43	11.05	0.01	0.00
PRMSE_0	7.40	20.01	10.18	7.40	7.31
R^2	0.77				

* N=150; unrestricted; standardized chi-square errors.

outcome from the experiment as the dependent variable and dimensions of the experiment as the independent variables. We estimate the following response surface regression

$$MSE(\hat{\beta}) = \alpha_1 + \alpha_2 N + \alpha_3 \text{stdt} + \alpha_4 \text{stdchi} + \alpha_5 \text{restrict} + \mu_i, \quad i = 1, \ldots, M \qquad (3.33)$$

where N is the sample size for the estimation sample, stdt is a dummy variable that equals 1 when the errors are drawn from a standardized t- distribution, stdchi is a dummy variable that equals 1 when the errors are drawn from a standardized chi-square distribution, restrict is a dummy variable that equals 1 when we place inequality restrictions on the parameter estimates, M is the number of Monte Carlo experiments, and $MSE(\hat{\beta})$ is the MSE of the estimator. We also estimate a response surface with $MSE(\hat{\beta})\,/\,MSE(OLS)$ as the dependent variable.

We estimate response surfaces for OLS, GME1, GME2, GME3, and GME4 where either PS1 or PSR1 is the GME parameter support matrix. For each of the response surface regressions we have $M = 18$ observations corresponding to the dimensions of our Monte Carlo experiments, which are summarized in Table 3.1, and we estimate the response surface regressions using OLS. Table 3.34 summarizes the response surface estimates with $MSE(\hat{\beta})$ as the dependent variable and Table 3.35 summarizes the response surface estimates with $MSE(\hat{\beta})\,/\,MSE(OLS)$ as the dependent variable.

Table 3.34 Response Surfaces for MSE ($\hat{\beta}$)

Variable	OLS	GME1	GME2	GME3	GME4
Intercept	226.22**	309.45**	139.16**	163.70**	173.14**
	(24.41)	(31.50)	(43.78)	(37.65)	(35.21)
Sample Size	-1.17**	-0.17	-0.66	-1.02**	-1.08**
	(0.19)	(0.24)	(0.34)	(0.29)	(0.27)
Stdt	-5.15	164.23**	125.41**	43.82	20.66
	(18.91)	(24.40)	(33.92)	(29.16)	(27.27)
Stdchi	-3.17	29.42	26.87	2.98	0.85
	(18.91)	(24.40)	(33.92)	(29.16)	(27.27)
Restrict	-45.62**	-126.23**	62.41**	159.93**	212.62**
	(15.44)	(19.92)	(27.69)	(23.81)	(22.27)
R^2	0.78	0.88	0.65	0.82	0.89
$\hat{\sigma}^2$	32.76	42.26	58.74	50.51	47.23

* M =18; Parameter support = PS1, PSR1; standard errors in parentheses.

Table 3.35 Response Surfaces for MSE ($\hat{\beta}$) / MSE(OLS)

Variable	IRLS	GME1	GME2	GME3	GME4
Intercept	.4712**	-16.78**	-0.55	3.28	7.19**
	(.0442)	(7.16)	(1.16)	(1.81)	(3.24)
Sample Size	.0014**	0.14**	0.01	-0.026	-0.065**
	(.0004)	(0.06)	(0.01)	(0.014)	(0.025)
Stdt	.0096	8.99	4.88**	3.36**	2.75
	(.0361)	(5.55)	(0.90)	(1.40)	(2.51)
Stdchi	.0007	2.86	0.75	0.26	0.32
	(.0361)	(5.55)	(0.90)	(1.40)	(2.51)
Restrict	——	17.04**	2.39**	6.09**	13.48**
		(4.53)	(0.74)	(1.14)	(2.05)
R^2	0.76	0.64	0.78	0.75	0.80
$\hat{\sigma}^2$	0.04	9.61	1.56	2.43	4.35

* M =9 for IRLS and 18 for GME; Parameter support = PS1, PSR1; standard errors in parentheses.

The response surface regressions on $MSE(\hat{\beta})$ show that increasing the sample size reduces the MSE for all of the alternative estimators, although the sample size is not significant in the response surface regressions for GME1 and GME2. The type of errors drawn has a much greater effect on the GME estimator than on the OLS estimator. Drawing errors drawn from the standardized *t*- distribution

(rather than the standard normal distribution) increases the MSE for the GME estimator, particularly for GME1 and GME2. Drawing errors from a standardized chi-square distribution also increases the MSE for the GME estimator although the effect is not significant for any of the GME estimators. As expected, imposing inequality restrictions reduces the MSE of the OLS estimator. However, the effects of restrictions are very different for the GME estimators. The observed effect of imposing inequality restrictions in GME is largely dependent on the prior mean rather than the restrictions themselves. As we widen the error bounds, which increases the degree of shrinkage towards the prior mean, the restrictions increase the MSE of the GME estimator. We further examine the impact of inequality restrictions and prior means on the GME parameter estimates in Chapter 4.

3.6 Conclusions

We examine GME estimation in linear regression models using actual data as a basis for Monte Carlo sampling experiments both with and without inequality restrictions placed on the parameters. We find that GME estimation has lower risk than OLS in terms of $MSE(\hat{\beta})$ and out-of-sample prediction error when the underlying errors have a standard normal or standardized chi-square distribution. However, GME does not perform well when the errors are drawn from a standardized *t*- distribution. In addition, we find that the GME estimates are very sensitive to the parameter and error support matrices, which we specify *a priori* based on prior information or economic theory.

Golan, Judge, and Miller (1996) specify a block diagonal parameter support matrix, which can be used to specify inequality restrictions such as $\beta_k > 0$. We specify a new parameter support matrix that allows us to specify inequality restrictions involving multiple parameters such as $\beta_k > \beta_j$. The GME estimator does not perform well in Monte Carlo experiments with inequality restrictions placed on the parameters. However, this is due to the impact that the parameter restrictions have on the prior means of the parameters. We discuss this issue further in Chapter 4, where we examine the cost of specifying inequality restrictions in GME.

GME estimation in the linear regression model performs well in many of our Monte Carlo experiments. The data are collinear and GME shrinks the parameters responsible for the collinearity

toward their prior means, which we generally specify to be zero. In addition, GME performs especially well compared to OLS when we have a small sample size. One concern that researchers may have with GME is the fact that we must specify a set of support points for each parameter. However, we find that even when we specify very wide parameter bounds GME still has lower risk measures than OLS in some of our Monte Carlo experiments, although the risk gains are very small. This is analogous to specifying a noninformative prior in Bayesian inference. In Chapter 4, we consider the effects of specifying incorrect prior information. We conclude that GME is a viable estimation technique in the linear regression model, particularly when we have good prior information about the signs or ranges of the parameters, the data are collinear, or we have a small sample size.

3.7 References

Belsley, D., E. Kuh, R. E. Welsch. (1980). *Regression diagnostics.* New York: John Wiley and Sons.

Davidson, R., J.G. MacKinnon. (1993). *Estimation and Inference in Econometrics.* New York: Oxford University Press.

Hill, R. C., P. A. Cartwright, J. F. Arbaugh. (1991). The use of biased predictors in marketing research, *International Journal of Forecasting* 7: 271-282.

Golan, A., G. Judge, D. Miller. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* New York: John Wiley and Sons.

Golan, A., G. Judge, J. Perloff. (1997). Estimation and inference with censored and ordered multinomial response data, *Journal of Econometrics* 79: 23-51.

Hendry, D.F. (1984). Monte Carlo experimentation in econometrics. In Z. Griliches, M.D. Intriligator (Eds.) *Handbook of Econometrics*, Vol. II. Amsterdam, North-Holland: Elsevier Science Publishers: 937-976.

Jaynes, E. T. (1957). Information theory and statistical mechanics, *Physical Review* 106: 620-630.

_______. (1957b). Information theory and statistical mechanics II, *Physical Review* 108: 171-190.

Judge, G. G., W. E. Griffith, R. C. Hill, H. Lutkepohl, T. Lee. (1985). The theory and practice of econometrics 2nd ed. New York: John Wiley and Sons.

Pukelsheim. F. (1994). The three sigma rule, *American Statistician* 48: 88-91.

CHAPTER 4

MONTE CARLO EXPERIMENTS EXAMINING THE COST OF IMPOSING LINEAR INEQUALITY RESTRICTIONS ON THE GME ESTIMATOR

4.1 Introduction

In this chapter we examine GME estimation in the general linear model (GLM) with inequality restrictions placed on the parameters. We carry out Monte Carlo experiments comparing the GME estimator with inequality restrictions to the ordinary least squares (OLS) and inequality restricted least squares (IRLS) estimators. We consider both correct and incorrect inequality restrictions and compare the estimators on the basis of the mean squared error (MSE) of the parameter estimates and the MSE of the predicted values.

In Chapter 3, we discuss GME estimation in the context of linear regression models. Golan, Judge, and Miller (1996) reparameterize the GLM such that the unknown parameters are in the form of probabilities in order to estimate the linear regression model with GME. To do this, we must specify a set of support points for each unknown β_k and estimate the probability that β_k is equal to each of these support points, which are based on nonsample information or economic theory. Thus, the GME estimator always has inequality restrictions since each parameter must be bounded. Since economic theory does not usually provide parameter bounds we consider the cost of imposing incorrect prior information.

We distinguish between the general boundary restrictions, which are necessary in GME, and sign or other inequality restrictions that are not necessary but which reflect prior information. For example, we could constrain the coefficient on price for a normal good to be negative. In this chapter, we examine the cost of specifying inequality restrictions through the GME parameter support matrix. We discuss how to specify prior information through the GME parameter support matrix in section 4.2, describe our sampling experiments in section 4.3, present results in section 4.4, and give conclusions in section 4.5.

4.2 Imposing Inequality Restrictions through the GME Parameter Support Matrix

In Chapter 3, we carry out Monte Carlo experiments in a linear regression model both with and without inequality restrictions placed on the parameters. We show that the block diagonal parameter support matrix developed by Golan, Judge, and Miller (1996) allows us to impose inequality restrictions involving a single parameter and we develop a new parameter support matrix that allows us to impose inequality restrictions involving more than one parameter, such as $\beta_1 > \beta_2$. We impose inequality restrictions that reflect correct prior information about the parameters, which are known in our sampling experiments. However, since we must bound each parameter it is important to know the cost of specifying incorrect nonsample information. In addition, we find that even when we impose correct inequality restrictions the MSE of the GME parameter estimates may increase since the prior means of the unknown parameters change.

Golan, Judge, and Miller (1996, p.140) consider the cost of imposing incorrect sign information about a parameter, but only in a limited way. They estimate a linear regression model using the generalized cross-entropy (GCE) estimator, which allows us to specify prior distributions other than a uniform distribution. Golan et al. specify parameter sign information by placing more prior weight on either the positive or negative parameter support points. The parameter support is given by $Z_k = [-10, 10]$ with prior weights of $[.375 \quad .625]$ or $[.625 \quad .375]$. Thus, the prior mean is equal to 2.5 or –2.5. Golan et al. find that the risk improves only slightly when the prior means are specified to take the correct signs versus when they are specified to take the incorrect signs.

We can impose the same type of sign restrictions in GME by specifying a non-symmetric support vector for the restricted parameter. In GME we can only vary the parameter support points to reflect prior information whereas in GCE we may vary the both the prior weights and the parameter support points to reflect the nonsample information. In both cases we can restrict the prior mean to take a particular sign, but still allow the parameter estimate to take any value between the parameter support bounds. Golan, Judge, and Miller (1996) impose this type of information, but do not actually restrict $\beta_k > 0$ since they specify a parameter support $Z_k = [-10, 10]$ in each case. We consider models with

prior information specified only through the prior means of the parameters as well as models in which we impose binding inequality restrictions.

4.3 Monte Carlo Design

We carry out several Monte Carlo experiments comparing the GME estimator to the OLS and IRLS estimators. We evaluate the estimators in terms of the MSE of the parameter estimates and the MSE of the predicted values. In this section we describe the experimental design and discuss variations in the experimental design. We use the tuna data from Hill, Cartwright, and Arbaugh (1991), which is described in Chapter 3, in our Monte Carlo experiments.

4.3.1 Monte Carlo Experiments

We follow the same Monte Carlo design as in Chapter 3. We generate new data during each Monte Carlo iteration using the true parameter values as

$$y = X\beta_{true} + e \qquad e \sim N(0,1)$$

and $$y_0 = X_0\beta_{true} + e_0 \qquad e_0 \sim N(0,1),$$

where y is an $N \times 1$ vector of generated observations, X is an $N \times K$ matrix of explanatory variables, y_0 is an $N_0 \times 1$ vector of generated observations for a hold-out sample, X_0 is an $N_0 \times K$ matrix of explanatory variables for the hold-out sample, and e and e_0 are vectors of random errors drawn from a standard normal distribution.

In Chapter 3, we impose inequality restrictions that we know to be true. Now we consider both correct and incorrect inequality restrictions in order to examine the cost of imposing inequality restrictions in GME as a function of the constraint specification error. We consider single parameter restrictions, which can be written as

$$\beta_i \geq r_i \qquad \text{for any } i = 1,2,\ldots,K, \tag{4.1}$$

and let $\delta_i = r_i - \beta_i$ represent the constraint specification error. We examine the empirical risk of the IRLS and restricted GME estimators as we vary δ_i.

We consider two types of parameter restrictions. First, following Golan, Judge, and Miller (1996) we vary the prior mean of the GME estimator, but specify the parameter support such that the

true parameter value lies between the boundaries of the parameter support vector. We refer to these "restrictions" as prior mean restrictions. In addition, we impose binding inequality restrictions. We carry out two sets of Monte Carlo experiments and impose inequality restrictions on a single parameter in each set of experiments. We consider the following parameter restrictions:

1) We know the true $\beta_2 = -4.5$ and we impose the series of restrictions, which represent increasing constraint specification error, β_2 > -10, -5, -4.5, -3, -1.5, 0, and 3. Recall that for the prior mean restrictions we only change the prior mean of the parameter estimate, but do not actually impose binding inequality restrictions. The true parameter value lies between the maximum and minimum points of the parameter support vector for β_2.
2) We know the true $\beta_3 = 3.5$ and we impose the series of restrictions, which represent increasing constraint specification error, β_3 < 10, 5, 3.5, 3, 1, 0, and -1. The true parameter value lies between the maximum and minimum points of the parameter support vector for β_3.

Within each set of experiments we vary the sample size, the GME error support, and the GME parameter support. We consider estimation samples of size $N = 150$, $N = 100$, and $N = 50$. Following Golan, Judge, and Perloff (1997) we calculate the sample standard deviation of y assuming it follows a uniform distribution between its minimum and maximum values; the sample standard deviation of y is equal to $s_y = 1.5$ for the tuna data. We construct two GME estimators using error bounds of $(-2s_y \quad 0 \quad 2s_y)$ for our GME2 estimator and $(-3s_y \quad 0 \quad 3s_y)$ for our GME3 estimator. Finally, we specify two different GME parameter supports for each of the binding inequality restrictions. We consider a restricted parameter support with a relatively low prior mean (R1) and one with a higher prior mean (R2) for each restriction. Thus, we have experiments IRLS, GME2-PM, GME2-R1, GME2-R2, GME3-PM, GME3-R1, and GME3-R2 for each restriction, where PM denotes the prior mean only type of restrictions and R1 and R2 denote binding inequality restrictions on the GME estimator.

4.3.2 Dimensions of the Experimental Design

The dimensions of the experimental design include the following:

1) We vary the size of the estimation sample data set. In addition to the sample size of 150 observations we carry out experiments with sample sizes of 100 and 50 observations.

2) We vary the parameter and error support matrices for the GME estimator to reflect different nonsample information.

3) We vary the specification error of the inequality restrictions and compare the empirical risk of the estimators as a function of the specification error.

4.4 Monte Carlo Results

In this section, we present results which examine risk measures for the IRLS and GME estimators in a linear regression model with inequality restrictions placed on the parameter estimates. Section 4.4.1 gives results for experiments with restrictions placed on β_2, section 4.4.2 gives results for experiments with restrictions placed on β_3, and section 4.4.3 gives response surface estimates.

4.4.1 Inequality Restrictions on β_2

In this section, we present results for Monte Carlo experiments with inequality restrictions placed on β_2, which is the coefficient for the regular price of the target brand. The true value of the parameter is $\beta_2 = -4.5$. Table 4.1 summarizes the dimensions for this set of experiments, Table 4.2 summarizes the parameter support for the PM only restrictions, Table 4.3 summarizes the parameter support for R1, and Table 4.4 summarizes the parameter support for R2.

We compare the GME and IRLS estimators to the OLS estimator as we vary the specification error and sample size. We have three different sample sizes and seven different specification errors and we refer to the experiments as IR11-IR17, IR21-IR27, and IR31-IR37. Tables 4.5-4.11 give risk measures of the OLS, IRLS, and GME2 estimators in sampling experiments IR11-IR17. Tables 4.12-4.18 give risk measures of the OLS, IRLS, and GME3 estimators in sampling experiments IR11-IR17. We only present results using a sample size of 150 observations here. The results for sample sizes of 100 and 50 observations are given in Appendix E. As expected, risk measures increase for each of the alternative estimators as we decrease the sample size.

Table 4.1 Dimensions of Monte Carlo Experiments with Restrictions on β_2

Experiment	Sample Size	Restriction	Specification Error
IR 11	150	$\beta_2 \geq -10$	$\delta_t = -5.5$
IR 12	150	$\beta_2 \geq -5$	$\delta_t = -.5$
IR 13	150	$\beta_2 \geq -4.5$	$\delta_t = 0$
IR 14	150	$\beta_2 \geq -3$	$\delta_t = 1.5$
IR 15	150	$\beta_2 \geq -1.5$	$\delta_t = 3$
IR 16	150	$\beta_2 \geq 0$	$\delta_t = 4.5$
IR 17	150	$\beta_2 \geq 3$	$\delta_t = 7.5$
IR 21	100	$\beta_2 \geq -10$	$\delta_t = -5.5$
IR 22	100	$\beta_2 \geq -5$	$\delta_t = -.5$
IR 23	100	$\beta_2 \geq -4.5$	$\delta_t = 0$
IR 24	100	$\beta_2 \geq -3$	$\delta_t = 1.5$
IR 25	100	$\beta_2 \geq -1.5$	$\delta_t = 3$
IR 26	100	$\beta_2 \geq 0$	$\delta_t = 4.5$
IR 27	100	$\beta_2 \geq 3$	$\delta_t = 7.5$
IR 31	50	$\beta_2 \geq -10$	$\delta_t = -5.5$
IR 32	50	$\beta_2 \geq -5$	$\delta_t = -.5$
IR 33	50	$\beta_2 \geq -4.5$	$\delta_t = 0$
IR 34	50	$\beta_2 \geq -3$	$\delta_t = 1.5$
IR 35	50	$\beta_2 \geq -1.5$	$\delta_t = 3$
IR 36	50	$\beta_2 \geq 0$	$\delta_t = 4.5$
IR 37	50	$\beta_2 \geq 3$	$\delta_t = 7.5$

Table 4.2 Support Points and Prior Mean of β_2 for PM Restrictions

Restriction	$z1$	$z2$	$z3$	$z4$	$z5$	Prior Mean
$\beta_2 \geq -10$	-30	-20	-10	0	20	-8
$\beta_2 \geq -5$	-20	-10	-5	0	20	-3
$\beta_2 \geq -4.5$	-20	-10	-5	5	20	-2
$\beta_2 \geq -3$	-20	-10	0	5	20	-1
$\beta_2 \geq -1.5$	-20	-10	0	10	20	0
$\beta_2 \geq 0$	-20	-5	5	10	20	2
$\beta_2 \geq 3$	-20	0	10	20	30	8

Table 4.3 Support Points and Prior Mean of β_2 for R1

Restriction	$z1$	$z2$	$z3$	$z4$	$z5$	Prior Mean
$\beta_2 \geq -10$	-10	-7.5	-5	-2.5	0	-5
$\beta_2 \geq -5$	-5	-2.5	0	2.5	5	0
$\beta_2 \geq -4.5$	-4.5	-2.5	0	2	10	1
$\beta_2 \geq -3$	-3	-1	1	3	10	2
$\beta_2 \geq -1.5$	-1.5	0	1.5	5	10	3
$\beta_2 \geq 0$	0	2.5	5	7.5	10	5
$\beta_2 \geq 3$	3	6	9	12	15	9

Table 4.4 Support Points and Prior Mean of β_2 for R2

Restriction	$z1$	$z2$	$z3$	$z4$	$z5$	Prior Mean
$\beta_2 \geq -10$	-10	-5	0	5	10	0
$\beta_2 \geq -5$	-5	-2	0	5	12	2
$\beta_2 \geq -4.5$	-4.5	-1.5	0	6	15	3
$\beta_2 \geq -3$	-3	0	3	5	15	4
$\beta_2 \geq -1.5$	-1.5	0	5	10	20	6.7
$\beta_2 \geq 0$	0	5	10	15	20	10
$\beta_2 \geq 3$	3	10	15	20	30	15.6

Table 4.5 Monte Carlo IR 11 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	49.40	36.39	29.85	40.24
RMSE ($\hat{\beta}$)	7.04	7.03	6.03	5.46	6.34
Var ($\hat{\beta}$)	49.44	49.35	30.63	24.67	27.43
Bias^2($\hat{\beta}$)	0.05	0.05	5.76	5.18	12.81
PRMSE	11.58	11.58	11.64	11.66	11.68
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.52	7.50	7.53
R^2	0.77				

* N=150; $\beta_2 \geq -10$; $\delta_t = -5.5$; normal errors.

Table 4.6 Monte Carlo IR 12 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	45.25	36.77	48.60	52.74
RMSE ($\hat{\beta}$)	7.04	6.73	6.06	6.97	7.26
Var ($\hat{\beta}$)	49.44	44.33	29.93	25.22	26.07
Bias^2($\hat{\beta}$)	0.05	0.92	6.84	23.39	26.67
PRMSE	11.58	11.59	11.64	11.74	11.75
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.55	7.56
R^2	0.77				

* N=150; $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table 4.7 Monte Carlo IR 13 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	45.04	37.25	51.39	55.48
RMSE ($\hat{\beta}$)	7.04	6.71	6.10	7.17	7.45
Var ($\hat{\beta}$)	49.44	43.50	30.01	25.57	26.15
Bias^2($\hat{\beta}$)	0.05	1.53	7.23	25.82	29.33
PRMSE	11.58	11.60	11.64	11.75	11.76
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.56	7.57
R^2	0.77				

* N=150; $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table 4.8 Monte Carlo IR 14 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	47.33	37.73	63.13	69.06
RMSE ($\hat{\beta}$)	7.04	6.88	6.14	7.95	8.31
Var ($\hat{\beta}$)	49.44	41.49	30.03	25.41	25.88
Bias^2($\hat{\beta}$)	0.05	5.84	7.70	37.72	43.19
PRMSE	11.58	11.63	11.65	11.81	11.83
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.52	7.60	7.61
R^2	0.77				

* N=150; $\beta_2 \geq -3$; $\delta_t = 1.5$; normal errors.

Table 4.9 Monte Carlo IR 15 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	56.29	38.25	77.42	79.35
RMSE ($\hat{\beta}$)	7.04	7.50	6.18	8.80	8.91
Var ($\hat{\beta}$)	49.44	40.52	30.19	25.43	26.01
Bias^2($\hat{\beta}$)	0.05	15.77	8.07	51.98	53.34
PRMSE	11.58	11.69	11.65	11.88	11.88
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.55	7.52	7.64	7.64
R^2	0.77				

* N=150; $\beta_2 \geq -1.5$; $\delta_t = 3$; normal errors.

Table 4.10 Monte Carlo IR 16 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	73.15	39.43	108.35	106.72
RMSE ($\hat{\beta}$)	7.04	8.55	6.28	10.41	10.33
Var ($\hat{\beta}$)	49.44	40.17	30.26	25.89	26.49
Bias^2($\hat{\beta}$)	0.05	32.98	9.18	82.47	80.23
PRMSE	11.58	11.77	11.65	12.03	12.01
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.60	7.52	7.73	7.72
R^2	0.77				

* N=150; $\beta_2 \geq 0$; $\delta_t = 4.5$; normal errors.

Table 4.11 Monte Carlo IR 17 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	130.35	41.30	167.81	152.86
RMSE ($\hat{\beta}$)	7.04	11.42	6.43	12.95	12.36
Var ($\hat{\beta}$)	49.44	40.03	30.84	26.73	26.69
Bias^2($\hat{\beta}$)	0.05	90.33	10.46	141.09	126.17
PRMSE	11.58	12.04	11.66	12.31	12.24
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
$PRMSE_0$	7.54	7.75	7.53	7.90	7.85
R^2	0.77				

* N=150; $\beta_2 \geq 3 ; \delta_t = 7.5$; normal errors.

Table 4.12 Monte Carlo IR 11 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	49.40	32.67	27.02	45.76
RMSE ($\hat{\beta}$)	7.04	7.03	5.72	5.20	6.76
Var ($\hat{\beta}$)	49.44	49.35	18.47	14.49	15.73
Bias^2($\hat{\beta}$)	0.05	0.05	14.21	12.53	30.04
PRMSE	11.58	11.58	11.64	11.67	11.72
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
$PRMSE_0$	7.54	7.53	7.47	7.46	7.51
R^2	0.77				

* N=150; $\beta_2 \geq -10 ; \delta_t = -5.5$; normal errors.

Table 4.13 Monte Carlo IR 12 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	45.25	35.37	53.77	65.98
RMSE ($\hat{\beta}$)	7.04	6.73	5.95	7.33	8.12
Var ($\hat{\beta}$)	49.44	44.33	17.69	14.77	15.24
Bias^2($\hat{\beta}$)	0.05	0.92	17.67	39.00	50.74
PRMSE	11.58	11.59	11.66	11.78	11.82
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.48	7.53	7.57
R^2	0.77				

* N=150; $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table 4.14 Monte Carlo IR 13 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	45.04	36.88	60.24	71.79
RMSE ($\hat{\beta}$)	7.04	6.71	6.07	7.76	8.47
Var ($\hat{\beta}$)	49.44	43.50	17.76	15.00	15.39
Bias^2($\hat{\beta}$)	0.05	1.53	19.12	45.24	56.40
PRMSE	11.58	11.60	11.67	11.80	11.85
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.48	7.55	7.59
R^2	0.77				

* N=150; $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table 4.15 Monte Carlo IR 14 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	47.33	38.51	74.93	90.05
RMSE ($\hat{\beta}$)	7.04	6.88	6.21	8.66	9.49
Var ($\hat{\beta}$)	49.44	41.49	17.76	14.92	15.19
Bias^2($\hat{\beta}$)	0.05	5.84	20.75	60.01	74.86
PRMSE	11.58	11.63	11.67	11.87	11.94
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.49	7.59	7.64
R^2	0.77				

* N=150; $\beta_2 \geq -3$; $\delta_t = 1.5$; normal errors.

Table 4.16 Monte Carlo IR 15 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	56.29	40.15	92.25	109.72
RMSE ($\hat{\beta}$)	7.04	7.50	6.34	9.60	10.47
Var ($\hat{\beta}$)	49.44	40.52	17.92	14.90	15.42
Bias^2($\hat{\beta}$)	0.05	15.77	22.23	77.35	94.30
PRMSE	11.58	11.69	11.68	11.95	12.02
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.55	7.49	7.65	7.70
R^2	0.77				

* N=150; $\beta_2 \geq -1.5$; $\delta_t = 3$; normal errors.

Table 4.17 Monte Carlo IR 16 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	73.15	44.25	134.21	160.65
RMSE ($\hat{\beta}$)	7.04	8.55	6.65	11.58	12.67
Var ($\hat{\beta}$)	49.44	40.17	17.95	14.93	15.53
Bias^2($\hat{\beta}$)	0.05	32.98	26.30	119.28	145.12
PRMSE	11.58	11.77	11.70	12.15	12.26
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.60	7.51	7.77	7.84
R^2	0.77				

* N=150; $\beta_2 \geq 0 : \delta_t = 4.5$; normal errors.

Table 4.18 Monte Carlo IR 17 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	130.35	51.14	216.98	223.78
RMSE ($\hat{\beta}$)	7.04	11.42	7.15	14.73	14.96
Var ($\hat{\beta}$)	49.44	40.03	18.66	15.05	15.53
Bias^2($\hat{\beta}$)	0.05	90.33	32.49	201.92	208.24
PRMSE	11.58	12.04	11.72	12.52	12.54
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.75	7.53	8.00	8.02
R^2	0.77				

* N=150; $\beta_2 \geq 3 ; \delta_t = 7.5$; normal errors.

The Monte Carlo results indicate that as the specification error $\delta_t \to -\infty$, the risk of the IRLS estimator converges to the risk of the OLS estimator (Judge et al. 1985, p. 824). In our experiments, the empirical risk of the IRLS estimator is minimized when $\delta_t = 0$. In contrast, the empirical risk of the GME estimator increases with the specification error over the entire range of δ_t that we examined. As noted earlier, this is due to the impact of the prior mean as well as the inequality restriction on the

parameter estimates. When we change the parameter support to impose an inequality restriction in GME we affect more than just the minimum value that the parameter estimate can take. The risk of the IRLS estimator compared to the OLS estimator is a function of the constraint specification error while the risk of the GME estimator is a function of both the constraint specification error and the prior mean of the parameter estimate. In section 4.4.3, we estimate response functions to examine the relative importance of these effects.

The results show that the risk of the GME estimator with PM only restrictions is lower than the risk of the OLS estimator for all of our experiments. Golan, Judge, and Miller (1996) note that incorrectly specifying the prior mean has little impact on the risk of the GME estimator. This is true as long as we specify a relatively wide parameter support vector. The GME-R1 and GME-R2 estimators have the lowest empirical risk of the alternative estimators when the specification error is very small. However, the risk rises very quickly for these estimators as the specification error increases. Figures 4.1 and 4.2 graph the empirical risk of the IRLS, GME2-R1, and GME2-R2 estimators compared to the OLS estimator, using sample sizes of 150 and 100, as we vary the constraint specification error.

4.4.2 Inequality Restrictions on β_3

In this section, we present results for Monte Carlo experiments with inequality restrictions placed on β_3, which is the coefficient for the discount percentage of the target brand. The true value of the parameter is $\beta_3 = 3.5$. Table 4.19 summarizes the dimensions for this set of experiments, Table 4.20 summarizes the parameter support for the PM only restrictions, Table 4.21 summarizes the parameter support for R1, and Table 4.22 summarizes the parameter support for R2.

We compare the GME and IRLS estimators to the OLS estimator as we vary the specification error and the sample size. Tables 4.23-4.29 give risk measures of the OLS, IRLS, and GME2 estimators in sampling experiments IR11-IR17. Tables 4.30-4.36 give risk measures of the OLS, IRLS, and GME3 estimators in sampling experiments IR11-IR17. We only present results using a sample size of 150 observations here. The results for sample sizes of 100 and 50 observations are given in Appendix E.

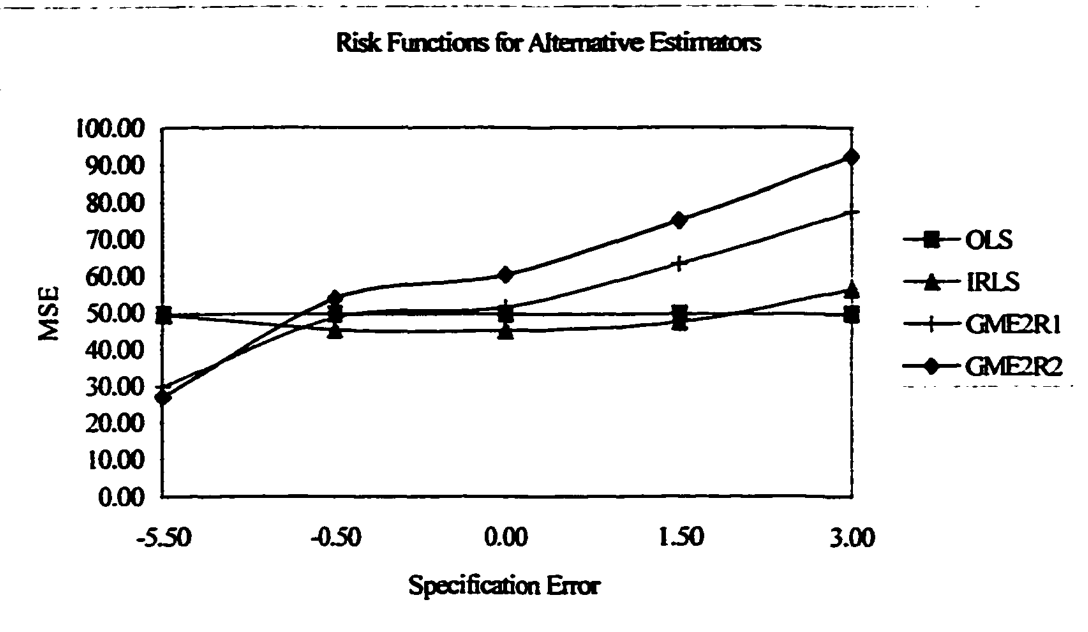

Figure 4.1 Empirical Risk of OLS, IRLS, and GME Estimators (N=150)

* Restrictions on β_2

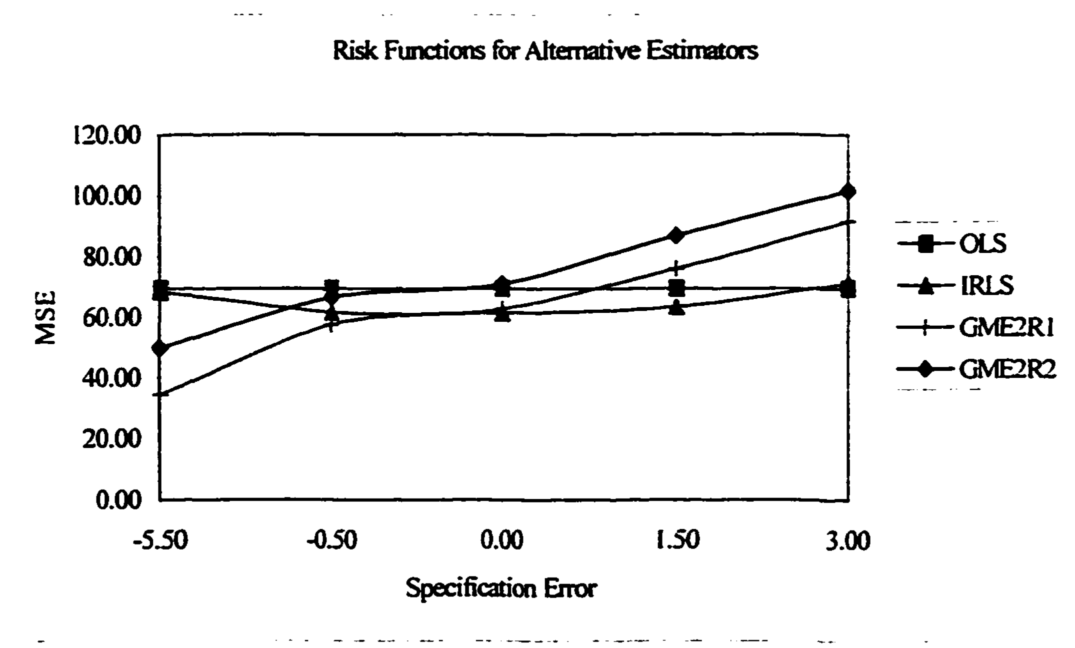

Figure 4.2 Empirical Risk of OLS, IRLS, and GME Estimators (N=100)

* Restrictions on β_2

Table 4.19 Dimensions of Monte Carlo Experiments with Restrictions on β_3

Experiment	Sample Size	Restriction	Specification Error
IR 11	150	$\beta_3 \leq 10$	$\delta_t = -6.5$
IR 12	150	$\beta_3 \leq 5$	$\delta_t = -1.5$
IR 13	150	$\beta_3 \leq 3.5$	$\delta_t = 0$
IR 14	150	$\beta_3 \leq 3$	$\delta_t = .5$
IR 15	150	$\beta_3 \leq 1$	$\delta_t = 2.5$
IR 16	150	$\beta_3 \leq 0$	$\delta_t = 3.5$
IR 17	150	$\beta_3 \leq -1$	$\delta_t = 4.5$
IR 21	100	$\beta_3 \leq 10$	$\delta_t = -6.5$
IR 22	100	$\beta_3 \leq 5$	$\delta_t = -1.5$
IR 23	100	$\beta_3 \leq 3.5$	$\delta_t = 0$
IR 24	100	$\beta_3 \leq 3$	$\delta_t = .5$
IR 25	100	$\beta_3 \leq 1$	$\delta_t = 2.5$
IR 26	100	$\beta_3 \leq 0$	$\delta_t = 3.5$
IR 27	100	$\beta_3 \leq -1$	$\delta_t = 4.5$
IR 31	50	$\beta_3 \leq 10$	$\delta_t = -6.5$
IR 32	50	$\beta_3 \leq 5$	$\delta_t = -1.5$
IR 33	50	$\beta_3 \leq 3.5$	$\delta_t = 0$
IR 34	50	$\beta_3 \leq 3$	$\delta_t = .5$
IR 35	50	$\beta_3 \leq 1$	$\delta_t = 2.5$
IR 36	50	$\beta_3 \leq 0$	$\delta_t = 3.5$
IR 37	50	$\beta_3 \leq -1$	$\delta_t = 4.5$

Table 4.20 Support Points and Prior Mean of β_3 for PM Restrictions

Restriction	$z1$	$z2$	$z3$	$z4$	$z5$	Prior Mean
$\beta_3 \leq 10$	-20	0	10	20	30	8
$\beta_3 \leq 5$	-20	0	5	15	20	4
$\beta_3 \leq 3.5$	-20	0	5	10	20	3
$\beta_3 \leq 3$	-20	-5	0	10	20	1
$\beta_3 \leq 1$	-20	-10	0	10	20	0
$\beta_3 \leq 0$	-20	-15	0	5	20	-2
$\beta_3 \leq -1$	-20	-15	-5	0	20	-4

Table 4.21 Support Points and Prior Mean of β_3 for R1

Restriction	$z1$	$z2$	$z3$	$z4$	$z5$	Prior Mean
$\beta_3 \leq 10$	0	2.5	5	7.5	10	5
$\beta_3 \leq 5$	-5	-2.5	0	2.5	5	0
$\beta_3 \leq 3.5$	-10	-5	0	1.5	3.5	-2
$\beta_3 \leq 3$	-10	-5	-3	0	3	-3
$\beta_3 \leq 1$	-15	-10	-1	0	1	-5
$\beta_3 \leq 0$	-15	-10	-5	-3	0	-6.6
$\beta_3 \leq -1$	-15	-12	-9	-5	-1	-8.4

Table 4.22 Support Points and Prior Mean of β_3 for R2

Restriction	$z1$	$z2$	$z3$	$z4$	$z5$	Prior Mean
$\beta_3 \leq 10$	-10	-5	0	5	10	0
$\beta_3 \leq 5$	-12	-5	0	2	5	-2
$\beta_3 \leq 3.5$	-15	-5	0	1.5	3.5	-3
$\beta_3 \leq 3$	-15	-10	-5	0	3	-5.4
$\beta_3 \leq 1$	-20	-10	-5	-1	1	-7
$\beta_3 \leq 0$	-20	-15	-10	-5	0	-10
$\beta_3 \leq -1$	-25	-20	-10	-5	-1	-12.2

Table 4.23 Monte Carlo IR 11 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	49.49	38.02	36.96	38.58
RMSE ($\hat{\beta}$)	7.04	7.04	6.17	6.08	6.21
Var ($\hat{\beta}$)	49.44	49.44	30.21	29.43	29.97
Bias^2($\hat{\beta}$)	0.05	0.05	7.82	7.54	8.60
PRMSE	11.58	11.58	11.65	11.66	11.65
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.54	7.52	7.50	7.52
R^2	0.77				

* N=150; $\beta_3 \leq 10$; $\delta_t = -6.5$; normal errors.

Table 4.24 Monte Carlo IR 12 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	49.27	38.08	40.83	40.80
RMSE ($\hat{\beta}$)	7.04	7.02	6.17	6.39	6.39
Var ($\hat{\beta}$)	49.44	49.22	30.18	29.52	29.62
Bias^2($\hat{\beta}$)	0.05	0.05	7.90	11.31	11.18
PRMSE	11.58	11.58	11.65	11.71	11.71
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.52	7.55	7.55
R^2	0.77				

* N=150; $\beta_3 \leq 5$; $\delta_t = -1.5$; normal errors.

Table 4.25 Monte Carlo IR 13 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	48.71	38.11	43.43	43.19
RMSE ($\hat{\beta}$)	7.04	6.98	6.17	6.59	6.57
Var ($\hat{\beta}$)	49.44	48.39	30.17	29.50	29.48
Bias^2 ($\hat{\beta}$)	0.05	0.32	7.94	13.93	13.71
PRMSE	11.58	11.60	11.65	11.77	11.77
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.60	7.59
R^2	0.77				

* N=150; $\beta_3 \leq 3.5 : \delta_t = 0$; normal errors.

Table 4.26 Monte Carlo IR 14 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	48.96	38.21	45.73	44.71
RMSE ($\hat{\beta}$)	7.04	7.00	6.18	6.76	6.69
Var ($\hat{\beta}$)	49.44	48.13	30.18	29.64	29.61
Bias^2 ($\hat{\beta}$)	0.05	0.82	8.03	16.09	15.10
PRMSE	11.58	11.62	11.65	11.82	11.80
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.64	7.62
R^2	0.77				

* N=150; $\beta_3 \leq 3 ; \delta_t = 0.5$; normal errors.

Table 4.27 Monte Carlo IR 15 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	56.96	38.25	55.44	55.46
RMSE ($\hat{\beta}$)	7.04	7.55	6.18	7.45	7.45
Var ($\hat{\beta}$)	49.44	47.78	30.19	30.12	30.10
Bias^2($\hat{\beta}$)	0.05	9.18	8.07	25.33	25.36
PRMSE	11.58	11.86	11.65	12.07	12.07
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.69	7.52	7.81	7.81
R^2	0.77				

* N=150; $\beta_3 \leq 1$; $\delta_t = 2.5$; normal errors.

Table 4.28 Monte Carlo IR 16 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	65.62	38.34	65.11	62.36
RMSE ($\hat{\beta}$)	7.04	8.10	6.19	8.07	7.90
Var ($\hat{\beta}$)	49.44	47.80	30.19	31.08	30.89
Bias^2($\hat{\beta}$)	0.05	17.82	8.15	34.04	31.48
PRMSE	11.58	12.09	11.65	12.28	12.22
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.86	7.52	7.97	7.92
R^2	0.77				

* N=150; $\beta_3 \leq 0$; $\delta_t = 3.5$; normal errors.

Table 4.29 Monte Carlo IR 17 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	77.24	38.44	77.77	77.64
RMSE ($\hat{\beta}$)	7.04	8.79	6.20	8.82	8.81
Var ($\hat{\beta}$)	49.44	47.80	30.20	33.72	33.71
Bias^2($\hat{\beta}$)	0.05	29.44	8.24	44.05	43.93
PRMSE	11.58	12.38	11.65	12.54	12.54
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	8.08	7.52	8.15	8.15
R^2	0.77				

* N=150; $\beta_3 \leq -1$; $\delta_t = 4.5$; normal errors.

Table 4.30 Monte Carlo IR 11 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	49.49	39.32	37.88	41.53
RMSE ($\hat{\beta}$)	7.04	7.04	6.27	6.15	6.44
Var ($\hat{\beta}$)	49.44	49.44	17.97	17.04	17.55
Bias^2($\hat{\beta}$)	0.05	0.05	21.35	20.84	23.98
PRMSE	11.58	11.58	11.68	11.71	11.70
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.54	7.49	7.47	7.51
R^2	0.77				

* N=150; $\beta_3 \leq 10$; $\delta_t = -6.5$; normal errors.

Table 4.31 Monte Carlo IR 12 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	49.27	39.55	46.66	47.29
RMSE ($\hat{\beta}$)	7.04	7.02	6.29	6.83	6.88
Var ($\hat{\beta}$)	49.44	49.22	17.92	17.08	17.22
Bias^2($\hat{\beta}$)	0.05	0.05	21.63	29.58	30.06
PRMSE	11.58	11.58	11.68	11.82	11.83
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.49	7.59	7.60
R^2	0.77				

* N=150; $\beta_3 \leq 5$; $\delta_t = -1.5$; normal errors.

Table 4.32 Monte Carlo IR 13 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	48.71	39.67	51.25	50.62
RMSE ($\hat{\beta}$)	7.04	6.98	6.30	7.16	7.12
Var ($\hat{\beta}$)	49.44	48.39	17.90	17.10	17.11
Bias^2($\hat{\beta}$)	0.05	0.32	21.77	34.16	33.52
PRMSE	11.58	11.60	11.68	11.93	11.91
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.49	7.67	7.66
R^2	0.77				

* N=150; $\beta_3 \leq 3.5$; $\delta_t = 0$; normal errors.

Table 4.33 Monte Carlo IR 14 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE $(\hat{\beta})$	49.49	48.96	40.00	56.48	55.07
RMSE $(\hat{\beta})$	7.04	7.00	6.32	7.52	7.42
Var $(\hat{\beta})$	49.44	48.13	17.90	17.13	17.21
$\text{Bias}^2(\hat{\beta})$	0.05	0.82	22.10	39.35	37.86
PRMSE	11.58	11.62	11.68	12.05	12.01
PRMSE $(\hat{p}+\hat{e})$			0.00	0.00	0.00
PRMSE_0	7.54	7.52	7.49	7.77	7.75
R^2	0.77				

* N=150; $\beta_3 \le 3 : \delta_t = 0.5$; normal errors.

Table 4.34 Monte Carlo IR 15 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE $(\hat{\beta})$	49.49	56.96	40.15	62.35	66.29
RMSE $(\hat{\beta})$	7.04	7.55	6.34	7.90	8.14
Var $(\hat{\beta})$	49.44	47.78	17.92	17.02	17.09
$\text{Bias}^2(\hat{\beta})$	0.05	9.18	22.23	45.33	49.20
PRMSE	11.58	11.86	11.68	12.20	12.29
PRMSE $(\hat{p}+\hat{e})$			0.00	0.00	0.00
PRMSE_0	7.54	7.69	7.49	7.88	7.96
R^2	0.77				

* N=150; $\beta_3 \le 1 : \delta_t = 2.5$; normal errors.

Table 4.35 Monte Carlo IR 16 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	65.62	40.48	78.16	74.36
RMSE ($\hat{\beta}$)	7.04	8.10	6.36	8.84	8.62
Var ($\hat{\beta}$)	49.44	47.80	17.93	17.14	17.17
$\text{Bias}^2(\hat{\beta})$	0.05	17.82	22.55	61.02	57.20
PRMSE	11.58	12.09	11.68	12.58	12.48
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	7.54	7.86	7.50	8.18	8.11
R^2	0.77				

* N=150; $\beta_3 \leq 0$; $\delta_t = 3.5$; normal errors.

Table 4.36 Monte Carlo IR 17 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	77.24	40.83	87.22	85.06
RMSE ($\hat{\beta}$)	7.04	8.79	6.39	9.34	9.22
Var ($\hat{\beta}$)	49.44	47.80	17.94	17.22	17.17
$\text{Bias}^2(\hat{\beta})$	0.05	29.44	22.89	70.00	67.89
PRMSE	11.58	12.38	11.68	12.80	12.75
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	7.54	8.08	7.50	8.35	8.31
R^2	0.77				

* N=150; $\beta_3 \leq -1$; $\delta_t = 4.5$; normal errors.

The results for the experiments with restrictions on β_3 are very similar to those for the experiments with restrictions on β_2. The main difference is that the empirical risk does not increase as quickly as we increase the specification error in this set of experiments. Note that in both sets of experiments, the variance of the alternative estimators does not change much as we increase the specification error. However, the bias of the alternative estimators increases very rapidly for the GME

estimators with binding restrictions as we increase the specification error. Figures 4.3 and 4.4 graph the empirical risk of the IRLS, GME2-R1, and GME2-R2 estimators compared to the OLS estimator as we vary the constraint specification error.

4.4.3 Response Surfaces for Monte Carlo Experiments

In this section, we estimate response surfaces for our Monte Carlo experiments. Hendry (1984) and Davidson and MacKinnon (1993) discuss response surfaces as a means of summarizing the results from a set of Monte Carlo experiments. We estimate the following response surface regression

$$\frac{MSE(\hat{\beta})}{MSE(OLS)} = \alpha_1 + \alpha_2 N + \alpha_3 \delta_i + \alpha_4 \delta_i^2 + \alpha_5 \rho_i + \alpha_6 \rho_i^2 + \mu_i, \qquad i = 1, \ldots, M \tag{4.2}$$

where N is the sample size of the estimation sample, δ_i is the constraint specification error, $\rho_i = PM_i - \beta_i$, M is the number of Monte Carlo experiments, and $MSE(\hat{\beta})$ is the MSE of the estimator of interest, either IRLS or GME. We include the squared terms since the MSE of the IRLS estimator is minimized when $\delta_i = 0$ and the MSE of the GME estimator should be minimized for $\rho_i = 0$ (which would shrink the parameter estimates toward the true parameter value). Thus, we do not expect a linear relationship between MSE and δ_i or ρ_i. We estimate response surfaces for IRLS, GME2-PM, and GME2-R (where GME2-R includes observations for both GME2-R1 and GME2-R2). Note that for the IRLS regressions $\rho_i = 0$ and for the GME-PM regressions $\delta_i = 0$. Tables 4.37 and 4.38 summarize the response surface estimates.

The response surface shows that the ratio of the MSE for the restricted estimator to the MSE for the OLS estimator increases as the sample size increases. Thus, the risk gains due to inequality restrictions in either IRLS or GME are greater the smaller the sample size. As expected, increasing the constraint specification error, δ, increases the risk for our restricted estimators. Additionally, the sign of the coefficient for the prior mean specification error, ρ, is positive but not significant. This is consistent with the observation by Golan, Judge, and Miller (1996) that changing the prior mean has only a small impact on risk measures for the GME estimator. Finally, the coefficient for the constraint specification error is greater for GME than for IRLS. This is because when we change the constraint

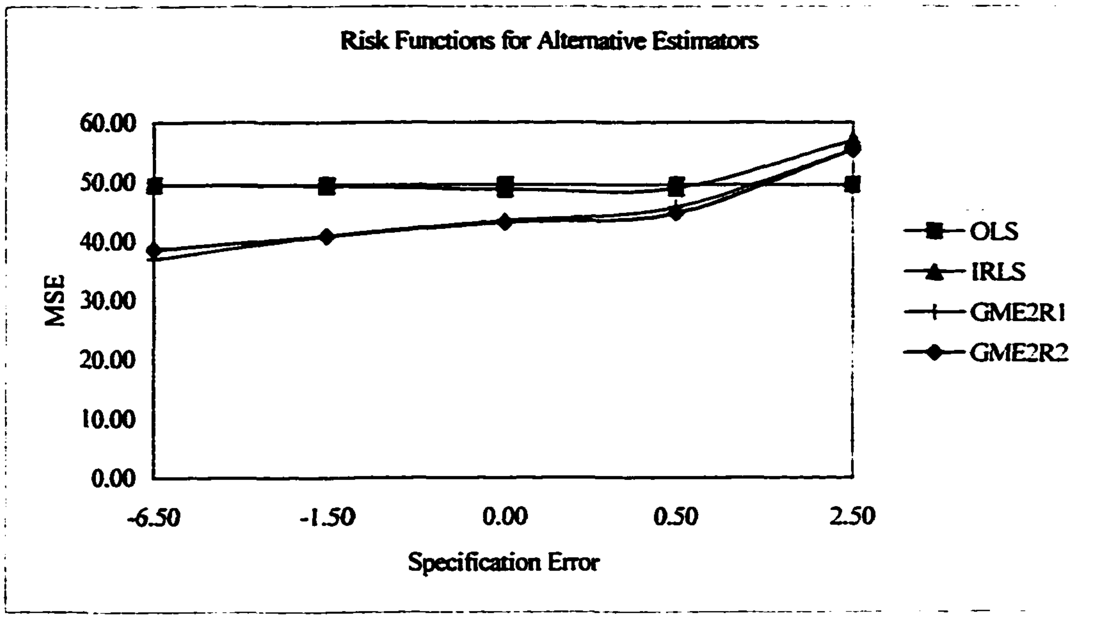

Figure 4.3 Empirical Risk of OLS, IRLS, and GME Estimators (N=150)

* Restrictions on β_3

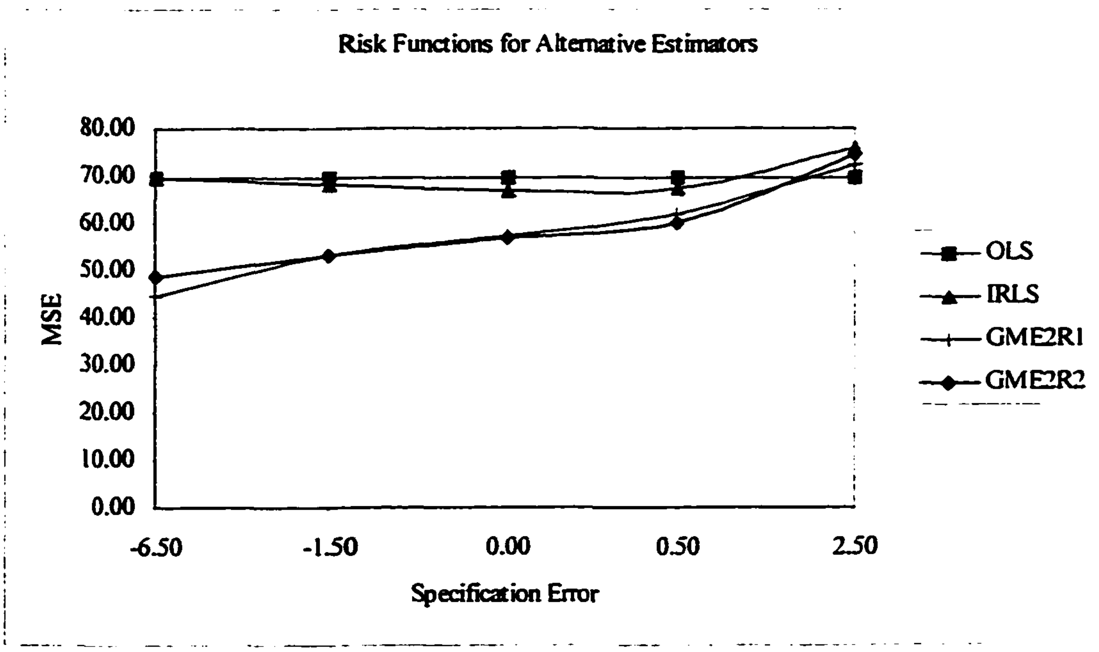

Figure 4.4 Empirical Risk of OLS, IRLS, and GME Estimators (N=100)

* Restrictions on β_3

Table 4.37 Response Surfaces for Monte Carlo Experiments with Restrictions on β_2

Variable	IRLS	GME2-PM	GME2-R
Intercept	.5645**	.1798**	-.3294
	(.1277)	(.0357)	(.2279)
Sample Size	.0029**	.0040**	.0088**
	(.0011)	(.0003)	(.0010)
δ_t	.0444**	——	.0937**
	(.0130)		(.0244)
δ_t^2	.0136**	——	.0144**
	(.0026)		(.0032)
ρ_t	——	.0059	.0546
		(.0052)	(.0403)
ρ_t^2	——	.0003	-.0017
		(.0005)	(.0017)
R^2	0.80	0.91	0.90
$\hat{\sigma}^2$	0.21	0.06	0.25
M	21	21	42

* Standard errors in parentheses.

Table 4.38 Response Surfaces for Monte Carlo Experiments with Restrictions on β_3

Variable	IRLS	GME2-PM	GME2-R
Intercept	.8247**	.2126**	.1688
	(.0549)	(.0348)	(.1123)
Sample Size	.0015**	.0039**	.0053**
	(.0005)	(.0003)	(.0005)
δ_t	.0423**	——	.0677**
	(.0064)		(.0175)
δ_t^2	.0074**	——	.0081**
	(.0016)		(.0024)
ρ_t	——	.0057	.0054
		(.0042)	(.0206)
ρ_t^2	——	.0002	-.0003
		(.0008)	(.0011)
R^2	0.77	0.91	0.87
$\hat{\sigma}^2$	0.09	0.06	0.13
M	21	21	42

* Standard errors in parentheses.

specification error in GME we must also change the parameter support. The GME estimator is sensitive to changes in the parameter support since the prior mean may change and changes in the parameter support also affect the relative weights placed on the unknown parameters and errors in GME.

4.5 Conclusions

In this chapter, we examined the cost of imposing inequality restrictions on the GME estimator in a linear regression problem. This is an important question in a linear regression problem since we must impose a maximum and minimum value on each parameter estimate in order to obtain GME estimates. We find that when the constraint specification error is negative (i.e. the restriction is correct) the GME estimator has lower risk than the OLS and IRLS estimators do. However, even when we impose correct inequality restrictions in GME, the risk measures for GME may be higher than for OLS or IRLS since the prior means and parameter support points change. Changing the prior mean shrinks the GME estimator towards a different value while changing the parameter support points changes the relative importance of the GME parameter and error weights. For example, if we specify a very narrow range for the parameter support, GME places relatively more weight on the unknown errors than the data. However, if we specify a wide parameter support the GME error estimates are very small. We used relatively narrow parameter support vectors to keep the prior means from becoming too large.

GME performs best when the constraint specification error is very small, which implies that we should use GME if we are confident that the prior information is correct. In addition, the GME-PM type of restriction provides a good alternative if we are not very sure of our prior information. Thus, GME-PM is a conservative approach to imposing prior information in GME. However, GME-PM does not impose binding inequality restrictions.

Golan, Judge, and Miller (1996) estimate linear regression models using a GCE estimator, which allows us to specify different prior weights on the support points, but they do not impose binding inequality restrictions in their experiments. However, GCE may allow an improved way to impose inequality restrictions using entropy. Under GCE, we could specify binding restrictions and a relatively wide error support, yet still keep the prior mean consistent with our prior information by placing more prior weight on certain parameter values. This is an area where we plan to extend this research.

4.6 References

Davidson, R., J.G. MacKinnon. (1993). *Estimation and Inference in Econometrics.* New York: Oxford University Press.

Golan, A., G. Judge, D. Miller. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* New York: John Wiley and Sons.

Golan, A., G. Judge, J. M. Perloff. (1997). Estimation and inference with censored and ordered multinomial response data, *Journal of Econometrics* 79: 23-51.

Hendry, D.F. (1984). Monte Carlo experimentation in econometrics. In Z. Griliches, M.D. Intriligator (Eds.) *Handbook of Econometrics*, Vol. II. Amsterdam, North-Holland: Elsevier Science Publishers, 937-976.

Hill, R. C., P. A. Cartwright, J. F. Arbaugh. (1991). The use of biased predictors in marketing research, *International Journal of Forecasting* 7: 271-282.

Judge, G. G., R.C. Hill, W. E. Griffith, H. Lutkepohl, T. Lee. (1988). Introduction to the theory and practice of econometrics 2nd ed. New York: John Wiley and Sons.

CHAPTER 5

MAXIMUM ENTROPY ESTIMATION IN CENSORED AND TRUNCATED REGRESSION MODELS

5.1 Introduction

In this chapter, we discuss ME estimation in censored and truncated regression models. Censored and truncated models are known as limited dependent variable models since we have limited observations for the dependent variable. In censored samples, all values of the dependent variable below a certain threshold are reported at the threshold value. Thus, we do not observe the true value of the dependent variable. However, we fully observe the independent variables in a censored sample. In a truncated sample, we only observe the data when the dependent variable, y, is above the threshold value; we observe neither the dependent nor the independent variables when y is below the threshold.

We estimate a censored regression model using the GME estimator discussed in Golan, Judge, and Perloff (1997) and Golan, Judge, and Miller (1996). We develop a GME estimator for two types of models involving self-selectivity, which is a special case of truncation. Self-selection means that individuals are either in the sample or not based on the outcome of a separate decision model. The decision may be made by the individual or by someone else. For example, in our credit scoring model individuals who are denied loans by the bank are not in the sample. In addition, individuals who choose not to apply for a loan because there is a high probability that they would be denied are not in the sample.

5.2 Generalized Maximum Entropy Estimation in a Censored Regression Model

In censored samples, we have observations on the independent variables for all individuals, but we do not observe the true dependent variable for individuals below a threshold value. Instead, the observed dependent variable is set to the threshold value for individuals below the threshold. For example, in measuring consumer expenditures on automobiles during a given year, we observe many households whose expenditures are zero. Tobin (1958) proposed a solution to this problem, which is

known as the censored regression model or the tobit model. The tobit formulation assumes a linear regression model

$$y_i^* = x_i'\beta + e_i, \qquad e_i \sim N(0,\sigma^2) \tag{5.1}$$

for y_i^*, which is an unobservable latent variable that represents the desired expenditure level. We observe

$$y_i = \begin{cases} y_i^* & \text{if } y_i^* > 0 \\ 0 & \text{if } y_i^* \leq 0. \end{cases} \qquad i = 1,\ldots,N \tag{5.2}$$

We have N_0 observations where $y_i = 0$ and N_1 observations where $y_i = y_i^*$ ($N = N_0 + N_1$). Judge et al. (1988, p. 797) show that the OLS estimator applied to the N_1 uncensored observations is biased and inconsistent. The OLS estimator applied to all N observations is also biased and inconsistent. This is because the unconditional expectation of y_i for a randomly drawn observation is

$$E[y_i|x_i] = \Phi\left(\frac{x_i'\beta}{\sigma}\right)(x_i'\beta + \sigma\lambda_i),$$

where

$$\lambda_i = \frac{\phi(x_i'\beta/\sigma)}{\Phi(x_i'\beta/\sigma)}. \tag{5.3}$$

$\Phi(z)$ denotes the standard normal cumulative distribution function evaluated at z and $\phi(z)$ denotes the standard normal probability density function evaluated at z. Thus, OLS estimation on y_i does not yield consistent estimates for the population regression function $E[y_i^*] = x_i'\beta$. Therefore, we use maximum likelihood to obtain the tobit estimates.

The probability that an observation is censored is given by

$$\text{Prob}[y_i = 0] = 1 - \Phi(x_i'\beta/\sigma) = 1 - F_i, \tag{5.4}$$

while the probability that an observation is not censored is $\Phi(x_i'\beta/\sigma) = F_i$. The probability density function for individuals who do make a purchase is

$$f(y_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y_t - x_t'\beta)^2}{\sigma^2}\right\},$$

and the log-likelihood function is given by

$$\ln L = \sum_{y_t=0} \ln(1 - F_t) - \sum_{y_t>0} \frac{1}{2}\left[\ln(2\pi) + \ln\sigma^2 + \frac{(y_t - x_t'\beta)^2}{\sigma^2}\right]. \tag{5.5}$$

Amemiya (1973b) showed that maximizing the log-likelihood function (5.5) leads to estimators that are consistent and asymptotically efficient.

Golan, Judge, and Perloff (1997) discuss GME estimation of the tobit model. Their GME solution for the censored regression problem closely follows the Golan, Judge, and Miller (1996) GME solution for a linear regression problem, which we discuss in Chapter 3. As in the linear regression problem the entropy function does not converge under the ME formulation. Therefore, we estimate the tobit model using the GME formulation, which includes the error term in the entropy function an its constraints.

Following Golan, Judge, and Perloff (1997), we let y_1 represent the uncensored observations for which $y_t = y_t^*$, and let y_0 represent the censored observations for which $y_t^* \leq 0$ and $y_t = 0$. The tobit model is written in matrix form as

$$\begin{bmatrix} y_1 = y_1^* \\ y_0 = 0 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_0 \end{bmatrix}\beta + \begin{bmatrix} e_1 \\ e_0 \end{bmatrix}, \tag{5.6}$$

where X_1 is an $N_1 \times K$ matrix of explanatory variables for the uncensored observations, X_0 is an $N_0 \times K$ matrix of explanatory variables for the censored observations, e_1 is an $N_1 \times 1$ vector of random errors for the uncensored observations, e_0 is a $N_0 \times 1$ vector of random errors for the censored observations, and β is a $K \times 1$ vector of unknown parameters. As in the linear regression problem, we reparameterize the unknown parameters and errors such that they take the form of probabilities, which can be estimated with GME. That is, we specify a matrix of support points for the unknown parameters and errors and we estimate the unknown probabilities associated with the support points. We rewrite the unknown parameter vector as

$$\beta = Zp = \begin{bmatrix} z_1' & 0 & \cdots & 0 \\ 0 & z_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_K' \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_K \end{bmatrix}, \qquad k = 1,2,\ldots,K \tag{5.7}$$

where β is a $K \times 1$ vector of unknown parameters, Z is a $K \times KM$ matrix of support points, and p is a $KM \times 1$ vector of unknown weights such that $p_{km} > 0$ and $p_k' i_M = 1$ for all k. We rewrite the unknown error vector as

$$e = Vw = \begin{bmatrix} v_1' & 0 & \cdots & 0 \\ 0 & v_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_N' \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}, \qquad n = 1,2,\ldots,N \tag{5.8}$$

where e is an $N \times 1$ vector of random errors, V is an $N \times NJ$ matrix of support points, and w is an $NJ \times 1$ vector of unknown weights such that $w_{ij} > 0$ and $w_i' i_J = 1$ for all i. The reparameterized GME tobit model is written as

$$\begin{bmatrix} y_1 = y_1^* \\ y_0 = 0 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_0 \end{bmatrix} Zp + \begin{bmatrix} V_1 w_1 \\ V_0 w_0 \end{bmatrix}. \tag{5.9}$$

We can take two approaches in solving the censored regression GME model developed by Golan, Judge, and Perloff (1997). We first discuss the solution given by Golan et al. and then present an alternative GME solution for equation (5.9). Golan et al. obtain GME estimates for the censored regression model by solving the constrained optimization problem

$$\max H(p,w) = -p' \ln(p) - w_1' \ln(w_1) - w_0' \ln(w_0) \tag{5.10}$$

subject to

$$y_1 = X_1 Zp + V_1 w_1 \tag{5.11}$$

$$0 \geq X_0 Zp + V_0 w_0 \tag{5.12}$$

$$(I_K \otimes i_M') p = i_K \tag{5.13}$$

$$(I_{N_1} \otimes i_J') w_1 = i_{N_1} \tag{5.14}$$

$$(I_{N_0} \otimes i_J')w_0 = i_{N_0}. \tag{5.15}$$

Equations (5.11) and (5.12) are the data constraints while equations (5.13)-(5.15) are the additivity constraints which require that the probabilities associated with the parameter and error supports sum to one. The Lagrangian for GME tobit problem is

$$\mathcal{L} = -p'\ln(p) - w_1'\ln(w_1) - w_0'\ln(w_0) + \lambda_1'(X_1Zp + V_1w_1 - y_1) + \lambda_0'(-X_0Zp - V_0w_0)$$

$$+\gamma'\left[i_K - (I_K \otimes i_M')p\right] + \delta_1'\left[i_{N_1} - (I_{N_1} \otimes i_J')w_1\right] + \delta_0'\left[i_{N_0} - (I_{N_0} \otimes i_J')w_0\right],$$

where λ_1 is a $N_1 \times 1$ vector of Lagrange multipliers for the data constraint on the uncensored observations, λ_0 is a $N_0 \times 1$ vector of Lagrange multipliers for the data constraint on the censored observations, γ is a $K \times 1$ vector of Lagrange multipliers for the additivity constraint on the unknown probabilities for the parameter support, δ_1 is a $N_1 \times 1$ vector of Lagrange multipliers for the additivity constraints on the error weights for the uncensored observations, and δ_0 is a $N_0 \times 1$ vector of Lagrange multipliers for the additivity constraints on the error weights for the censored observations. The first-order and the Kuhn-Tucker conditions for the GME tobit problem are

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{KM} - \ln(p) + ZX_1'\lambda_1 - ZX_0'\lambda_0 - (I_K \otimes i_M)\gamma \le 0$$

$$p > 0, p \cdot \frac{\partial \mathcal{L}}{\partial p} = 0 \Rightarrow \frac{\partial \mathcal{L}}{\partial p} = 0$$

$$\frac{\partial \mathcal{L}}{\partial w_1} = -i_{N_1J} - \ln(w_1) + V_1'\lambda_1 - (I_{N_1} \otimes i_J)\delta_1 = 0$$

$$\frac{\partial \mathcal{L}}{\partial w_0} = -i_{N_0J} - \ln(w_0) - V_0'\lambda_0 - (I_{N_0} \otimes i_J)\delta_0 \le 0$$

$$w_0 > 0, w_0 \cdot \frac{\partial \mathcal{L}}{\partial w_0} = 0 \Rightarrow \frac{\partial \mathcal{L}}{\partial w_0} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = X_1Zp + V_1w_1 - y_1 = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_0} = -X_0 Zp - V_0 w_0 \geq 0$$

$$\lambda_0 \geq 0, \lambda_0 \cdot \frac{\partial \mathcal{L}}{\partial \lambda_0} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_K - (I_K \otimes i'_M)p = 0$$

$$\frac{\partial \mathcal{L}}{\partial \delta_1} = i_{N_1} - (I_{N_1} \otimes i'_J)w_1 = 0$$

$$\frac{\partial \mathcal{L}}{\partial \delta_0} = i_{N_0} - (I_{N_0} \otimes i'_J)w_0 = 0\,.$$

Note that $\lambda_0 = 0$ unless the latent value of the censored observations $(y_0^* = X_0 Zp + V_0 w_0)$ is equal to zero. Inserting $\lambda_0 = 0$ and solving the KM FOC's for the unknown p's yields

$$\hat{p} = \exp(ZX_1'\lambda_1)\exp\left[-i_{KM} - (I_K \otimes i_M)\gamma\right],$$

which implies that

$$\hat{p}_{km} = \exp(z_{km} x_k' \hat{\lambda}_1)\exp(-1-\gamma_k)\,.$$

Since the additivity constraint requires that $\sum_{m=1}^{M} p_{km} = 1$ and $\exp(-1-\gamma_k)$ is constant for a given parameter, the optimal probabilities may be rewritten as

$$\hat{p}_{km} = \frac{\exp(z_{km} x_k' \hat{\lambda}_1)}{\sum_{m=1}^{M} \exp(z_{km} x_k' \hat{\lambda}_1)} = \frac{\exp(z_{km} x_k' \hat{\lambda}_1)}{\Omega_k(\hat{\lambda}_1)}, \tag{5.16}$$

where $\Omega_k(\hat{\lambda}_1)$ represents the partition function for the k^{th} parameter. Solving the $N_1 J$ FOC's for w_1 yields

$$\hat{w}_1 = \exp(V_1'\lambda_1)\exp\left[-i_{N_1 J} - (I_{N_1} \otimes i_J)\delta_1\right],$$

which implies that

$$\hat{w}_{nj} = \exp(v_{nj} \hat{\lambda}_{1n})\exp(-1-\delta_{1n})$$

for the uncensored observations. Solving the $N_0 J$ FOC's for w_0 yields

$$\hat{w}_0 = \exp(-V_0'\lambda_0)\exp\left[-i_{N_0 J} - (I_{N_0} \otimes i_J)\delta_0\right],$$

which implies that

$$\hat{w}_{nj} = \exp(-1 - \delta_{0n})$$

for the censored observations (since $\lambda_0 = 0$). Since the additivity constraint requires that $\sum_{j=1}^{J} \hat{w}_{nj} = 1$ for all N observations and $\exp(-1 - \delta_n)$ is constant for a given error, we rewrite the optimal error weights as

$$\hat{w}_{nj} = \frac{\exp(v_{nj}\hat{\lambda}_{1n})}{\sum_{j=1}^{J}\exp(v_{nj}\hat{\lambda}_{1n})} = \frac{\exp(v_{nj}\hat{\lambda}_{1n})}{\Psi_n(\hat{\lambda}_{1n})}, \tag{5.17}$$

for the uncensored observations, and

$$\hat{w}_{nj} = \frac{\exp(-1 - \delta_{0n})}{\sum_{j=1}^{J}\exp(-1 - \delta_{0n})} = \frac{1}{J} \tag{5.18}$$

for the censored observations. Thus, if the error support is symmetric about zero, as we have assumed, the estimated errors are equal to zero for the censored observations. Hence, $X_0 Zp = V_0 w_0 = 0$ and the GME tobit solution proposed by Golan, Judge, and Perloff (1997) is equal to the GME linear regression solution, equations (3.10) and (3.19), applied to the N_1 uncensored observations.

An alternative solution is to stack the matrices in equation (5.9) and impose this equality as a single constraint. Thus, we obtain the alternative GME solution by solving the constrained optimization problem

$$\max H(p, w) = -p' \ln(p) - w' \ln(w) \tag{5.19}$$

subject to

$$y = XZp + Vw \tag{5.20}$$

$$(I_K \otimes i_M')p = i_K \tag{5.21}$$

$$(I_N \otimes i_J')w = i_N, \tag{5.22}$$

where $y = (y_1' \quad y_0')'$, $X = (X_1' \quad X_0')'$, and $w = (w_1' \quad w_0')'$. This GME tobit solution is equal to the GME linear regression solution applied to all N observations.

We know that the OLS estimator applied to either the uncensored observations or the entire sample is biased and inconsistent. While GME allows us to impose nonsample information through the parameter and error support matrices, the GME linear regression solution applied to the tobit model suffers from the same problems as OLS. Judge et al. (1988, p. 797) show that for the uncensored observations

$$y_t = x_t'\beta + \sigma \frac{\mathrm{f}_t}{\mathrm{F}_t} + u_t,$$

where f_t represents the standard normal probability density function, F_t represent the standard normal cumulative distribution function, and σ represents the standard deviation of e_t. The GME tobit solution proposed by Golan, Judge, and Perloff (1997) omits the term $\sigma \cdot \mathrm{f}_t / \mathrm{F}_t$, which is a function of x_t. Their GME approach does not include any constraints that take into account the probability of an observation being censored. In addition, their GME formulation assumes independence between p and w so that we can estimate them jointly in the entropy function. However, we know that the errors and parameters are not independent in the tobit model. Therefore, the entropy between the unknown errors and parameters is not additive.

We estimate a censored regression model using a small data sample from Judge et al. (1988, p. 800). The data consist of $N = 20$ observations generated from the regression function

$$y_t^* = \beta_1 + \beta_2 x_{t2} + e_t, \qquad e_t \sim N(0, \sigma^2) \tag{5.23}$$

where $\beta_1 = -9$, $\beta_2 = 1$, and $\sigma = 4$. The sample contains $N_0 = 6$ censored observations and $N_1 = 14$ uncensored observations.

Following Pukelsheim (1994), Golan, Judge, and Perloff (1997) suggest setting the error bounds as $v_{i1} = -3\sigma$ and $v_{i2} = 3\sigma$, where σ is the standard deviation of e. They calculate the

sample variance of y assuming it follows a uniform distribution between its minimum and maximum values, $s_y^2 = (y_{max} - y_{min})^2 / 12$, and use s_y in setting the error bounds. Since we have a censored sample, Golan et al. replace y_{min} with $\hat{y}_{min}$ when calculating s_y^2, where $\hat{y}_{min}$ is an inferred value based on the observed sample and the percentage of censored observations. We use the true latent values, which are known in our sample, to calculate s_y^2. The sample standard deviation of y, assuming is follows a uniform distribution between its maximum and minimum values, is $s_y = 8.5$. Therefore, we use $v_i' = \begin{pmatrix} -25 & 0 & 25 \end{pmatrix}$ as the GME error support using the 3σ-rule and call this model GME3. We also we estimate model GME2, which has an error support $v_i' = \begin{pmatrix} -10 & 0 & 10 \end{pmatrix}$ to examine GME estimation under a narrower error support. In both models we specify the parameter support for both β_1 and β_2 to be $z_i' = \begin{pmatrix} -20 & -10 & 0 & 10 & 20 \end{pmatrix}$.

Table 5.1 gives estimates obtained using OLS on all 20 observations, OLS on the 14 uncensored observations, GME on all 20 observations, GME on the 14 uncensored observations, and tobit. We let OLS-A, GME2-A, and GME3-A denote the estimators applies to all 20 observations and OLS-U, GME2-U, and GME3-U denote the estimators applied to the 14 uncensored observations.

Table 5.1 Parameter Estimates for Model (5.23)

Variable	OLS-A	OLS-U	GME2-A	GME2-U	GME3-A	GME3-U	Tobit
$\hat{\beta}_1$	-2.153	-1.491	-2.128	-1.350	-1.484	-0.652	-5.73
	(-1.47)	(-0.57)					(-2.58)
$\hat{\beta}_2$	0.668	0.653	0.678	0.653	0.620	0.596	0.901
	(5.48)	(3.46)					(5.30)
$\hat{\sigma}^2$	9.887	11.361					13.184
ln l							-41.255

* N =20; 14 uncensored and 6 censored observations; t-statistics in parentheses.

The results show that GME shrinks the parameters towards their prior means, which are zero in this case. The GME estimates are farther from the true parameters than either OLS or tobit. In addition, GME does not provide an estimate of $\hat{\sigma}^2$. Golan, Judge, and Perloff (1997) suggest using the GME

estimates to estimate values for $\hat{y}_0^*$ and re-estimating the model with GME using the true y_1^* and the estimated $\hat{y}_0^*$ as the dependent variables. However, this method does not seem to be a very good one if the original GME parameter estimates are biased and inconsistent. The only benefit of this GME tobit formulation is that it does allow a means of introducing nonsample information. For example, we could have shrunk the parameters toward their true values using GME. Unfortunately, economic theory does not generally provide such information.

5.3 Truncated Regression Models

In this section, we discuss ME in truncated regression models. To date, no one has estimated a truncated regression model using ME. We consider two different truncated regression models involving sample selectivity, the probit model with sample selection and the bivariate probit model. We describe the probit model with sample selection, from Kleit, Pierce, and Hill (1998), in section 5.3.1 and discuss ME estimation in the probit model with sample selection in section 5.3.2. We describe the bivariate probit model and the censored probit model in section 5.3.3.

5.3.1 The Probit Model with Sample Selection

Kleit, Pierce, and Hill (1998) discuss the probit model with sample selection selection using data on penalties administered by the Louisiana Department of Environmental Quality (DEQ). There are three levels of action that can be taken - no action, a mild penalty (compliance order), or a more severe penalty (financial penalty) - depending on whether or not a latent variable, y_i^*, crosses an unobservable threshold. We have complete observations when there is any type of action taken, but we do not have any observations for cases in which no action is taken. Let y_i^* represent a latent variable that is determined by

$$y_i^* = x_i'\beta + e_i, \tag{5.24}$$

where x_i' is a row vector of explanatory variables, β is a $K \times 1$ vector of unknown parameters, and e_i is an $N \times 1$ vector of random errors. In the regular probit model we observe

$$y_i = \begin{cases} 1 & \text{if } y_i^* > \mu_1 \\ 0 & \text{if } y_i^* \le \mu_1. \end{cases} \tag{5.25}$$

Since μ_1 is not identified and cannot be estimated we normalize $\mu_1 = 0$. In the probit model with sample selection we observe

$$y_i = \begin{cases} 1 & \text{if } y_i^* > \mu_1 \\ 0 & \text{if } \mu_1 \ge y_i^* > \mu_0. \end{cases} \tag{5.26}$$

We do not observe anything if $y_i^* < \mu_0$. Again, we normalize $\mu_1 = 0$.

Kleit, Pierce, and Hill (1998) note that ignoring the sample selectivity in this model leads to biased and inconsistent estimators. They show that the probability of a random individual being observed is

$$\Pr\left[y_i^* > \mu_0\right] = 1 - \Phi(\mu_0 - x_i'\beta), \tag{5.27}$$

where $\Phi(z)$ is the standard normal cumulative distribution function evaluated at z. For a given observation the probability that $y_i = 0$ is

$$\Pr\left[y_i = 0\right] = \frac{\Phi(-x_i'\beta) - \Phi(\mu_0 - x_i'\beta)}{1 - \Phi(\mu_0 - x_i'\beta)} = F_{0i},$$

and the probability that we observe $y_i = 1$ is

$$\Pr\left[y_i = 1\right] = \frac{\Phi(x_i'\beta)}{1 - \Phi(\mu_0 - x_i'\beta)} = F_{1i}.$$

The maximum likelihood solution to this problem is obtained by maximizing

$$L = \prod_{i=1}^{T} F_{0i}^{1-y_i} F_{1i}^{y_i} \tag{5.28}$$

with respect to β and μ_0.

5.3.3 ME Estimation in the Probit Model with Sample Selection

Our ME solution to the binary choice model with sample selection assumes that we know the probability that $x_i'\beta > \mu_0$. For example, we may know that a bank denies 10 percent of their loan applications. In Chapter 2, we write the binary choice problem as

$$y_{ij} = G(x_i'\beta_j) + e_{ij} = p_{ij} + e_{ij}. \tag{5.29}$$

In the binary choice model with sample selection we write the model as

$$y_{ij} = G(x_i'\beta_j) + e_{ij} = p_{ij} / \Pr(x_i'\beta_j > \mu_0) + e_{ij}, \tag{5.30}$$

where $\Pr(x_i'\beta_j > \mu_0)$ is the probability that the loan is granted.

Given this information, we estimate the ME probabilities by solving the constrained optimization problem

$$\max H(p) = -p' \ln(p) \tag{5.31}$$

subject to

$$.9(I_2 \otimes X')y = (I_2 \otimes X')p \tag{5.32}$$

$$[I_N \quad I_N]p = .9 \cdot i_N. \tag{5.33}$$

We solve this problem exactly as in Chapter 2. Now, for each individual the probabilities that they will either repay or default on their loan add up to .9, while the probability that the loan is denied is equal to .1. The obvious weakness with this specification is that the probability of the loan being denied is assumed to be constant for each individual. However, without data on individuals denied loans we cannot estimate individual probabilities for having a loan denied.

5.3.4 The Bivariate Probit Model

The bivariate probit model assumes that there are two separate binary equations to determine a sample outcome. We observe the binary outcome in the second equation if and only if the first equation results in a particular outcome. For example, we observe default or repayment of a loan only if the bank has granted the loan. For individuals whose loan applications are denied, we do not observe the binary outcome of default or repayment. There are several papers that deal with the bivariate probit model including Poirier (1980), Zellner and Lee (1965), Farber (1983), Meng and Schmidt (1985), and Terza and Tsai (1996). Morimune (1979) and Schmidt and Strauss (1975) examine a bivariate logit model.

In Chapter 2, we estimate a credit scoring model as a binary choice model in which we observe individuals to default or repay their loans. However, there is sample selection bias in credit scoring problems since we cannot observe repayment or default among individuals who are denied loans. If we

observe the personal and financial characteristics of the individuals who are denied loans we have a censored probit model. If we do not observe the personal and financial characteristics of those denied loans we have a truncated model, known as the bivariate probit model.

In the bivariate probit model, we have two equations

$$y_i^* = x_i'\beta + e_i \tag{5.34}$$

$$d_i^* = z_i'\alpha + v_i \tag{5.35}$$

where y_i^* and d_i^* are unobservable latent variables and $corr(e_i, v_i) = \rho$. We observe

$$d_i = \begin{cases} 1 & \text{if } d_i^* > 0 \\ 0 & \text{if } d_i^* \leq 0, \end{cases}$$

and we observe

$$y_i = \begin{cases} 1 & \text{if } y_i^* > 0 \\ 0 & \text{if } y_i^* \leq 0 \end{cases}$$

if and only if $d_i = 1$. For example, in the credit scoring model we only observe individuals who are granted loans. Thus, in our credit scoring model equation (5.35) represents the loan granting decision and equation (5.34) represents the decision to default or repay the loan.

Poirier (1980) estimates a bivariate probit model in which we only observe the product $y_i \cdot d_i$ rather than each choice individually. Meng and Schmidt (1985) discuss the censored probit model of the type that we have in the credit scoring problem. In the censored probit model, we have complete observations for the loan granting equation, but the repayment equation is only observed for individuals receiving a loan. In this model, the probabilities of the possible outcomes are

$$\Pr[d_i = 1] = q_i = \Phi(z_i'\alpha)$$

$$\Pr[d_i = 0] = 1 - q_i = 1 - \Phi(z_i'\alpha)$$

$$\Pr[y_i = 1, d_i = 1] = p_i q_i = F(x_i'\beta, z_i'\alpha; \rho)$$

$$\Pr[y_i = 0, d_i = 1] = (1 - p_i)q_i = \Phi(z_i'\alpha) - F(x_i'\beta, z_i'\alpha; \rho),$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function and $F(\cdot,\cdot;\rho)$ denotes the bivariate standard normal cumulative distribution function with correlation coefficient ρ. The joint probability density function of the observed y_t's and d_t's is

$$f(y_t,d_t)=p_t q_t^{y_t d_t}(1-p_t)q_t^{(1-y_t)d_t}(1-q_t)^{1-d_t}. \tag{5.36}$$

The likelihood function, $L=\underset{t=1}{\overset{N}{\pi}} f(y_t,d_t)$, is given by

$$L=F(x_t'\beta,z_t'\alpha;\rho)^{y_t d_t}[\Phi(z_t'\alpha)-F(x_t'B,z_t'\alpha;\rho)]^{(1-y_t)d_t}[1-\Phi(z_t'\alpha)]^{1-d_t} \tag{5.37}$$

The log-likelihood function is

$$\ln L=\sum_{t=1}^{N}\{y_t d_t \ln F(x_t'\beta,z_t'\alpha;\rho)+(1-y_t)d_t\ln[\Phi(z_t'\alpha)-F(x_t'\beta,z_t'\alpha;\rho)]$$

$$+(1-d_t)\ln[1-\Phi(z_t'\alpha)]\}$$

$$=\sum_{d_t=1}\{y_t\ln F(x_t'\beta,z_t'\alpha;\rho)+(1-y_t)\ln[\Phi(z_t'\alpha)-F(x_t'\beta,z_t'\alpha;\rho)]\}$$

$$+\sum_{t=1}^{N}(1-d_t)\ln[1-\Phi(z_t'\alpha)]. \tag{5.38}$$

Maximization of (5.38) yields the maximum likelihood estimates for the censored probit model. Manski and Lerman (1977) and Hoffman, Boyes, and Low (1989) estimate credit scoring models using a censored probit model with data on individuals who are denied loans.

5.4 Conclusions

To date, there has been very little research about ME and GME estimation in censored and truncated regression models. Golan, Judge, and Perloff (1997) and Golan, Judge, and Miller (1996) discuss GME estimation in the tobit model. However, their solution does not account for the probability of an observation being censored. We conclude that the current GME approaches do not provide a good alternative to traditional estimation techniques in censored and truncated regression models. However, given the scarcity of research about GME estimation in econometric models, it is likely that additional research will yield a viable GME estimator for censored and truncated regression models.

5.5 References

Amemiya, T. (1973b). Regression analysis when the dependent variable is truncated normal, *Econometrica* 41: 997-1016.

Farber, H. S. (1983). The determination of the union status of workers, *Econometrica* 51 (5): 1417-1437.

Golan, A., G. Judge, D. Miller. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* New York: John Wiley and Sons.

Golan, A., G. Judge, J. Perloff. (1997). Estimation and inference with censored and ordered multinomial response data, *Journal of Econometrics* 79: 23-51.

Boyes, W. J., D. L. Hoffman, S. A. Low. (1989). An econometric analysis of the bank credit scoring Problem. *Journal of Econometrics* 40: 3-14.

Judge, G. G., R.C. Hill, W. E. Griffith, H. Lutkepohl, T. Lee. (1988). Introduction to the theory and practice of econometrics 2nd ed. New York: John Wiley and Sons.

Kleit, A.N., M.A. Pierce, R.C. Hill. (1998). Environmental protection, agency motivations, and rent extraction: The regulation of water pollution in Louisiana, *Journal of Regulatory Economics* 13: 121-137.

Manski, C. F., S. R. Lerman. (1977). The estimation of choice probabilities from choice based samples, *Econometrica* 45 (8): 1977-1988.

Meng, C. L., P. Schmidt. (1985). On the cost of partial observability in the bivariate probit model, *International Economic Review* 26 (1): 71-85.

Morimune, K. (1979). Comparisons of normal and logistic models in the bivariate dichotomous analysis, *Econometrica* 47 (4): 957-975.

Pukelsheim. F. (1994). The three sigma rule, *American Statistician* 48: 88-91.

Poirier, D. J. (1980). Partial observability in bivariate probit models, *Journal of Econometrics* 12: 209-217.

Schmidt, P., R. P. Strauss. (1975). Estimation of models with jointly dependent qualitative variables: A simultaneous logit approach, *Econometrica* 43 (4): 745-755.

Terza, J. V., W. D. Tsai. (1996). An alternative specification for the censored probit model: Dealing with an uncooperative correlation coefficient. Working Paper Penn State.

Tobin, J. (1958). Estimation of relationships for limited dependent variables, *Econometrica* 26: 24-36.

Zellner, A., T. H. Lee. (1965). Joint estimation of relationships involving discrete random variables, *Econometrica* 33 (2): 382-394.

CHAPTER 6

CONCLUSIONS

6.1 Maximum Entropy Estimation

Maximum entropy (ME) estimation is a relatively new estimation technique in econometrics. Jaynes (1957a,b) argues that when estimating an unknown probability distribution we should choose the distribution that maximizes entropy and is compatible with our prior information. Denzau, Gibbons, and Greenberg (1989) and Soofi (1992) discuss ME estimation of the unknown probabilities in multinomial choice models. Estimates for the unknown parameters are obtained as the Lagrange multipliers in the entropy maximization problem. Golan, Judge, and Perloff (1996) extend this work by defining a generalized maximum entropy (GME) estimator that jointly estimates the unknown probabilities and errors. Golan, Judge, and Miller (1996) apply the GME estimator to estimate a wide range of econometric problems by reparameterizing the unknown parameters and errors such that they take the form of probabilities. We examine the ME and GME estimators in the context of several statistical models using real data as a basis for Monte Carlo experiments, as opposed to previous studies that used artificial data.

6.2 ME and GME Estimation in Discrete Choice Models

We examine both ME and GME estimation in discrete choice models using actual data as a basis for Monte Carlo sampling experiments. In discrete choice models, we have a known set of discrete support points for the dependent variable, y, and we estimate the probability that y is equal to each of these points. ME estimation yields the most uniform probability distribution that is compatible with the observed data. We find that the ME estimator is equivalent to the logit estimator when we specify the ME data constraint suggested by Soofi (1992) and Golan, Judge, and Perloff (1996).

The GME estimator includes the unknown errors in the entropy function and its constraints. GME allows us to estimate both the unknown probabilities and the unknown errors associated with these

probabilities. To obtain GME error estimates, we must convert the unknown errors to probabilities, which we do by specifying a set of support points for the errors *a priori*. We show that in discrete choice models GME shrinks the predicted probabilities for each alternative toward $1/J$, where J is the number of alternatives. The wider we specify the error bounds the greater the degree of shrinkage. By placing more weight on the errors we shrink the parameter estimates toward zero and the predicted probabilities toward $1/J$. In addition, we find that shrinkage is greater when we have an unbalanced sample. For balanced samples, we observe very little difference in prediction between the alternative estimators.

We find that GME performs well in Monte Carlo sampling experiments in terms of risk measures for the parameter estimates, the predicted latent variables, the predicted probabilities, and the marginal effects. We specify three GME estimators with different error support vectors; GME1 has wide error bounds of $[-1,1]$, GME2 has error bounds of $[-.1,.1]$, and GME3 has errors bounds of $[-.01,.01]$. The GME2 and GME3 estimators dominate ME-logit in terms of empirical risk measures in Monte Carlo sampling experiments. The GME1 estimator has the smallest variance of the alternative estimators since the parameter estimates are shrunk toward zero. However, the GME1 estimator does not predict well in a hold-out sample and has a higher MSE for the margnal effects than the alternative estimators. On the basis of our Monte Carlo experiments we conclude that GME estimation is a good alternative to logit estimation, particularly if we have an unbalanced sample where the shrinkage properties of GME are greater.

6.3 GME Estimation in a Linear Regression Model

We examine GME estimation in linear regression models using actual data as a basis for Monte Carlo sampling experiments both with and without inequality restrictions placed on the parameters. In addition, we develop a method to specify inequality restrictions in GME through the parameter support matrix. However, the restricted GME parameter estimates are sensitive to both the restrictions and the prior mean specified in the parameter support.

The GME estimator in the linear regression problems works similar to Stein-like rules since it allows us to introduce non-sample information. In the linear regression problem we must specify a support matrix based on prior information and we estimate the discrete probability distribution for each β_k associated with its support vector. We show the GME shrinks the parameter estimates toward the prior means of their support vectors and that the GME parameter estimates are very sensitive to the parameter and error support matrices that we specify *a priori*. While GME allows us a means of specifying nonsample information about the parameters the fact that we must bound each parameter is a problem since economic theory does not usually provide such information.

We find that the GME estimator has lower risk than the OLS estimator in terms of the MSE of the parameter estimates and the out-of-sample prediction error when the underlying errors have a standard normal or standardized chi-square distribution. GME does not perform well when the errors are drawn from a standardized *t*- distribution. Our Monte Carlo experiments show that GME is superior to OLS if we have good nonsample information about the parameters. However, even when we have good nonsample information the GME estimator is very sensitive to the support points that the researcher must specify for the unknown parameters and errors.

6.4 Future Research Plans

In Chapter 5, we discuss ME estimation in censored and truncated regression models. This is an area of research that has not yet been explored in great detail. Golan, Judge, and Perloff (1997) propose a GME solution to the tobit model, but it is not a completely satisfactory solution. There have been no entropy solutions proposed for truncated regression models. This is one area that we will continue to research.

Another area of future research is to impose nonsample information about the parameters in discrete choice models. The ME estimator for discrete models developed by Soofi (1992) and Golan, Judge, and Perloff (1996) estimates the unknown probabilities of choosing each alternative and only yields parameter estimates indirectly as Lagrange multipliers for the data constraint. Thus, there is no means for imposing nonsample information about the parameters in a discrete model as there is in the linear regression model. Finally, we will further explore the shrinkage properties of GME estimation in

the linear regression model and how to best specify nonsample information. One possibility that has not been discussed in the literature is to place priors on values such as elacticities rather than the parameters themselves.

ME and GME estimation perform very well in Monte Carlo sampling experiments using real data which shows that they have great potential for econometric applications. However, entropy methods are not fully understood at this time and there is a lot of potential for future research on how best to apply entropy estimation in econometric modeling.

6.5 References

Denzau, A. T., P. C. Gibbons, E. Greenberg. (1989). Bayesian estimation of proportions with a cross-entropy prior, *Communications in Statistics-Theory and Methods* 18: 1843-1861.

Golan, A., G. Judge, D. Miller. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* New York: John Wiley and Sons.

Golan, A., G. Judge, J. M. Perloff. (1996). A maximum entropy approach to recovering information from multinomial response data, *Journal of the American Statistical Association* 91 (June): 841-853.

________. (1997). Estimation and inference with censored and ordered multinomial response data, *Journal of Econometrics* 79: 23-51.

Jaynes, E. T. (1957a). Information theory and statistical mechanics, *Physical Review* 106: 620-630.

________. (1957b). Information theory and statistical mechanics II, *Physical Review* 108: 171-190.

Soofi, E. S. (1992). A generalizable formulation of conditional logit with diagnostics, *Journal of the American Statistical Association* 87(Sept.): 812-816.

APPENDIX A

SOLUTIONS TO MAXIMUM ENTROPY PROBLEMS

A.1 General Entropy Problem

The Objective Function:

$$\max H(p) = -\sum_{i=1}^{N} p_i \ln(p_i) \tag{A.1}$$

subject to

$$\sum_{i=1}^{N} p_i = 1. \tag{A.2}$$

The Lagrangian:

$$\mathcal{L} = -\sum_{i=1}^{N} p_i \ln(p_i) + \lambda(1 - \sum_{i=1}^{N} p_i).$$

First-Order Conditions:

$$\frac{\partial \mathcal{L}}{\partial p_i} = -1 - \ln(p_i) - \lambda = 0 \qquad \forall i$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 1 - \sum_{i=1}^{N} p_i = 0.$$

Solve FOC's for p:

$$\ln(p_i) = -1 - \lambda$$

$$\Rightarrow \hat{p}_i = \exp(-1 - \hat{\lambda}) \qquad \forall i.$$

Because the probabilities must sum to one, we can rewrite the optimal ME probabilities as

$$\hat{p}_i = \frac{\exp(-1 - \hat{\lambda})}{\sum_{i=1}^{N} \exp(-1 - \hat{\lambda})} = \frac{1}{N}. \tag{A.3}$$

Second-Order Conditions:

$$\frac{\partial^2 \mathcal{L}}{\partial p_i^2} = \begin{bmatrix} -p_1^{-1} & 0 & \cdots & 0 \\ 0 & -p_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -p_N^{-1} \end{bmatrix}.$$

Because p_i is positive for all i, all of the diagonal elements of the Hessian are negative and the matrix of second derivatives is negative definite, thus ensuring a unique global maximum.

A.2 Jaynes' Dice Problem – Primal Solution

The Objective Function:

$$\max H(p) = -\sum_{i=1}^{6} p_i \ln(p_i) \tag{A.4}$$

subject to

$$\sum_{i=1}^{6} p_i x_i = y \tag{A.5}$$

$$\sum_{i=1}^{6} p_i = 1. \tag{A.6}$$

The Lagrangian:

$$\mathcal{L} = -\sum_{i=1}^{6} p_i \ln(p_i) + \lambda(y - \sum_{i=1}^{6} p_i x_i) + \gamma(1 - \sum_{i=1}^{6} p_i).$$

First-Order Conditions:

$$\frac{\partial \mathcal{L}}{\partial p_i} = -1 - \ln(p_i) - x_i\lambda - \gamma = 0 \qquad \forall i$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = y - \sum_{i=1}^{6} p_i x_i = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = 1 - \sum_{i=1}^{6} p_i = 0.$$

Solve FOC's for p:

$$\ln(p_i) = -1 - x_i\lambda - \gamma$$

$$\Rightarrow \hat{p}_i = \exp(-1 - x_i\hat{\lambda} - \hat{\gamma}) \qquad \forall i.$$

Because the additivity constraint requires that the probabilities must sum to one, we can rewrite the optimal ME probabilities as

$$\hat{p}_i = \frac{\exp(-1-\hat{\gamma})\exp(-x_i\hat{\lambda})}{\exp(-1-\hat{\gamma})\sum_{j=1}^{6}\exp(-x_i\hat{\lambda})} = \frac{\exp(-x_i\hat{\lambda})}{\sum_{j=1}^{6}\exp(-x_j\hat{\lambda})} = \frac{\exp(-x_i\hat{\lambda})}{\Omega(\hat{\lambda})}. \tag{A.7}$$

Second-Order Conditions:

$$\frac{\partial^2 \mathcal{L}}{\partial p_i^{\,2}} = \begin{bmatrix} -p_1^{-1} & 0 & \cdots & 0 \\ 0 & -p_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -p_N^{-1} \end{bmatrix}.$$

Because p_i is positive for all i, all of the diagonal elements of the Hessian are negative and the matrix of second derivatives is negative definite, thus ensuring a unique global maximum.

A.2.1 Jaynes' Dice Problem - Dual Solution

We substitute the optimal ME probabilities into the original Lagrangian to form the concentrated Lagrangian:

$$M(\lambda) = -\sum_{i=1}^{6}\hat{p}_i(\lambda)\ln(\hat{p}_i(\lambda)) + \lambda(y - \sum_{i=1}^{6}\hat{p}_i(\lambda)x_i)$$

$$= -\sum_{i=1}^{6}\hat{p}_i(\lambda)\ln\left[\frac{\exp(-x_i\lambda)}{\Omega(\lambda)}\right] + \lambda y - \lambda\sum_{i=1}^{6}\hat{p}_i(\lambda)x_i$$

$$= -\sum_{i=1}^{6}\hat{p}_i(\lambda)\left[-x_i\lambda - \ln(\Omega(\lambda))\right] + \lambda y - \lambda\sum_{i=1}^{6}\hat{p}_i(\lambda)x_i$$

$$= -\lambda\sum_{i=1}^{6}\hat{p}_i(\lambda)x_i + \ln(\Omega(\lambda)) + \lambda y - \lambda\sum_{i=1}^{6}\hat{p}_i(\lambda)x_i$$

$$= \lambda y + \ln(\Omega(\lambda))$$

$$= \lambda y + \ln\left[\sum_{j=1}^{6}\exp(-x_j\lambda)\right].$$

First-Order Condition:

$$\frac{\partial \mathrm{M}}{\partial \lambda}=y+\sum_{j=1}^{6}\frac{-x_j \exp(-x_j\lambda)}{\Omega(\lambda)}=0.$$

Second-Order Condition:

$$\frac{\partial^2 \mathrm{M}}{\partial \lambda^2}=\sum_{j=1}^{6}\frac{x_j^2 \exp(-x_j\lambda)\Omega(\lambda)+x_j \exp(-x_j\lambda)\sum_{j=1}^{6}-x_j \exp(-x_j\lambda)}{[\Omega(\lambda)]^2}$$

$$=\sum_{j=1}^{6}\frac{x_j \exp(-x_j\lambda)\left[x_j\sum_{j=1}^{6}\exp(-x_j\lambda)-\sum_{j=1}^{6}x_j \exp(-x_j\lambda)\right]}{[\Omega(\lambda)]^2},$$

which is positive for all λ. Therefore, the concentrated Lagrangian is a strictly convex function of λ.

We minimize the concentrated Lagrangian to obtain $\hat{\lambda}$, which yields the optimal probabilities.

A.3 Binary Choice Problem – ME Primal Solution

The Objective Function:

$$\max H(p)=-p'\ln(p) \tag{A.8}$$

subject to

$$(I_2\otimes X')y=(I_2\otimes X')p \tag{A.9}$$

$$\begin{bmatrix}I_N & I_N\end{bmatrix}p=i_N\,. \tag{A.10}$$

The Lagrangian:

$$\mathcal{L}=-p'\ln(p)+\lambda'[(I_2\otimes X')p-(I_2\otimes X')y]+\gamma'\left[i_N-\begin{bmatrix}I_N & I_N\end{bmatrix}p\right],$$

where $\lambda=\begin{bmatrix}\lambda_1' & \lambda_2'\end{bmatrix}'$ is a $2K\times 1$ vector of parameter estimates.

First-Order Conditions:

$$\frac{\partial \mathcal{L}}{\partial p}=-i_{2N}-\ln(p)+(I_2\otimes X)\lambda-\begin{bmatrix}I_N & I_N\end{bmatrix}'\gamma=0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda}=(I_2\otimes X')p-(I_2\otimes X')y=0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_N - [I_N \quad I_N]p = 0 .$$

Solve FOC's for p:

$$\ln(p) = (I_2 \otimes X)\hat{\lambda} - i_{2N} - (I_N \; I_N)'\hat{\gamma}$$

$$\hat{p} = \exp\left[(I_2 \otimes X)\hat{\lambda}\right]\exp\left[-i_{2N} - (I_N \; I_N)'\hat{\gamma}\right]$$

$$\Rightarrow \hat{p}_{ij} = \exp(x_i'\hat{\lambda}_j)\exp(-1 - \hat{\gamma}_i) .$$

Since additivity requires that $p_{i1} + p_{i2} = 1$ and $\exp(-\hat{\gamma}_i - 1)$ is constant for a given individual, the optimal ME probabilities can be rewritten as

$$\hat{p}_{ij} = \frac{\exp(-1 - \hat{\gamma}_i)\exp(x_i'\hat{\lambda}_j)}{\exp(-1 - \hat{\gamma}_i)\sum_{j=1}^{2}\exp(x_i'\hat{\lambda}_j)} = \frac{\exp(x_i'\hat{\lambda}_j)}{\sum_{j=1}^{2}\exp(x_i'\hat{\lambda}_j)} .$$

Replacing λ with β and normalizing $\hat{\beta}_2 = 0$ yields

$$\hat{p}_{ij} = \frac{\exp(x_i'\hat{\beta}_j)}{1 + \exp(x_i'\hat{\beta}_1)} = \frac{\exp(x_i'\hat{\beta}_j)}{\Omega_i(\hat{\beta})}, \qquad j = 1,2 . \tag{A.11}$$

Second-Order Conditions:

$$\frac{\partial^2 \mathcal{L}}{\partial p \partial p'} = -\frac{1}{p}$$

on the diagonal elements and zero on the off-diagonals. Therefore, the Hessian is negative definite and we obtain a unique global maximum.

A.3.1 Binary Choice Problem – ME Dual Solution

We substitute the optimal ME probabilities into the original Lagrangian to form the concentrated Lagrangian:

$$M(\lambda) = -\hat{p}'(\lambda)\ln(\hat{p}(\lambda)) + \lambda'\left[(I_2 \otimes X')\hat{p}(\lambda) - (I_2 \otimes X')y\right]$$

$$= -\hat{p}'(\lambda)\ln\left[\frac{\exp(x_i'\lambda_j)}{\Omega_i(\lambda)}\right] + \lambda'\left[(I_2 \otimes X')\hat{p}(\lambda) - (I_2 \otimes X')y\right]$$

$$= -\hat{p}'(\lambda)\left[x_i'\lambda_j - \ln(\Omega_i(\lambda))\right] + \hat{p}'(\lambda)(I_2 \otimes X)\lambda - y'(I_2 \otimes X)\lambda$$

$$= -\hat{p}'(\lambda)(I_2 \otimes X)\lambda + \sum_i \ln\left[\Omega_i(\lambda)\right] + \hat{p}'(\lambda)(I_2 \otimes X)\lambda - y'(I_2 \otimes X)\lambda$$

$$= -y'(I_2 \otimes X)\lambda + \sum_i \ln\left[\sum_j \exp(x_i'\lambda_j)\right].$$

Normalizing $\lambda_2 = 0$ yields

$$M(\lambda) = -y'X\lambda + \sum_i \ln\left[1 + \exp(x_i'\lambda)\right].$$

First-Order Condition:

$$\frac{\partial \mathrm{M}}{\partial \lambda} = -X'y + \sum_i \frac{x_i \exp(x_i'\lambda)}{\Omega_i(\lambda)} = 0.$$

Second-Order Condition:

$$\frac{\partial^2 \mathrm{M}}{\partial \lambda \partial \lambda'} = \sum_i \frac{x_i x_i' \exp(x_i'\lambda)\left[1 + \exp(x_i'\lambda)\right] - x_i x_i'\left[\exp(x_i'\lambda)\right]^2}{\left[\Omega_i(\lambda)\right]^2}$$

$$= \sum_i \frac{x_i x_i' \exp(x_i'\lambda) + x_i x_i'\left[\exp(x_i'\lambda)\right]^2 - x_i x_i'\left[\exp(x_i'\lambda)\right]^2}{\left[\Omega_i(\lambda)\right]^2}$$

$$= \sum_i \frac{x_i x_i' \exp(x_i'\lambda)}{\left[\Omega_i(\lambda)\right]^2},$$

which is positive for all λ. Therefore, the concentrated Lagrangian is a strictly convex function of λ.

We minimize the concentrated Lagrangian to obtain $\hat{\lambda}$, which yields the optimal probabilities.

A.4 Binary Choice Problem – GME Primal Solution

The Objective Function:

$$H(p,w) = -p'\ln(p) - w'\ln(w) \tag{A.12}$$

subject to

$$(I_2 \otimes X')y = (I_2 \otimes X')p + (I_2 \otimes X')Vw \tag{A.13}$$

$$\left[I_N \quad I_N\right]p = i_N \tag{A.14}$$

$$(I_{2N} \otimes i'_M)w = i_{2N}. \tag{A.15}$$

The Lagrangian:

$$\mathcal{L} = -p'\ln(p) - w'\ln(w) + \lambda'[(I_2 \otimes X')p + (I_2 \otimes X')Vw - (I_2 \otimes X')y]$$

$$+\gamma'[i_N - [I_N \quad I_N]p] + \tau'[i_{2N} - (I_{2N} \otimes i'_M)w],$$

where $\lambda = [\lambda'_1 \quad \lambda'_2]'$ is a $2K \times 1$ vector of parameter estimates.

First-Order Conditions:

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{2N} - \ln(p) + (I_2 \otimes X)\lambda - [I_N \quad I_N]'\gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = (I_2 \otimes X')p + (I_2 \otimes X')Vw - (I_2 \otimes X')y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_N - [I_N \quad I_N]p = 0$$

$$\frac{\partial \mathcal{L}}{\partial w} = -i_{2NM} - \ln(w) + V'(I_2 \otimes X)\lambda - (I_{2N} \otimes i_M)\tau = 0$$

$$\frac{\partial \mathcal{L}}{\partial \tau} = i_{2N} - (I_{2N} \otimes i'_M)w = 0.$$

Solve FOC's for p:

$$\ln(p) = (I_2 \otimes X)\hat{\lambda} - i_{2N} - [I_N \quad I_N]'\hat{\gamma}$$

$$\hat{p} = \exp\left[(I_2 \otimes X)\hat{\lambda}\right]\exp\left[-i_{2N} - [I_N \quad I_N]'\hat{\gamma}\right]$$

$$\Rightarrow \hat{p}_{ij} = \exp(x'_i\hat{\lambda}_j)\exp(-1 - \hat{\gamma}_i).$$

Since additivity requires that $p_{i1} + p_{i2} = 1$ and $\exp(-\hat{\gamma}_i - 1)$ is constant for a given individual, the optimal ME probabilities can be rewritten as

$$\hat{p}_{ij} = \frac{\exp(-1 - \hat{\gamma}_i)\exp(x'_i\hat{\lambda}_j)}{\exp(-1 - \hat{\gamma}_i)\sum_{j=1}^{2}\exp(x'_i\hat{\lambda}_j)} = \frac{\exp(x'_i\hat{\lambda}_j)}{\sum_{j=1}^{2}\exp(x'_i\hat{\lambda}_j)}.$$

Replacing λ with β and normalizing $\hat{\beta}_2 = 0$ yields

$$\hat{p}_{ij} = \frac{\exp(x_i'\hat{\beta}_j)}{1+\exp(x_i'\hat{\beta}_1)} = \frac{\exp(x_i'\hat{\beta}_j)}{\Omega_i(\hat{\beta})}, \qquad j = 1,2. \tag{A.16}$$

Solve FOC's for w:

$$\ln(w) = V'(I_2 \otimes X)\hat{\lambda} - i_{2NM} - (I_{2N} \otimes i_M)\hat{\tau}$$

$$w = \exp\left[V'(I_2 \otimes X)\hat{\lambda}\right]\exp\left[-i_{2NM} - (I_{2N} \otimes i_M)\hat{\tau}\right]$$

$$\Rightarrow w_{ijm} = \exp(x_i'\hat{\lambda}_j v_{ijm})\exp\left[-1 - \hat{\tau}_{ij}\right].$$

Because additivity requires that $\sum_{m=1}^{M} w_{ijm} = 1$, and $\exp(-\tau_{ij} - 1)$ is constant for a given i and j as we vary m, the optimal GME error weights may be rewritten as

$$w_{ijm} = \frac{\exp(-1-\hat{\tau}_{ij})\exp(x_i'\hat{\lambda}_j v_{ijm})}{\exp(-1-\hat{\tau}_{ij})\sum_{m=1}^{M}\exp(x_i'\hat{\lambda}_j v_{ijm})} = \frac{\exp(x_i'\hat{\lambda}_j v_{ijm})}{\sum_{m=1}^{M}\exp(x_i'\hat{\lambda}_j v_{ijm})}$$

We again replace λ with β and normalize $\hat{\beta}_2 = 0$ to obtain the optimal GME error weights

$$\hat{w}_{ijm} = \frac{\exp(x_i'\hat{\beta}_j v_{ijm})}{\sum_{m=1}^{M}\exp(x_i'\hat{\beta}_j v_{ijm})} = \frac{\exp(x_i'\hat{\beta}_j v_{ijm})}{\Psi_{ij}(\hat{\beta}_j)}. \tag{A.17}$$

Second-Order Conditions:

$$\frac{\partial^2 \mathcal{L}}{\partial p \partial p'} = -\frac{1}{p}$$

on the diagonal elements and zero on the off-diagonals;

$$\frac{\partial^2 \mathcal{L}}{\partial w \partial w'} = -\frac{1}{w}$$

on the diagonal elements and zero on the off-diagonals. Therefore, the Hessian is negative definite and we obtain a unique global maximum.

A.4.1 Binary Choice Problem – GME Dual Solution

We substitute the optimal GME probabilities into the original Lagrangian to form the concentrated Lagrangian:

$$M(\lambda) = -\hat{p}'(\lambda)\ln(\hat{p}'(\lambda)) - \hat{w}'(\lambda)\ln(\hat{w}'(\lambda))$$

$$+\lambda'\left[(I_2 \otimes X')\hat{p}(\lambda) + (I_2 \otimes X')V\hat{w}(\lambda) - (I_2 \otimes X')y\right]$$

$$= -\hat{p}'(\lambda)\ln\left[\frac{\exp(x_i'\lambda_j)}{\Omega_i(\lambda)}\right] - \hat{w}'(\lambda)\ln\left[\frac{\exp(x_i'\lambda_j v_{ijm})}{\Psi_{ij}(\lambda_j)}\right]$$

$$+\lambda'\left[(I_2 \otimes X')\hat{p}(\lambda) + (I_2 \otimes X')V\hat{w}(\lambda) - (I_2 \otimes X')y\right]$$

$$= -\hat{p}'(\lambda)\left[x_i'\lambda_j - \ln(\Omega_i(\lambda))\right] - \hat{w}'(\lambda)\left[x_i'\lambda_j v_{ijm} - \ln(\Psi_{ij}(\lambda_j))\right]$$

$$+\hat{p}'(\lambda)(I_2 \otimes X)\lambda + \hat{w}'(\lambda)V'(I_2 \otimes X)\lambda - y'(I_2 \otimes X)\lambda$$

$$= -\hat{p}'(\lambda)(I_2 \otimes X)\lambda + \sum_i \ln\left[\Omega_i(\lambda)\right] - \hat{w}'(\lambda)V'(I_2 \otimes X)\lambda$$

$$+\sum_i\sum_j \ln\left[\Psi_{ij}(\lambda_j)\right] + \hat{p}'(\lambda)(I_2 \otimes X)\lambda + \hat{w}'(\lambda)V'(I_2 \otimes X)\lambda$$

$$-y'(I_2 \otimes X)\lambda$$

$$= -y'(I_2 \otimes X)\lambda + \sum_i \ln\left[\sum_j \exp(x_i'\lambda_j)\right] + \sum_i\sum_j \ln\left[\sum_m \exp(x_i'\lambda_j v_{ijm})\right].$$

Normalizing $\lambda_2 = 0$ yields

$$M(\lambda) = -y'X\lambda + \sum_i \ln\left[1 + \exp(x_i'\lambda)\right] + \sum_i\sum_j \ln\left[\sum_m \exp(x_i'\lambda v_{ijm})\right]$$

$$= -y'X\lambda + \sum_i \ln\left[1 + \exp(x_i'\lambda)\right] + \sum_i \ln\left[\sum_m \exp(x_i'\lambda v_{ijm})\right] + \sum_i \ln\left[\sum_m \exp(0)\right].$$

First-Order Condition:

$$\frac{\partial \mathrm{M}}{\partial \lambda} = -X'y + \sum_i \frac{x_i \exp(x_i'\lambda)}{\Omega_i(\lambda)} + \sum_i\left[\sum_m \frac{x_i v_{ijm} \exp(x_i'\lambda v_{ijm})}{\Psi_{ij}(\lambda)}\right]$$

Second-Order Condition:

$$\frac{\partial^2 \mathrm{M}}{\partial\lambda\partial\lambda'}=\sum_i \frac{x_i x_i' \exp(x_i'\lambda)[1+\exp(x_i'\lambda)]-x_i x_i'[\exp(x_i'\lambda)]^2}{[\Omega_i(\lambda)]^2}+$$

$$\sum_i\left[\sum_m \frac{x_i x_i' v_{ijm}^2 \exp(x_i'\lambda v_{ijm})\Psi_{ij}-x_i x_i' v_{ijm}\exp(x_i'\lambda v_{ijm})\sum_m v_{ijm}\exp(x_i'\lambda v_{ijm})}{\Psi_{ij}^2}\right]$$

$$=\sum_i \frac{x_i x_i' \exp(x_i'\lambda)}{[\Omega_i(\lambda)]^2}+\sum_i\left\{\sum_m \frac{x_i x_i' v_{ijm}\exp(x_i'\lambda v_{ijm})\left[v_{ijm}\Psi_{ij}-\sum_m v_{ijm}\exp(x_i'\lambda v_{ijm})\right]}{\Psi_{ij}^2}\right\},$$

which is positive for all λ. Therefore, the concentrated Lagrangian is a strictly convex function of λ.

We minimize the concentrated Lagrangian to obtain $\hat{\lambda}$, which yields the optimal probabilities.

A.5 Multinomial Choice Problem – ME Primal Solution

The Objective Function:

$$\max H(p)=-p'\ln(p) \tag{A.18}$$

subject to

$$(I_J\otimes X')y=(I_J\otimes X')p \tag{A.19}$$

$$\begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J}\end{bmatrix}p=i_N\,. \tag{A.20}$$

The Lagrangian:

$$\mathcal{L}=-p'\ln(p)+\lambda'[(I_J\otimes X')p-(I_J\otimes X')y]+\gamma'\left[i_N-\begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J}\end{bmatrix}p\right]$$

where $\lambda=\begin{bmatrix}\lambda_1' & \lambda_2' & \cdots & \lambda_J'\end{bmatrix}'$ is a $KJ\times 1$ vector of parameter estimates.

First-Order Conditions:

$$\frac{\partial\mathcal{L}}{\partial p}=-i_{NJ}-\ln(p)+(I_J\otimes X)\lambda-\begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J}\end{bmatrix}'\gamma=0$$

$$\frac{\partial\mathcal{L}}{\partial\lambda}=(I_J\otimes X')p-(I_J\otimes X')y=0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_N - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix} p = 0.$$

Solve FOC's for p:

$$\ln(p) = (I_J \otimes X)\lambda - i_{NJ} - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix}' \gamma$$

$$\hat{p} = \exp\left[(I_J \otimes X)\hat{\lambda}\right] \exp\left[-i_{NJ} - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix}' \hat{\gamma}\right]$$

$$\Rightarrow \hat{p}_{ij} = \exp(x_i'\hat{\lambda}_j)\exp(-1-\hat{\gamma}_i).$$

Since additivity requires that $p_{i1} + p_{i2} + \ldots + p_{iJ} = 1$ and $\exp(-\hat{\gamma}_i - 1)$ is constant for a given individual, the optimal ME probabilities can be rewritten as

$$\hat{p}_{ij} = \frac{\exp(-1-\hat{\gamma}_i)\exp(x_i'\hat{\lambda}_j)}{\exp(-1-\hat{\gamma}_i)\sum_{j=1}^{J}\exp(x_i'\hat{\lambda}_j)} = \frac{\exp(x_i'\hat{\lambda}_j)}{\sum_{j=1}^{J}\exp(x_i'\hat{\lambda}_j)}.$$

Replacing λ with β and normalizing $\hat{\beta}_J = 0$ yields

$$\hat{p}_{ij} = \frac{\exp(x_i'\hat{\beta}_j)}{1+\sum_{j=1}^{J-1}\exp(x_i'\hat{\beta}_j)} = \frac{\exp(x_i'\hat{\beta}_j)}{\Omega_i(\hat{\beta})}. \qquad \text{(A.21)}$$

Second-Order Conditions:

$$\frac{\partial^2 \mathcal{L}}{\partial p \partial p'} = -\frac{1}{p}$$

on the diagonal elements and zero on the off-diagonals. Therefore, the Hessian is negative definite and we obtain a unique global maximum.

A.5.1 Multinomial Choice Problem – ME Dual Solution

We substitute the optimal ME probabilities into the original Lagrangian to form the concentrated Lagrangian:

$$M(\lambda) = -\hat{p}'(\lambda)\ln(\hat{p}(\lambda)) + \lambda'\left[(I_J \otimes X')\hat{p}(\lambda) - (I_J \otimes X')y\right]$$

$$= -\hat{p}'(\lambda)\ln\left[\frac{\exp(x_i'\lambda_j)}{\Omega_i(\lambda)}\right] + \lambda'[(I_J \otimes X')\hat{p}(\lambda) - (I_J \otimes X')y]$$

$$= -\hat{p}'(\lambda)\left[x_i'\lambda_j - \ln(\Omega_i(\lambda))\right] + \hat{p}'(\lambda)(I_J \otimes X)\lambda - y'(I_J \otimes X)\lambda$$

$$= -\hat{p}'(\lambda)(I_J \otimes X)\lambda + \sum_i \ln[\Omega_i(\lambda)] + \hat{p}'(\lambda)(I_J \otimes X)\lambda - y'(I_J \otimes X)\lambda$$

$$= -y'(I_J \otimes X)\lambda + \sum_i \ln\left[\sum_j \exp(x_i'\lambda_j)\right].$$

Normalizing $\lambda_J = 0$ yields

$$M(\lambda) = -y'(I_{J-1} \otimes X)\lambda + \sum_i \ln\left[1 + \sum_{j=1}^{J-1} \exp(x_i'\lambda_j)\right].$$

Since

$$y'(I_J \otimes X)\lambda = \sum_i \sum_j y_{ij} x_i' \lambda_j ,$$

we write the concentrated Lagrangian as

$$M(\lambda) = -\sum_{i=1}^{N} \sum_{j=1}^{J-1} y_{ij} x_i' \lambda_j + \sum_i \ln\left[1 + \sum_{j=1}^{J-1} \exp(x_i'\lambda_j)\right].$$

First-Order Condition:

$$\frac{\partial \mathrm{M}}{\partial \lambda} = -\sum_i y_{ij} x_i + \sum_i \frac{x_i \exp(x_i'\lambda_j)}{1 + \sum_{j=1}^{J-1} \exp(x_i'\lambda_j)}$$

$$= -\sum_i \left[y_{ij} - \frac{\exp(x_i'\lambda_j)}{1 + \sum_{j=1}^{J-1} \exp(x_i'\lambda_j)} \right] x_i$$

$$= -\sum_i (y_{ij} - p_{ij})\, x_i = 0 .$$

Second-Order Condition:

$$\frac{\partial^2 \mathrm{M}}{\partial \lambda_j \partial \lambda_j'} = \sum_i \frac{x_i x_i' \exp(x_i'\lambda_j)\Omega_i - x_i x_i' \exp(x_i'\lambda_j)\exp(x_i'\lambda_j)}{[\Omega_i(\lambda)]^2}$$

$$= \sum_i x_i x_i' \left[\frac{\exp(x_i'\lambda_j)}{\Omega_i} - \frac{\exp(x_i'\lambda_j)\exp(x_i'\lambda_j)}{\Omega_i^2} \right]$$

$$= \sum_i x_i x_i' \left[p_{ij} - p_{ij}^2 \right].$$

$$\frac{\partial^2 \mathrm{M}}{\partial \lambda_j \partial \lambda_k'} = \sum_i \frac{0 - x_i x_i' \exp(x_i'\lambda_j)\exp(x_i'\lambda_k)}{[\Omega_i(\lambda)]^2}$$

$$= \sum_i -x_i x_i' p_{ij} p_{ik}.$$

A.6 Multinomial Choice Problem – GME Primal Solution

The Objective Function:

$$H(p,w) = -p' \ln(p) - w' \ln(w) \tag{A.22}$$

subject to

$$(I_J \otimes X')y = (I_J \otimes X')p + (I_J \otimes X')Vw \tag{A.23}$$

$$\begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix} p = i_N \tag{A.24}$$

$$(I_{NJ} \otimes i_M')w = i_{NJ}. \tag{A.25}$$

The Lagrangian:

$$\mathcal{L} = -p' \ln(p) - w' \ln(w) + \lambda' [(I_J \otimes X')p + (I_J \otimes X')Vw - (I_J \otimes X')y]$$

$$+ \gamma' \left[i_N - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix} p \right] + \tau' \left[i_{NJ} - (I_{NJ} \otimes i_M')w \right],$$

where $\lambda = \begin{bmatrix} \lambda_1' & \lambda_2' & \cdots & \lambda_J' \end{bmatrix}'$ is a $KJ \times 1$ vector of parameter estimates.

First-Order Conditions:

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{2N} - \ln(p) + (I_J \otimes X)\lambda - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix}' \gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = (I_J \otimes X')p + (I_J \otimes X')Vw - (I_J \otimes X')y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_N - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix} p = 0$$

$$\frac{\partial \mathcal{L}}{\partial w} = -i_{NJM} - \ln(w) + V'(I_J \otimes X)\lambda - (I_{NJ} \otimes i_M)\tau = 0$$

$$\frac{\partial \mathcal{L}}{\partial \tau} = i_{NJ} - (I_{NJ} \otimes i'_M)w = 0\,.$$

Solve FOC's for p:

$$\ln(p) = (I_J \otimes X)\lambda - i_{NJ} - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix}' \gamma$$

$$\hat{p} = \exp\left[(I_J \otimes X)\hat{\lambda}\right] \exp\left[-i_{NJ} - \begin{bmatrix} I_{N_1} & I_{N_2} & \cdots & I_{N_J} \end{bmatrix}' \hat{\gamma}\right]$$

$$\Rightarrow \hat{p}_{ij} = \exp(x_i'\hat{\lambda}_j)\exp(-1 - \hat{\gamma}_i)\,.$$

Since additivity requires that $p_{i1} + p_{i2} + \ldots + p_{iJ} = 1$ and $\exp(-\hat{\gamma}_i - 1)$ is constant for a given individual, the optimal ME probabilities can be rewritten as

$$\hat{p}_{ij} = \frac{\exp(-1 - \hat{\gamma}_i)\exp(x_i'\hat{\lambda}_j)}{\exp(-1 - \hat{\gamma}_i)\sum_{j=1}^{J}\exp(x_i'\hat{\lambda}_j)} = \frac{\exp(x_i'\hat{\lambda}_j)}{\sum_{j=1}^{J}\exp(x_i'\hat{\lambda}_j)}\,.$$

Replacing λ with β and normalizing $\hat{\beta}_J = 0$ yields

$$\hat{p}_{ij} = \frac{\exp(x_i'\hat{\beta}_j)}{1 + \sum_{j=1}^{J-1}\exp(x_i'\hat{\beta}_j)} = \frac{\exp(x_i'\hat{\beta}_j)}{\Omega_i(\hat{\beta})}\,. \qquad \text{(A.26)}$$

Solve FOC's for w:

$$\ln(w) = V'(I_J \otimes X)\lambda - i_{NJM} - (I_{NJ} \otimes i_M)\tau$$

$$w = \exp\left[V'(I_J \otimes X)\hat{\lambda}\right]\exp\left[-i_{NJM} - (I_{NJ} \otimes i_M)\hat{\tau}\right]$$

$$\Rightarrow w_{ijm} = \exp(x_i'\hat{\lambda}_j v_{ijm})\exp\left[-1 - \hat{\tau}_{ij}\right].$$

Because additivity requires that $\sum_{m=1}^{M} w_{ijm} = 1$, and $\exp(-\tau_{ij} - 1)$ is constant for a given i and j as we vary m, the optimal GME error weights may be rewritten as

$$w_{ijm} = \frac{\exp(-1-\hat{\tau}_{ij})\exp(x_i'\hat{\lambda}_j v_{ijm})}{\exp(-1-\hat{\tau}_{ij})\sum_{m=1}^{M}\exp(x_i'\hat{\lambda}_j v_{ijm})} = \frac{\exp(x_i'\hat{\lambda}_j v_{ijm})}{\sum_{m=1}^{M}\exp(x_i'\hat{\lambda}_j v_{ijm})}$$

We again replace λ with β and normalize $\hat{\beta}_2 = 0$ to obtain the optimal GME error weights

$$\hat{w}_{ijm} = \frac{\exp(x_i'\hat{\beta}_j v_{ijm})}{\sum_{m=1}^{M}\exp(x_i'\hat{\beta}_j v_{ijm})} = \frac{\exp(x_i'\hat{\beta}_j v_{ijm})}{\Psi_{ij}(\hat{\beta}_j)}, \tag{A.27}$$

Second-Order Conditions:

$$\frac{\partial^2 \mathcal{L}}{\partial p \partial p'} = -\frac{1}{p}$$

on the diagonal elements and zero on the off-diagonals;

$$\frac{\partial^2 \mathcal{L}}{\partial w \partial w'} = -\frac{1}{w}$$

on the diagonal elements and zero on the off-diagonals. Therefore, the Hessian is negative definite and we obtain a unique global maximum.

A.6.1 Multinomial Choice Problem – GME Dual Solution

We substitute the optimal GME probabilities into the original Lagrangian to form the concentrated Lagrangian:

$$\begin{aligned}
M(\lambda) &= -\hat{p}'(\lambda)\ln(\hat{p}'(\lambda)) - \hat{w}'(\lambda)\ln(\hat{w}'(\lambda)) \\
&\quad + \lambda'\left[(I_J \otimes X')\hat{p}(\lambda) + (I_J \otimes X')V\hat{w}(\lambda) - (I_J \otimes X')y\right] \\
&= -\hat{p}'(\lambda)\ln\left[\frac{\exp(x_i'\lambda_j)}{\Omega_i(\lambda)}\right] - \hat{w}'(\lambda)\ln\left[\frac{\exp(x_i'\lambda_j v_{ijm})}{\Psi_{ij}(\lambda_j)}\right] \\
&\quad + \lambda'\left[(I_J \otimes X')\hat{p}(\lambda) + (I_J \otimes X')V\hat{w}(\lambda) - (I_J \otimes X')y\right] \\
&= -\hat{p}'(\lambda)\left[x_i'\lambda_j - \ln(\Omega_i(\lambda))\right] - \hat{w}'(\lambda)\left[x_i'\lambda_j v_{ijm} - \ln(\Psi_{ij}(\lambda_j))\right] \\
&\quad + \hat{p}'(\lambda)(I_J \otimes X)\lambda + \hat{w}'(\lambda)V'(I_J \otimes X)\lambda - y'(I_J \otimes X)\lambda \\
&= -\hat{p}'(\lambda)(I_2 \otimes X)\lambda + \sum_i \ln\left[\Omega_i(\lambda)\right] - \hat{w}'(\lambda)V'(I_2 \otimes X)\lambda
\end{aligned}$$

$$+\sum_t\sum_j \ln\left[\Psi_{tj}(\lambda_j)\right]+\hat{p}'(\lambda)(I_2\otimes X)\lambda+\hat{w}'(\lambda)V'(I_2\otimes X)\lambda-y'(I_2\otimes X)\lambda$$

$$=-y'(I_2\otimes X)\lambda+\sum_t \ln\left[\sum_j \exp(x_t'\lambda_j)\right]+\sum_t\sum_j \ln\left[\sum_m \exp(x_t'\lambda_j v_{tjm})\right].$$

Normalizing $\lambda_J=0$ yields

$$M(\lambda)=-y'(I_{J-1}\otimes X)\lambda+\sum_t \ln\left[1+\sum_{j=1}^{J-1}\exp(x_t'\lambda_j)\right]+\sum_t \ln\left[\sum_m \exp(x_t'\lambda_1 v_{t1m})\right]$$

$$+\sum_t \ln\left[\sum_m \exp(x_t'\lambda_2 v_{t2m})\right]+\ldots+\sum_t \ln\left[\sum_m \exp(x_t'\lambda_{J-1} v_{t(J-1)m})\right]+\sum_t \ln\left[\sum_m \exp(0)\right]$$

$$=-\sum_t\sum_j y_{tj}x_t'\lambda_j+\sum_t \ln\left[\sum_j \exp(x_t'\lambda_j)\right]+\sum_t\sum_j \ln\left[\sum_m \exp(x_t'\lambda_j v_{tjm})\right].$$

First-Order Condition:

$$\frac{\partial \mathrm{M}}{\partial \lambda}=-\sum_t y_{tj}x_t+\sum_t \frac{x_t\exp(x_t'\lambda_j)}{1+\sum_{j=1}^{J-1}\exp(x_t'\lambda_j)}+\sum_t\left[\sum_m \frac{x_t v_{tjm}\exp(x_t'\lambda_j v_{tjm})}{\Psi_{tj}(\lambda_j)}\right]$$

$$=-\sum_t\left[y_{tj}-\frac{\exp(x_t'\lambda_j)}{1+\sum_{j=1}^{J-1}\exp(x_t'\lambda_j)}-\sum_m \frac{v_{tjm}\exp(x_t'\lambda_j v_{tjm})}{\Psi_{tj}(\lambda_j)}\right]x_t$$

$$=-\sum_t\left[y_{tj}-p_{tj}-\sum_m v_{tjm}w_{tjm}\right]x_t=0.$$

Second-Order Condition:

$$\frac{\partial^2 \mathrm{M}}{\partial \lambda_j \partial \lambda_j'}=\sum_t x_t x_t'\left[p_{tj}-p_{tj}^2+\sum_m v_{tjm}w_{tjm}(v_{tjm}-\sum_m v_{tjm}w_{tjm})\right]$$

$$\frac{\partial^2 \mathrm{M}}{\partial \lambda_j \partial \lambda_k'}=\sum_t -x_t x_t' p_{tj}p_{tk}.$$

A.7 Linear Regression Problem – GME Primal Solution

The Objective Function:

$$\max H(p,w)=-p'\ln(p)-w'\ln(w) \qquad \text{(A.28)}$$

subject to

$$y = XZp + Vw \tag{A.29}$$

$$(I_K \otimes i'_M)p = i_K \tag{A.30}$$

$$(I_N \otimes i'_J)w = i_N\,. \tag{A.31}$$

The Lagrangian:

$$\mathcal{L} = -p'\ln(p) - w'\ln(w) + \lambda'(XZp + Vw - y) + \gamma'[i_K - (I_K \otimes i'_M)p] + \delta'[i_N - (I_N \otimes i'_J)w].$$

First-Order Conditions:

$$\frac{\partial \mathcal{L}}{\partial p} = -i_{km} - \ln(p) + Z'X'\lambda - (I_K \otimes i_M)\gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial w} = -i_{NJ} - \ln(w) + V'w - (I_N \otimes i_J)\delta = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = XZp + Vw - y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = i_K - (I_K \otimes i'_M)p = 0$$

$$\frac{\partial \mathcal{L}}{\partial \delta} = i_N - (I_N \otimes i'_J)w = 0.$$

Solve FOC's for p:

$$\ln(p) = Z'X'\lambda - i_{KM} - (I_K \otimes i_M)\gamma$$

$$\Rightarrow \hat{p} = \exp(Z'X'\hat{\lambda})\exp[-i_{KM} - (I_K \otimes i_M)\hat{\gamma}]$$

Since the additivity constraint requires that $\sum_{m=1}^{M} p_{km} = 1$ and $\exp(-1-\gamma_k)$ is constant for a given parameter, the optimal GME probabilities can be rewritten as

$$\hat{p}_{km} = \frac{\exp(z_{km}x'_k\hat{\lambda})}{\sum_{m=1}^{M}\exp(z_{km}x'_k\hat{\lambda})} = \frac{\exp(z_{km}x'_k\hat{\lambda})}{\Omega_k(\hat{\lambda})}. \tag{A.32}$$

Solve FOC's for w:

$$\ln(w) = V'\lambda - i_{NJ} - (I_N \otimes i_J)\delta$$

$$\hat{w} = \exp(V'\hat{\lambda})\exp\left[-i_{NJ} - (I_N \otimes i_J)\hat{\delta}\right]$$

$$\Rightarrow \hat{w}_{nj} = \exp(v_{nj}\hat{\lambda}_n)\exp(-1-\hat{\delta}_n).$$

Since the addivity constraint requires that $\sum_{j=1}^{J} w_{nj} = 1$ and $\exp(-1-\delta_n)$ is constant for a given error, we rewrite the optimal error probabilities as

$$\hat{w}_{nj} = \frac{\exp(v_{nj}\hat{\lambda}_n)}{\sum_{j=1}^{J}\exp(v_{nj}\hat{\lambda}_n)} = \frac{\exp(v_{nj}\hat{\lambda}_n)}{\Psi_n(\hat{\lambda}_n)}. \tag{A.33}$$

Second-Order Conditions:

$$\frac{\partial^2 \mathcal{L}}{\partial p \partial p'} = -\frac{1}{p}$$

on the diagonal elements and zero on the off-diagonals;

$$\frac{\partial^2 \mathcal{L}}{\partial w \partial w'} = -\frac{1}{w}$$

on the diagonal elements and zero on the off-diagonals. Therefore, the Hessian is negative definite and we obtain a unique global maximum.

A.7.1 Linear Regression Problem – GME Dual Solution

We substitute the optimal GME probabilities into the original Lagrangian to form the concentrated Lagrangian:

$$M(\lambda) = -\hat{p}'(\lambda)\ln(\hat{p}(\lambda)) - \hat{w}'(\lambda)\ln(\hat{w}(\lambda)) + \lambda'\left[XZ\hat{p}(\lambda) + V\hat{w}(\lambda) - y\right]$$

$$= -\hat{p}'(\lambda)\ln\left[\frac{\exp(z_{km}x_k'\lambda)}{\sum_m \exp(z_{km}x_k'\lambda)}\right] - \hat{w}'(\lambda)\ln\left[\frac{\exp(v_{nj}\lambda_n)}{\sum_j \exp(v_{nj}\lambda_n)}\right]$$

$$+\lambda'\left[XZ\hat{p}(\lambda) + V\hat{w}(\lambda) - y\right]$$

$$= -y'\lambda + \sum_k \ln\left[\Omega_k(\lambda)\right] + \sum_n \ln\left[\Psi_n(\lambda_n)\right]$$

$$= -y'\lambda + \sum_k \ln\left[\sum_m \exp(z_{km}x_k'\lambda)\right] + \sum_n \ln\left[\sum_j \exp(v_{nj}\lambda_n)\right].$$

First-Order Condition:

$$\frac{\partial \mathrm{M}}{\partial \lambda} = -y + \sum_k \left[\sum_m \frac{x_k z_{km} \exp(z_{km}x_k'\lambda)}{\Omega_k(\lambda)}\right] + \left[\sum_j \frac{v_{nj} \exp(v_{nj}\lambda_n)}{\Psi_n(\lambda_n)}\right].$$

APPENDIX B

GAUSS PROGRAMS TO OBTAIN ME AND GME SOLUTIONS

B.1 Program 1 – Jaynes' Dice Problem (ME Primal Solution)

```
/* Randy Campbell
** Jaynes' Dice Problem - Constrained Optimization
*/

new; library co;
et=hsec; " Date " datestr(date) "  Time " timestr(time);
output reset;

" ***** Maximum Entropy Solution to Jaynes' Dice Problem ***** ";

y=1.5;
do while y < 6;

#include co.ext;
coset;

let x[6,1] = 1 2 3 4 5 6;        /* The support */

proc fct(p);                  /* Write objective function */
 local a,j;
 a=zeros(6,1);

 j = 1;                      /* Define 0*ln(0) to be zero */
 do while j <= 6;
  if p[j] > 0;
   a[j] = ln(p[j]);
  else;
   a[j] = 0;
  endif;
 j = j+1;
 endo;

 retp(p'a);                  /* Objective function */
endp;

proc eqp(p);                  /* Define equality constraints */
   local c;
   c=zeros(2,1);
   c[1]=p'x-y;               /* Data constraint */
   c[2]=sumc(p)-1;              /* Additivity constraint */
   retp(c);
endp;
```

```
_co_Bounds = { 0 1 };          /* Bounds for the parameters */
_co_EqProc = &eqp;

start = { 1, 1, 1, 1, 1, 1 };
output file = c:\randy\programs\dice.out; output on;

format /rdn 1,1;
print;
print "Maximum Entropy Solution to Jaynes' Dice Problem";
print "Observed Average is Equal to " y;
print;

{ p,f,g,ret } = co( &fct,start );
call coprt(p,f,g,ret);

  y=y+.5;
  print;
endo;

output off;
```

B.2 Program 2 – Jaynes' Dice Problem (ME Dual Solution)

```
/* Randy Campbell
** Jaynes' Dice Problem: Dual - Unconstrained problem
*/

new;
et=hsec; " Date " datestr(date) "  Time " timestr(time);
output reset;

" ***** Maximum Entropy Dual Solution to Jaynes' Dice Problem ***** ";

y=1.5;
do while y < 6;

library co;
#include co.ext;

coset;

let x[6,1] = 1 2 3 4 5 6;      /* The support */
p = zeros(6,1);                /* The ME probabilities */

proc fct(lam);                 /* Write objective function */
   local omega,z;

   z = exp(-x*lam);
   omega = sumc(z);            /* Define omega */
   p = exp(-x*lam)/omega;      /* obtain ME probabilities */

   retp(lam*y+ln(omega));      /* The objective function */
```

```
endp;

start = { 1 };
output file = c:\randy\programs\dual.out; output on;

format /rdn 1,1;
print;
print "Maximum Entropy Dual Solution to Jaynes' Dice Problem";
print "Observed Average is Equal to " y;

{ x,f,g,ret } = co( &fct,start );
call coprt(x,f,g,ret);

format /rdn 1,4;
print;
print "The Maximum Entropy probabilities are: ";
print p;

  y=y+.5;
  print;
endo;

output off;
```

B.3 Program 3 – Binary Choice Problem (ME Dual Solution)

```
/* Randy Campbell
** c:\randy\sirmans\me.pgm
** ME (Dual) Estimation and Prediction in Sirmans Dataset
*/

new; library co pgraph; graphset;
et=hsec; " Date " datestr(date) "  Time " timestr(time);
beggtime = time;

#include co.ext;
coset;

@ -- Specify Variables -- @

t = 78;                    /* Total number of observations */
load dat[t,16] = c:\sirmans\tbl30.asc; /* Load data */
y1 = 1 - dat[.,1];             /* Set variable rate to 1 */
y2 = dat[.,1];                 /* Set fixed rate to 0 */
yd = y1|y2;
x1 = ones(t,1)~dat[.,2:16];
x = x1./10;
k = cols(x);
smpvar = sumc(y1);
smpfix = t-smpvar;
obs = seqa(1,1,t);
```

```
@ -- The Entropy Function -- @

proc fct(lam);                     /* Write objective function */
  local lnomega,sumomega;
  lnomega = ln(1+exp(x*lam));
  sumomega = sumc(lnomega);          /* Define sum of omegas */
  retp(-y1'x*lam+sumomega);         /* The objective function */
endp;

@ -- Analytical Gradient -- @

proc gp(lam);
  local q;
    q = exp(x*lam)./(1+exp(x*lam));  /* Sum of 1st derivatives */
  retp(-x'y1+x'q);
endp;

_co_Bounds = { -1e3 1e3 };
_co_GradProc = &gp;
_co_GradCheckTol = 1e-3;
_co_DirTol = 1e-5;
_co_GradMethod = 0;
_co_MaxIters = 12;

start = zeros(k,1);                /* Starting values */

@ -- ME Estimation -- @

{ b,f,g,ret } = co( &fct,start );
call coprt(b,f,g,ret);
b = b./10;

@ -- Prediction -- @

pvhat = exp(x1*b)./(1+exp(x1*b));    /* Variable rate prediction */
pfhat = 1./(1+exp(x1*b));            /* Fixed rate prediction */
yvhat = pvhat .>= 0.5;
cs = (y1 .== yvhat);
cs00 = (y1 .== yvhat .and y1 .== 0);  /* Number 0's predicted as 0's */
cs01 = (y1 ./= yvhat .and y1 .== 0);  /* Number 0's predicted as 1's */
cs10 = (y1 ./= yvhat .and y1 .== 1);  /* Number 1's predicted as 0's */
cs11 = (y1 .== yvhat .and y1 .== 1);  /* Number 1's predicted as 1's */
pcts = sumc(cs)/t;
num00 = sumc(cs00);
num01 = sumc(cs01);
num10 = sumc(cs10);
num11 = sumc(cs11);
pcts00 = sumc(cs00)/smpfix;
pcts01 = sumc(cs01)/smpfix;
pcts10 = sumc(cs10)/smpvar;
pcts11 = sumc(cs11)/smpvar;
```

```
@ -- Output -- @

output file = c:\randy\sirmans\me.out on; output reset;

"Memory to start with: " coreleft; print;
" **** ME estimation on Sirmans data **** "; print;
format /rd 10,5;
"The ME parameter estimates ";
"Variable    Estimate";
"intercept  "  b[1];
"BA        "  b[2];
"BS        "  b[3];
"NW        "  b[4];
"FI        "  b[5];
"PTS       "  b[6];
"MAT       "  b[7];
"MOB       "  b[8];
"MC        "  b[9];
"FTB       "  b[10];
"SE        "  b[11];
"YLD       "  b[12];
"MARG      "  b[13];
"CB        "  b[14];
"STL       "  b[15];
"LA        "  b[16];
print;

"Avg. variable probability " meanc(pvhat);
"Min. variable probability " minc(pvhat);
"Max. variable probability " maxc(pvhat); print;

"Avg. fixed probability " meanc(pfhat);
"Min. fixed probability " minc(pfhat);
"Max. fixed probability " maxc(pfhat); print;

format /rd 8,0;

"Number fixed mortgages (y=0)        " smpfix;
"  Number fixed predicted (y=0)      " num00;
"  Number variable predicted (y=1)   " num01;
"Number variable mortgages (y=1)     " smpvar;
"  Number fixed predicted (y=0)      " num10;
"  Number variable predicted (y=1)   " num11; print;

endtime = time;
timing = (endtime - beggtime)';
"Time used " timing;

output off;

end;
```

B.4 Program 4 – Binary Choice Problem (GME Dual Solution)

```
/* Randy Campbell
** c:\randy\sirmans\gme2m3.pgm
** GME (Dual) Estimation in Sirmans Dataset
*/

new; library co pgraph; graphset;
et=hsec; " Date " datestr(date) "  Time " timestr(time);
beggtime = time;

#include co.ext;
coset;

@ -- Specify Variables -- @

t = 78;                          /* Total number of observations */
load dat[t,16] = c:\sirmans\tbl30.asc; /* Load data */
y1 = 1 - dat[.,1];               /* Set variable rate to 1 */
y2 = dat[.,1];                   /* Set fixed rate to 0 */
yd = y1|y2;
x1 = ones(t,1)~dat[.,2:16];
x = x1./10;
k = cols(x);
vi = { -.1, 0, .1 };             /* The error support */
m=rows(vi);
smpvar = sumc(y1);
smpfix = t-smpvar;

@ -- The Entropy Function -- @

proc fct(lam);                   /* Write objective function */
 local lnomega,sumomega,psi1,lnpsi1,sumpsi1,psi2,lnpsi2,sumpsi2;
 lnomega = ln(1+exp(x*lam));
 sumomega = sumc(lnomega);          /* Define sum of omegas */
 psi1 = sumc(exp(x*lam*vi')');
 lnpsi1 = ln(psi1);
 sumpsi1 = sumc(lnpsi1);            /* Define sum of psi */
 psi2 = m*ones(t,1);
 lnpsi2 = ln(psi2);
 sumpsi2 = sumc(lnpsi2);
 retp(-y1'x*lam+sumomega+sumpsi1+sumpsi2); /* The objective function */
endp;

@ -- Analytical Gradient -- @

proc gp(lam);
 local psij,q,e;
  q = exp(x*lam)./(1+exp(x*lam));
  psij = sumc(exp(x*lam*vi')');
  e = (exp(x*lam*vi')./psij)*vi;
 retp(-x'y1+x'q+x'e);
endp;
```

```
_co_Bounds = { -1e3 1e3 };
_co_GradProc = &gp;
_co_GradCheckTol = 1e-3;
_co_DirTol = 1e-5;
_co_GradMethod = 0;
_co_MaxIters = 12;

start = .001*ones(k,1);            /* Starting values */

@ -- GME Estimation -- @

{ b,f,g,ret } = co( &fct,start );
call coprt(b,f,g,ret);
b = b./10;

@ -- Prediction -- @

phat = exp(x1*b)./(1+exp(x1*b));      /* In sample prediction */
psi = sumc(exp(x1*b*vi')');
ehat = (exp(x1*b*vi')./psi)*vi;
pvhat = phat+ehat;                 /* Variable rate prediction */
pfhat = 1./(1+exp(x1*b));            /* Fixed rate prediction */
yvhat = (pvhat .>= 0.5);
cs = (y1 .== yvhat);
cs00 = (y1 .== yvhat .and y1 .== 0);  /* Number 0's predicted as 0's */
cs01 = (y1 ./= yvhat .and y1 .== 0);  /* Number 0's predicted as 1's */
cs10 = (y1 ./= yvhat .and y1 .== 1);  /* Number 1's predicted as 0's */
cs11 = (y1 .== yvhat .and y1 .== 1);  /* Number 1's predicted as 1's */
pcts = sumc(cs)/t;
num00 = sumc(cs00);
num01 = sumc(cs01);
num10 = sumc(cs10);
num11 = sumc(cs11);
pcts00 = sumc(cs00)/smpfix;
pcts01 = sumc(cs01)/smpfix;
pcts10 = sumc(cs10)/smpvar;
pcts11 = sumc(cs11)/smpvar;

@ -- Output -- @

output file = c:\randy\sirmans\gme2m3.out on; output reset;

"Memory to start with: " coreleft; print;
" **** GME estimation on Sirmans data (M=3) **** "; print;
format /rd 10,5;
"The GME2 parameter estimates";
"Variable    Estimate";
"intercept  " b[1];
"BA         " b[2];
"BS         " b[3];
"NW         " b[4];
"FI         " b[5];
```

```
"PTS      " b[6];
"MAT      " b[7];
"MOB      " b[8];
"MC       " b[9];
"FTB      " b[10];
"SE       " b[11];
"YLD      " b[12];
"MARG     " b[13];
"CB       " b[14];
"STL      " b[15];
"LA       " b[16];
print;

"Avg. variable probability " meanc(pvhat);
"Min. variable probability " minc(pvhat);
"Max. variable probability " maxc(pvhat); print;

"Avg. fixed probability " meanc(pfhat);
"Min. fixed probability " minc(pfhat);
"Max. fixed probability " maxc(pfhat); print;

format /rd 8,0;

"Number fixed mortgages (y=0)       " smpfix;
"  Number fixed predicted (y=0)     " num00;
"  Number variable predicted (y=1)  " num01;
"Number variable mortgages (y=1)    " smpvar;
"  Number fixed predicted (y=0)     " num10;
"  Number variable predicted (y=1)  " num11; print;

format /rd 10,5;

"% correctly predicted " pcts;
"% variable correct    " pcts11;
"% fixed correct       " pcts00; print;

format /rd 5,2;
"M equals  " m;
"error weights vector  " vi'; print;

endtime = time;
timing = (endtime - beggtime)';
"Time used  " timing;

output off;

end;
```

B.5 Program 5 – Multinomial Choice Program (ME Dual Solution)

```
/* Randy Campbell
** c:\randy\pse\mmepse.pgm
** Multinomial ME (Dual) Estimation and Prediction
```

```
*/

new; library co pgraph; graphset;
et=hsec; " Date " datestr(date) "  Time " timestr(time);
beggtime = time;

#include co.ext;
coset;

@ -- Specify Variables -- @

t = 9450;                       /* Number of observations drawn */
load dat[t,11] = c:\randy\pse\jan29.dat;      /* Load data */
x = ones(t,1)~dat[.,3:11];
k = cols(x);
yval = dat[.,1];                 /* Actual value of y */
j = maxc(yval)+1;                  /* Number of choices */
alt = seqa(0,1,j);
y = (yval .== (ones(t,1).*alt'));
sumy = sumc(y);

@ -- The Entropy Function -- @

proc fct(lam);                   /* Write objective function */
  local bmat,yhat,lstar;
  bmat = (reshape(lam,j-1,k))';      /* Create k by j-1 matrix */
  bmat = bmat~zeros(k,1);            /* Create k by j matrix */
  yhat = x*bmat;
  lstar = sumc(sumc((y.*yhat)')) - sumc(ln(sumc(exp(yhat)')));
  retp(-lstar);                  /* The objective function */
endp;

@ -- Analytical Gradient -- @

proc gp(lam);
  local bmat,yhat,gvec,jiter,phat;
    bmat = (reshape(lam,j-1,k))';     /* Create k by j-1 matrix */
    bmat = bmat~zeros(k,1);           /* Create k by j matrix */
    yhat = x*bmat;
    phat = exp(yhat)./sumc(exp(yhat)');/* Selection probabilities */
    gvec = zeros((j-1)*k,1);
    jiter = 1;
    do while jiter le (j-1);
      gvec[((jiter-1)*k+1):(jiter*k)] =
         sumc((y[.,jiter] - phat[.,jiter]).*x);
      jiter = jiter+1;
    endo;
  retp(-gvec);
endp;

_co_Bounds = { -1e3 1e3 };
_co_GradProc = &gp;
_co_GradCheckTol = 1e-3;
```

```
_co_DirTol = 1e-5;
_co_GradMethod = 0;
_co_MaxIters = 12;

start = zeros(k*(j-1),1);          /* Starting values */

@ -- ME Estimation -- @

{ b,f,g,ret } = co( &fct,start );
call coprt(b,f,g,ret);

@ -- Multinomial ME Prediction -- @

bmat = (reshape(b,j-1,k))';         /* Create k by j-1 matrix */
bmat = bmat~zeros(k,1);             /* Create k by j matrix */

yhat = x*bmat;
phat = exp(yhat)./sumc(exp(yhat)');  /* Selection probabilities */
ypred = (phat .== maxc(phat'));
csmat = (y .== ypred);
cs = (sumc(csmat') .== j);          /* Correct predictions */
cs1 = (cs .== 1 .and y[.,1] .== 1);
cs2 = (cs .== 1 .and y[.,2] .== 1);
cs3 = (cs .== 1 .and y[.,3] .== 1);
cs4 = (cs .== 1 .and y[.,4] .== 1);
pcts = sumc(cs)/t;
pcts1 = sumc(cs1)/sumy[1];
pcts2 = sumc(cs2)/sumy[2];
pcts3 = sumc(cs3)/sumy[3];
pcts4 = sumc(cs4)/sumy[4];

@ -- Output -- @

output file = c:\randy\pse\mmepse.out on; output reset;

format 10,6;
" **** Multinomial ME Coefficients **** "; print;
"Coefficients   : " b'; print;
"Number of choices: " j;
"Number of parameters: " k;
"Number of Observations: " rows(y); print;
"% Correct Predictions: " pcts;
"% Correct Predictions(y=1): " pcts1;
"% Correct Predictions(y=2): " pcts2;
"% Correct Predictions(y=3): " pcts3;
"% Correct Predictions(y=4): " pcts4;

output off;

end;
```

B.6 Program 6 – Multinomial Choice Problem (GME Dual Solution)

```
/* Randy Campbell
** c:\randy\pse\mgme2pse.pgm
** Multinomial GME2 (Dual) Estimation and Prediction
*/

new; library co pgraph; graphset;
et=hsec; " Date " datestr(date) "  Time " timestr(time);
beggtime = time;

#include co.ext;
coset;

@ -- Specify Variables -- @

t = 9450;                        /* Number of observations drawn */
load dat[t,11] = c:\randy\pse\jan29.dat;      /* Load data */
x = ones(t,1)~dat[.,3:11];
k = cols(x);
yval = dat[.,1];                 /* Actual value of y */
j = maxc(yval)+1;                   /* Number of choices */
alt = seqa(0,1,j);
y = (yval .== (ones(t,1).*alt'));
sumy = sumc(y);

vij1 = -.1;                    /* Lower error bound */
vij2 = 0;                      /* Middle error bound */
vij3 = .1;                     /* Upper error bound */
vi = vij1|vij2|vij3;              /* The error support */
m = 3;

@ -- The Entropy Function -- @

proc fct(lam);                   /* Write objective function */
  local bmat,yhat,emat,lstar;
  bmat = (reshape(lam,j-1,k))';       /* Create k by j-1 matrix */
  bmat = bmat~zeros(k,1);            /* Create k by j matrix */
  yhat = x*bmat;
  emat = (yhat).*.vi';
  lstar = sumc(sumc((y.*yhat)')) - sumc(ln(sumc(exp(yhat)')))
        - sumc(ln(sumc(exp(emat)')));
  retp(-lstar);                 /* The objective function */
endp;

@ -- Analytical Gradient -- @

proc gp(lam);
  local bmat,yhat,emat,emat2,psij,enum,errmat,ehat,gvec,jiter,phat;
    bmat = (reshape(lam,j-1,k))';      /* Create k by j-1 matrix */
    bmat = bmat~zeros(k,1);           /* Create k by j matrix */
    yhat = x*bmat;
    phat = exp(yhat)./sumc(exp(yhat)');/* Selection probabilities */
```

```
    emat = (yhat).*.vi';
    emat2 = (reshape(emat,t*j,m));     /* NJ by M error matrix */
    psij = sumc(exp(emat2)');
    enum = exp(emat2)*vi;
    errmat = enum./psij;
    ehat = (reshape(errmat,t,j));
    gvec = zeros((j-1)*k,1);
    jiter = 1;
    do while jiter le (j-1);
      gvec[((jiter-1)*k+1):(jiter*k)] =
        sumc((y[.,jiter] - phat[.,jiter] - ehat[.,jiter]).*x);
      jiter = jiter+1;
    endo;
  retp(-gvec);
endp;

_co_Bounds = { -1e3 1e3 };
_co_GradProc = &gp;
_co_GradCheckTol = 1e-3;
_co_DirTol = 1e-5;
_co_GradMethod = 0;
_co_MaxIters = 12;

start = zeros(k*(j-1),1);          /* Starting values */

@ -- GME Estimation -- @

{ b,f,g,ret } = co( &fct,start );
call coprt(b,f,g,ret);

@ -- Multinomial GME Prediction -- @

bmat = (reshape(b,j-1,k))';         /* Create k by j-1 matrix */
bmat = bmat~zeros(k,1);             /* Create k by j matrix */

yhat = x*bmat;
phat = exp(yhat)./sumc(exp(yhat)');  /* Selection probabilities */
emat = (yhat).*.vi';
emat2 = (reshape(emat,t*j,m));      /* NJ by M error matrix */
psij = sumc(exp(emat2)');
enum = exp(emat2)*vi;
errmat = enum./psij;
ehat = (reshape(errmat,t,j));       /* Predicted Errors */
ypred = ((phat+ehat) .== maxc((phat+ehat)'));
csmat = (y .== ypred);
cs = (sumc(csmat') .== j);          /* Correct predictions */
cs1 = (cs .== 1 .and y[.,1] .== 1);
cs2 = (cs .== 1 .and y[.,2] .== 1);
cs3 = (cs .== 1 .and y[.,3] .== 1);
cs4 = (cs .== 1 .and y[.,4] .== 1);
pcts = sumc(cs)/t;
pcts1 = sumc(cs1)/sumy[1];
pcts2 = sumc(cs2)/sumy[2];
```

```
pcts3 = sumc(cs3)/sumy[3];
pcts4 = sumc(cs4)/sumy[4];

@ -- Output -- @

output file = c:\randy\pse\mgme2pse.out on; output reset;

format 10,6;
" **** Multinomial GME2 Coefficients **** "; print;
"Coefficients   : " b'; print;
"Number of choices: " j;
"Number of parameters: " k;
"Number of Observations: " rows(y); print;
"% Correct Predictions: " pcts;
"% Correct Predictions(y=1): " pcts1;
"% Correct Predictions(y=2): " pcts2;
"% Correct Predictions(y=3): " pcts3;
"% Correct Predictions(y=4): " pcts4;

endtime = time;
timing = (endtime - beggtime)';
"Time used  " timing;

output off;

end;
```

B.7 Program 7 – Linear Regression Problem (GME Dual Solution)

```
/* Randy Campbell
** c:\randy\tuna1\gme2ps1.pgm
** Monte Carlo Simulation - GME Estimation of Tuna Dataset
*/

new; library co pgraph; graphset;
et=hsec; " Date " datestr(date) "  Time " timestr(time);
beggtime = time;

#include co.ext;
coset;

@ -- Specify Variables -- @

t = 156;                         /* Total number of observations */
load dat[t,22] = c:\randy\tuna\dat\chain.dat; /* Load data */
chain = dat[.,1];
week = dat[.,2];
sales = dat[.,3]~dat[.,7]~dat[.,11]~dat[.,15];
apr = dat[.,4]~dat[.,8]~dat[.,12]~dat[.,16];
rpr = dat[.,5]~dat[.,9]~dat[.,13]~dat[.,17];
c = dat[.,6]~dat[.,10]~dat[.,14]~dat[.,18];

discount = (rpr-apr)./rpr;
```

```
mad = (c .== 2);
dis = (c .== 8);
dismad = (c .== 10);
q1 = (week .le 13);
q2 = (week .ge 14 .and week .le 26);
q3 = (week .ge 27 .and week .le 39);
ch1 = (chain .eq 1);
ch2 = (chain .eq 2);

y = ln(sales[.,1]);
x = ones(t,1)~rpr[.,1]~discount[.,1]~dis[.,1]~dismad[.,1]~rpr[.,2]~
   discount[.,2]~dis[.,2]~dismad[.,2]~rpr[.,3]~discount[.,3]~dis[.,3]~
   dismad[.,3]~rpr[.,4]~discount[.,4]~dis[.,4]~dismad[.,4];

@ -- Parameter Support -- @

z1 = { -20, -10, 0, 10, 20 };
z2 = { -20, -10, 0, 10, 20 };
z3 = { -20, -10, 0, 10, 20 };
z4 = { -10, -5, 0, 5, 10 };
z5 = { -10, -5, 0, 5, 10 };
z6 = { -20, -10, 0, 10, 20 };
z7 = { -20, -10, 0, 10, 20 };
z8 = { -10, -5, 0, 5, 10 };
z9 = { -10, -5, 0, 5, 10 };
z10 = { -20, -10, 0, 10, 20 };
z11 = { -20, -10, 0, 10, 20 };
z12 = { -10, -5, 0, 5, 10 };
z13 = { -10, -5, 0, 5, 10 };
z14 = { -20, -10, 0, 10, 20 };
z15 = { -20, -10, 0, 10, 20 };
z16 = { -10, -5, 0, 5, 10 };
z17 = { -10, -5, 0, 5, 10 };
m = rows(z1);
zmat = z1~z2~z3~z4~z5~z6~z7~z8~z9~z10~z11~z12~z13~z14~z15~z16~z17;
im = ones(m,1);

vi = { -3, 0, 3 };
j = rows(vi);
v = eye(t).*.vi';
ij = ones(j,1);

@ -- The Entropy Function -- @

proc fct(lam);                    /* Write objective function */
 local xtb,xmat,zxtb,omega,lnomega,sumomega,vtb,psi,lnpsi,sumpsi;
 xtb = x'lam;
 xmat = im*xtb';                  /* M by K matrix */
 zxtb = zmat.*xmat;
 omega = sumc(exp(zxtb));              /* K by 1 vector */
 lnomega = ln(omega);
 sumomega = sumc(lnomega);             /* Define sum of omega */
 vtb = vi*lam';                   /* J by T matrix */
```

```
  psi = sumc(exp(vtb));              /* T by 1 vector */
  lnpsi = ln(psi);
  sumpsi = sumc(lnpsi);              /* Define sum of psi */
  retp(-y'lam+sumomega+sumpsi);      /* The objective function */
endp;

@ -- Analytical Gradient -- @

proc gp(lam);
  local xtb,xmat,zxtb,expzxtb,omega,zomega,gmat,gvec,
     vtb,expvtb,psi,wvec,gvec2;
  xtb = x'lam;
  xmat = im*xtb';
  zxtb = zmat.*xmat;
  expzxtb = exp(zxtb);               /* M by K matrix */
  omega = sumc(expzxtb);             /* K by 1 vector */
  zomega = zmat.*expzxtb;            /* M by K matrix */
  gmat = zomega'./omega;             /* K by M matrix */
  gvec = sumc((x*gmat)');            /* T by 1 vector */
  vtb = vi*lam';                     /* J by T matrix */
  expvtb = exp(vtb);
  psi = sumc(expvtb);                /* T by 1 vector */
  wvec = (vi'expvtb)';               /* T by 1 vector */
  gvec2 = wvec./psi;                 /* T by 1 vector */
  retp(-y+gvec+gvec2);
endp;

_co_GradProc = &gp;
_co_GradCheckTol = 1e-3;
_co_DirTol = 1e-5;
_co_GradMethod = 0;
_co_MaxIters = 12;

start = zeros(t,1);                  /* Starting values */

@ -- GME Dual Estimation -- @

{ lam,f,g,ret } = co( &fct,start );
call coprt(lam,f,g,ret);

xtb = x'lam;
xmat = im*xtb';                      /* M by K matrix */
zxtb = zmat.*xmat;                   /* M by K matrix */
omega = sumc(exp(zxtb));             /* K by 1 vector */
phat = exp(zxtb')./omega;            /* K by M matrix */
bmat = phat*zmat;
bgme = diag(bmat);

vtb = vi*lam';                       /* J by T matrix */
psi = sumc(exp(vtb));                /* T by 1 vector */
what = (exp(vtb'))./psi;             /* T by J matrix */
wvec = reshape(what,t*j,1);          /* TJ by 1 matrix */
errgme = v*wvec;                     /* T by 1 matrix of errors */
```

```
@ -- Output -- @

output file = c:\randy\tuna1\gme2ps1.out on; output reset;

"Memory to start with: " coreleft; print;
format /rd 10,5;
" **** GME2 Estimates for Tuna data - All 3 Chains **** ";
print;
"Variable    Estimate ";
"intercept " bgme[1];
"price1     " bgme[2];
"discount1 " bgme[3];
"dis1       " bgme[4];
"dismad1    " bgme[5];
"price2     " bgme[6];
"discount2 " bgme[7];
"dis2       " bgme[8];
"dismad2    " bgme[9];
"price3     " bgme[10];
"discount3 " bgme[11];
"dis3       " bgme[12];
"dismad3    " bgme[13];
"price4     " bgme[14];
"discount4 " bgme[15];
"dis4       " bgme[16];
"dismad4    " bgme[17]; print;

format /rd 8,0;
"Number of Observations " t; print; print;

format /rd 2.0;
endtime = time;
timing = (endtime - beggtime)';
"Time used  " timing;

output off;

end;
```

APPENDIX C

TABLES FOR BINARY CHOICE MONTE CARLO EXPERIMENTS

Table C.1 Used Cars Data - Monte Carlo Prediction 1 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.3112	-2.2385	-2.2989	-2.2385	-0.4569	-2.0938	-2.2369
	(-30.43)	(-26.90)	(-29.10)	(-26.90)	(-47.36)	(-29.50)	(-26.93)
Cdcmi	-.00004	-.00008	-.00007	-.00008	-.00001	-.00007	-.00008
	(-3.61)	(-3.63)	(-3.89)	(-3.63)	(-3.88)	(-3.69)	(-3.63)
Cdfpay	.00163	.00293	.00328	.00293	.00041	.00260	.00292
	(5.36)	(5.26)	(5.26)	(5.26)	(5.25)	(5.26)	(5.26)
Cdtdp	-.00205	-.00332	-.00334	-.00332	-.00042	-.00290	-.00332
	(-2.71)	(-2.40)	(-2.54)	(-2.40)	(-2.54)	(-2.43)	(-2.40)
Age	-.00390	-.00777	-.00718	-.00777	-.00091	-.00666	-.00775
	(-4.26)	(-4.29)	(-4.49)	(-4.29)	(-4.50)	(-4.35)	(-4.30)
Amtfin	.00001	.00003	.00003	.00003	.00000	.00002	.00003
	(1.63)	(1.78)	(1.52)	(1.78)	(1.53)	(1.73)	(1.78)
Netrate	.00248	.00558	.00342	.00558	.00045	.00450	.00556
	(1.59)	(1.85)	(1.20)	(1.85)	(1.25)	(1.73)	(1.84)
Home	-.19754	-.37652	-.36012	-.37652	-.04570	-.32459	-.37588
	(-5.64)	(-5.58)	(-6.05)	(-5.58)	(-6.01)	(-5.69)	(-5.59)
Cosign	-.08344	-.15468	-.16096	-.15468	-.02031	-.13545	-.15444
	(-1.70)	(-1.66)	(-1.87)	(-1.66)	(-1.86)	(-1.70)	(-1.66)
Estimation Sample							
Percent Correct	80.92	80.97	80.50	80.97	80.90	80.96	80.97
Standard Deviation	.0108	.0105	.0102	.0105	.0125	.0109	.0105
% Repay. Correct	89.62	89.67	88.99	89.67	89.59	89.65	89.67
% Defaults Correct	19.98	20.08	21.05	20.08	20.07	20.09	20.08
Hold-out Sample							
Percent Correct	80.91	80.96	80.49	80.96	80.89	80.94	80.96
Standard Deviation	.0118	.0114	.0113	.0114	.0133	.0118	.0115
% Repay. Correct	89.61	89.66	88.99	89.66	89.58	89.64	89.66
% Defaults Correct	19.75	19.84	20.83	19.84	19.82	19.85	19.85

* N=20,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.2 Used Cars Data – Monte Carlo Prediction 2 (See Table 2.3 for design)

Variable	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-.20741	-.33803	-.33436	-.33803	-.08870	-.32818	-.33793
	(-6.39)	(-6.51)	(-6.42)	(-6.51)	(-6.47)	(-6.51)	(-6.51)
Cdcmi	-.00004	-.00007	-.00007	-.00007	-.00002	-.00007	-.00007
	(-5.69)	(-5.70)	(-5.75)	(-5.70)	(-5.76)	(-5.71)	(-5.70)
Cdfpay	.00203	.00336	.00322	.00336	.00086	.00326	.00336
	(8.05)	(8.84)	(8.59)	(8.84)	(8.74)	(8.84)	(8.84)
Cdtdp	-.00308	-.00497	-.00492	-.00497	-.00130	-.00483	-.00497
	(-4.40)	(-4.42)	(-4.40)	(-4.42)	(-4.41)	(-4.42)	(-4.42)
Age	-.00426	-.00687	-.00688	-.00687	-.00182	-.00667	-.00687
	(-6.95)	(-6.90)	(-6.94)	(-6.90)	(-6.96)	(-6.90)	(-6.90)
Amtfin	.00001	.00002	.00002	.00002	.00001	.00002	.00002
	(1.93)	(1.89)	(2.19)	(1.89)	(2.14)	(1.91)	(1.89)
Netrate	.00263	.00399	.00449	.00399	.00116	.00391	.00399
	(2.15)	(2.04)	(2.29)	(2.04)	(2.25)	(2.06)	(2.04)
Home	-.22829	-.36668	-.36717	-.36668	-.09711	-.35620	-.36657
	(-9.22)	(-9.18)	(-9.18)	(-9.18)	(-9.26)	(-9.19)	(-9.18)
Cosign	-.09541	-.15273	-.15301	-.15273	-.04048	-.14838	-.15268
	(-2.45)	(-2.45)	(-2.46)	(-2.45)	(-2.46)	(-2.45)	(-2.45)
Estimation Sample							
Percent Correct	58.11	58.12	58.11	58.12	58.11	58.12	58.12
Standard Deviation	.0035	.0035	.0035	.0035	.0035	.0035	.0035
% Repay. Correct	59.36	59.37	59.39	59.37	59.39	59.37	59.37
% Defaults Correct	56.86	56.86	56.82	56.86	56.83	56.86	56.86
Hold-out Sample							
Percent Correct	58.05	58.07	58.06	58.07	58.06	58.06	58.07
Standard Deviation	.0071	.0070	.0071	.0070	.0071	.0071	.0070
% Repay. Correct	59.33	59.34	59.37	59.34	59.37	59.35	59.34
% Defaults Correct	56.77	56.79	56.74	56.79	56.75	56.78	56.79

* N=20,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.3 Used Cars Data - Monte Carlo Prediction 3 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.3142	-2.2436	-2.3081	-2.2436	-0.4574	-2.0979	-2.2419
	(-15.77)	(-13.97)	(-15.04)	(-13.97)	(-24.54)	(-15.33)	(-13.99)
Cdcmi	-.00004	-.00008	-.00007	-.00008	-.00001	-.00007	-.00008
	(-1.87)	(-1.88)	(-1.97)	(-1.88)	(-1.95)	(-1.90)	(-1.88)
Cdfpay	.00169	.00302	.00338	.00302	.00042	.00269	.00302
	(2.74)	(2.70)	(2.61)	(2.70)	(2.62)	(2.69)	(2.70)
Cdtdp	-.00215	-.00352	-.00349	-.00352	-.00044	-.00306	-.00351
	(-1.29)	(-1.15)	(-1.20)	(-1.15)	(-1.19)	(-1.16)	(-1.15)
Age	-.00400	-.00798	-.00730	-.00798	-.00093	-.00682	-.00796
	(-2.33)	(-2.35)	(-2.43)	(-2.35)	(-2.42)	(-2.37)	(-2.35)
Amtfin	.00001	.00003	.00002	.00003	.00000	.00002	.00003
	(0.70)	(0.78)	(0.64)	(0.78)	(0.65)	(0.75)	(0.78)
Netrate	.00260	.00582	.00356	.00582	.00047	.00469	.00581
	(0.83)	(0.98)	(0.63)	(0.98)	(0.65)	(0.92)	(0.98)
Home	-.20155	-.38401	-.36477	-.38401	-.04628	-.33024	-.38335
	(-3.08)	(-3.05)	(-3.30)	(-3.05)	(-3.28)	(-3.11)	(-3.05)
Cosign	-.09000	-.16901	-.16842	-.16901	-.02128	-.14632	-.16872
	(-0.93)	(-0.91)	(-1.00)	(-0.91)	(-1.00)	(-0.93)	(-0.91)
Estimation Sample							
Percent Correct	80.45	80.54	80.19	80.54	80.41	80.52	80.53
Standard Deviation	.0214	.0209	.0198	.0209	.0245	.0215	.0209
% Repay. Correct	88.87	88.96	88.46	88.96	88.79	88.93	88.96
% Defaults Correct	21.42	21.49	22.20	21.49	21.61	21.52	21.49
Hold-out Sample							
Percent Correct	80.28	80.37	80.01	80.37	80.22	80.34	80.37
Standard Deviation	.0212	.0207	.0198	.0207	.0245	.0213	.0207
% Repay. Correct	88.78	88.87	88.36	88.87	88.69	88.84	88.87
% Defaults Correct	20.72	20.77	21.48	20.77	20.87	20.78	20.76

* N=5,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.4 Used Cars Data - Monte Carlo Prediction 4 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-.20570	-.33475	-.33188	-.33475	-.08786	-.32498	-.33465
	(-3.25)	(-3.29)	(-3.28)	(-3.29)	(-3.29)	(-3.29)	(-3.29)
Cdcmi	-.00005	-.00008	-.00007	-.00008	-.00002	-.00007	-.00008
	(-2.76)	(-2.77)	(-2.79)	(-2.77)	(-2.80)	(-2.77)	(-2.77)
Cdfpay	.00201	.00332	.00318	.00332	.00085	.00321	.00332
	(3.98)	(4.17)	(4.12)	(4.17)	(4.16)	(4.17)	(4.17)
Cdtdp	-.00311	-.00501	-.00495	-.00501	-.00131	-.00486	-.00501
	(-2.12)	(-2.13)	(-2.12)	(-2.13)	(-2.12)	(-2.13)	(-2.13)
Age	-.00429	-.00693	-.00693	-.00693	-.00183	-.00673	-.00693
	(-3.13)	(-3.11)	(-3.13)	(-3.11)	(-3.14)	(-3.11)	(-3.11)
Amtfin	.00002	.00002	.00003	.00002	.00001	.00002	.00002
	(1.08)	(1.05)	(1.19)	(1.05)	(1.17)	(1.06)	(1.05)
Netrate	.00251	.00383	.00431	.00383	.00112	.00375	.00383
	(1.06)	(1.01)	(1.13)	(1.01)	(1.12)	(1.02)	(1.01)
Home	-.22811	-.36665	-.36678	-.36665	-.09688	-.35610	-.36655
	(-4.52)	(-4.51)	(-4.51)	(-4.51)	(-4.54)	(-4.51)	(-4.51)
Cosign	-.09685	-.15511	-.15535	-.15511	-.04103	-.15066	-.15506
	(-1.29)	(-1.29)	(-1.30)	(-1.29)	(-1.30)	(-1.29)	(-1.29)
Estimation Sample							
Percent Correct	58.23	58.24	58.24	58.24	58.24	58.24	58.24
Standard Deviation	.0070	.0070	.0069	.0070	.0069	.0070	.0070
% Repay. Correct	59.51	59.51	59.55	59.51	59.54	59.52	59.51
% Defaults Correct	56.95	56.96	56.93	56.96	56.93	56.95	56.96
Hold-out Sample							
Percent Correct	57.92	57.94	57.93	57.94	57.93	57.93	57.94
Standard Deviation	.0069	.0068	.0069	.0068	.0069	.0068	.0068
% Repay. Correct	59.15	59.16	59.18	59.16	59.18	59.16	59.16
% Defaults Correct	56.70	56.71	56.68	56.71	56.68	56.70	56.71

* N=5,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.5 Used Cars Data - Monte Carlo Prediction 5 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.3191	-2.2554	-2.3258	-2.2554	-0.4586	-2.1075	-2.2537
	(-10.29)	(-8.98)	(-9.76)	(-8.98)	(-15.86)	(-9.88)	(-8.99)
Cdcmi	-.00004	-.00009	-.00007	-.00009	-.00001	-.00007	-.00009
	(-1.28)	(-1.29)	(-1.35)	(-1.29)	(-1.34)	(-1.30)	(-1.29)
Cdfpay	.00170	.00304	.00340	.00304	.00043	.00270	.00304
	(1.80)	(1.75)	(1.70)	(1.75)	(1.71)	(1.75)	(1.75)
Cdtdp	-.00183	-.00291	-.00291	-.00291	-.00037	-.00253	-.00291
	(-0.74)	(-0.64)	(-0.67)	(-0.64)	(-0.67)	(-0.65)	(-0.64)
Age	-.00392	-.00784	-.00704	-.00784	-.00089	-.00666	-.00782
	(-1.40)	(-1.41)	(-1.43)	(-1.41)	(-1.44)	(-1.43)	(-1.41)
Amtfin	.00001	.00003	.00003	.00003	.00000	.00002	.00003
	(0.52)	(0.57)	(0.46)	(0.57)	(0.47)	(0.55)	(0.57)
Netrate	.00231	.00539	.00298	.00539	.00040	.00428	.00538
	(0.48)	(0.58)	(0.34)	(0.58)	(0.35)	(0.53)	(0.58)
Home	-.20251	-.38503	-.36011	-.38503	-.04568	-.32934	-.38433
	(-1.92)	(-1.88)	(-2.04)	(-1.88)	(-2.03)	(-1.92)	(-1.88)
Cosign	-.08887	-.17071	-.16153	-.17071	-.02041	-.14534	-.17038
	(-0.60)	(-0.60)	(-0.64)	(-0.60)	(-0.64)	(-0.61)	(-0.60)
Estimation Sample							
Percent Correct	79.93	80.08	79.88	80.08	79.86	80.05	80.08
Standard Deviation	.0309	.0302	.0288	.0302	.0359	.0312	.0303
% Repay. Correct	87.97	88.14	87.88	88.14	87.85	88.10	88.14
% Defaults Correct	23.46	23.39	23.66	23.39	23.61	23.42	23.39
Hold-out Sample							
Percent Correct	79.51	79.63	79.46	79.63	79.41	79.60	79.63
Standard Deviation	.0301	.0295	.0281	.0295	.0355	.0305	.0295
% Repay. Correct	87.77	87.92	87.68	87.92	87.63	87.88	87.92
% Defaults Correct	21.78	21.67	22.05	21.67	21.96	21.75	21.67

* N=2,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.6 Used Cars Data - Monte Carlo Prediction 6 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-.21254	-.34458	-.34197	-.34458	-.09030	-.33445	-.34447
	(-2.08)	(-2.08)	(-2.08)	(-2.08)	(-2.09)	(-2.08)	(-2.08)
Cdcmi	-.00005	-.00007	-.00007	-.00007	-.00002	-.00007	-.00007
	(-1.81)	(-1.82)	(-1.83)	(-1.82)	(-1.83)	(-1.82)	(-1.82)
Cdfpay	.00209	.00342	.00328	.00342	.00087	.00331	.00342
	(2.64)	(2.67)	(2.72)	(2.67)	(2.73)	(2.68)	(2.67)
Cdtdp	-.00300	-.00484	-.00477	-.00484	-.00126	-.00469	-.00483
	(-1.36)	(-1.36)	(-1.36)	(-1.36)	(-1.36)	(-1.36)	(-1.36)
Age	-.00424	-.00685	-.00684	-.00685	-.00180	-.00665	-.00685
	(-2.11)	(-2.09)	(-2.11)	(-2.09)	(-2.12)	(-2.10)	(-2.09)
Amtfin	.00001	.00002	.00002	.00002	.00001	.00002	.00002
	(0.57)	(0.56)	(0.65)	(0.56)	(0.64)	(0.56)	(0.56)
Netrate	.00251	.00387	.00437	.00387	.00113	.00379	.00387
	(0.66)	(0.63)	(0.72)	(0.63)	(0.70)	(0.64)	(0.63)
Home	-.23290	-.37473	-.37449	-.37473	-.09872	-.36384	-.37462
	(-2.93)	(-2.92)	(-2.92)	(-2.92)	(-2.94)	(-2.93)	(-2.92)
Cosign	-.10228	-.16390	-.16392	-.16390	-.04320	-.15915	-.16385
	(-0.87)	(-0.87)	(-0.87)	(-0.87)	(-0.87)	(-0.87)	(-0.87)
Estimation Sample							
Percent Correct	58.36	58.38	58.37	58.38	58.37	58.37	58.38
Standard Deviation	.0107	.0108	.0107	.0108	.0107	.0108	.0108
% Repay. Correct	59.16	59.20	59.18	59.20	59.18	59.19	59.20
% Defaults Correct	57.56	57.56	57.56	57.56	57.56	57.56	57.56
Hold-out Sample							
Percent Correct	57.66	57.66	57.67	57.66	57.66	57.66	57.66
Standard Deviation	.0075	.0075	.0075	.0075	.0075	.0075	.0075
% Repay. Correct	58.45	58.47	58.47	58.47	58.47	58.47	58.47
% Defaults Correct	56.86	56.85	56.86	56.85	56.85	56.85	56.85

* N=2,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.7 Used Cars Data - Monte Carlo Prediction 7 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-1.3113	-2.2386	-2.2386	-0.4570	-2.0940	-2.2370
	(-30.64)	(-27.04)	(-27.04)	(-47.65)	(-29.66)	(-27.07)
Cdcmi	-.00004	-.00008	-.00008	-.00001	-.00007	-.00008
	(-3.62)	(-3.65)	(-3.65)	(-3.90)	(-3.70)	(-3.65)
Cdfpay	.00162	.00292	.00292	.00041	.00259	.00291
	(5.57)	(5.42)	(5.42)	(5.51)	(5.43)	(5.42)
Cdtdp	-.00205	-.00333	-.00333	-.00043	-.00290	-.00332
	(-2.75)	(-2.44)	(-2.44)	(-2.60)	(-2.47)	(-2.44)
Age	-.00390	-.00777	-.00777	-.00091	-.00666	-.00775
	(-4.26)	(-4.30)	(-4.30)	(-4.50)	(-4.35)	(-4.30)
Amtfin	.00001	.00003	.00003	.00000	.00002	.00003
	(1.75)	(1.88)	(1.88)	(1.68)	(1.84)	(1.88)
Netrate	.00254	.00564	.00564	.00047	.00457	.00563
	(1.73)	(1.95)	(1.95)	(1.47)	(1.86)	(1.95)
Home	-.19763	-.37655	-.37655	-.04576	-.32467	-.37591
	(-5.65)	(-5.59)	(-5.59)	(-6.05)	(-5.70)	(-5.59)
Cosign	-.08458	-.15696	-.15696	-.02053	-.13734	-.15672
	(-1.83)	(-1.79)	(-1.79)	(-1.99)	(-1.83)	(-1.79)
Estimation Sample						
Percent Correct	80.91	80.97	80.97	80.90	80.96	80.97
Standard Deviation	.0108	.0105	.0105	.0126	.0109	.0105
% Repay. Correct	89.61	89.67	89.67	89.58	89.65	89.67
% Defaults Correct	19.98	20.09	20.09	20.08	20.09	20.09
Hold-out Sample						
Percent Correct	80.90	80.96	80.96	80.89	80.94	80.96
Standard Deviation	.0118	.0115	.0115	.0134	.0118	.0115
% Repay. Correct	89.61	89.65	89.65	89.58	89.64	89.65
% Defaults Correct	19.77	19.85	19.85	19.82	19.85	19.85

* N=20,000; unbalanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.8 Used Cars Data - Monte Carlo Prediction 8 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-.20738	-.33800	-.33800	-.08870	-.32815	-.33789
	(-6.39)	(-6.52)	(-6.52)	(-6.48)	(-6.52)	(-6.52)
Cdcmi	-.00004	-.00007	-.00007	-.00002	-.00007	-.00007
	(-5.69)	(-5.70)	(-5.70)	(-5.76)	(-5.71)	(-5.70)
Cdfpay	.00203	.00336	.00336	.00086	.00325	.00336
	(8.15)	(9.00)	(9.00)	(8.80)	(8.99)	(9.00)
Cdtdp	-.00308	-.00497	-.00497	-.00130	-.00483	-.00497
	(-4.40)	(-4.43)	(-4.43)	(-4.42)	(-4.43)	(-4.43)
Age	-.00426	-.00687	-.00687	-.00182	-.00667	-.00687
	(-6.95)	(-6.90)	(-6.90)	(-6.96)	(-6.90)	(-6.90)
Amtfin	.00001	.00002	.00002	.00001	.00002	.00002
	(1.97)	(1.95)	(1.95)	(2.17)	(1.97)	(1.95)
Netrate	.00264	.00401	.00401	.00117	.00392	.00401
	(2.18)	(2.09)	(2.09)	(2.28)	(2.10)	(2.09)
Home	-.22830	-.36673	-.36673	-.09711	-.35625	-.36663
	(-9.23)	(-9.20)	(-9.20)	(-9.27)	(-9.21)	(-9.20)
Cosign	-.09545	-.15278	-.15278	-.04049	-.14843	-.15273
	(-2.46)	(-2.46)	(-2.46)	(-2.47)	(-2.46)	(-2.46)
Estimation Sample						
Percent Correct	58.11	58.12	58.12	58.11	58.12	58.12
Standard Deviation	.0035	.0035	.0035	.0035	.0035	.0035
% Repay. Correct	59.36	59.37	59.37	59.39	59.37	59.37
% Defaults Correct	56.86	56.86	56.86	56.83	56.86	56.86
Hold-out Sample						
Percent Correct	58.06	58.07	58.07	58.06	58.07	58.07
Standard Deviation	.0071	.0070	.0070	.0071	.0071	.0070
% Repay. Correct	59.33	59.34	59.34	59.37	59.34	59.34
% Defaults Correct	56.78	56.79	56.79	56.75	56.79	56.79

* N=20,000; balanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.9 Used Cars Data - Monte Carlo Prediction 9 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-1.3138	-2.2431	-2.2431	-0.4572	-2.0974	-2.2415
	(-16.09)	(-14.21)	(-14.21)	(-25.05)	(-15.60)	(-14.23)
Cdcmi	-.00004	-.00008	-.00008	-.00001	-.00007	-.00008
	(-1.98)	(-1.98)	(-1.98)	(-2.10)	(-2.01)	(-1.98)
Cdfpay	.00160	.00288	.00288	.00040	.00255	.00288
	(3.10)	(3.01)	(3.01)	(3.01)	(3.02)	(3.01)
Cdtdp	-.00223	-.00368	-.00368	-.00047	-.00321	-.00368
	(-1.47)	(-1.36)	(-1.36)	(-1.43)	(-1.37)	(-1.36)
Age	-.00401	-.00800	-.00800	-.00093	-.00684	-.00799
	(-2.36)	(-2.39)	(-2.39)	(-2.46)	(-2.41)	(-2.39)
Amtfin	.00002	.00003	.00003	.00000	.00003	.00003
	(1.07)	(1.11)	(1.11)	(1.05)	(1.10)	(1.11)
Netrate	.00306	.00656	.00656	.00061	.00538	.00655
	(1.20)	(1.29)	(1.29)	(1.08)	(1.25)	(1.29)
Home	-.20239	-.38521	-.38521	-.04664	-.33153	-.38454
	(-3.12)	(-3.07)	(-3.07)	(-3.34)	(-3.14)	(-3.07)
Cosign	-.09794	-.18394	-.18394	-.02283	-.15881	-.18363
	(-1.17)	(-1.15)	(-1.15)	(-1.25)	(-1.17)	(-1.15)
Estimation Sample						
Percent Correct	80.45	80.54	80.54	80.41	80.51	80.53
Standard Deviation	.0216	.0210	.0210	.0247	.0217	.0210
% Repay. Correct	88.87	88.96	88.96	88.79	88.93	88.96
% Defaults Correct	21.43	21.48	21.48	21.62	21.51	21.49
Hold-out Sample						
Percent Correct	80.30	80.39	80.39	80.25	80.36	80.39
Standard Deviation	.0213	.0208	.0208	.0247	.0214	.0208
% Repay. Correct	88.79	88.88	88.88	88.70	88.85	88.88
% Defaults Correct	20.82	20.87	20.87	20.98	20.88	20.87

* N=5,000; unbalanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.10 Used Cars Data - Monte Carlo Prediction 10 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-.20582	-.33492	-.33492	-.08795	-.32516	-.33481
	(-3.29)	(-3.33)	(-3.33)	(-3.33)	(-3.33)	(-3.33)
Cdcmi	-.00005	-.00008	-.00008	-.00002	-.00007	-.00008
	(-2.81)	(-2.81)	(-2.81)	(-2.84)	(-2.82)	(-2.81)
Cdfpay	.00197	.00324	.00324	.00083	.00314	.00324
	(4.30)	(4.54)	(4.54)	(4.46)	(4.54)	(4.54)
Cdtdp	-.00313	-.00505	-.00505	-.00132	-.00490	-.00505
	(-2.20)	(-2.21)	(-2.21)	(-2.20)	(-2.21)	(-2.21)
Age	-.00429	-.00693	-.00693	-.00183	-.00673	-.00693
	(-3.14)	(-3.11)	(-3.11)	(-3.15)	(-3.12)	(-3.11)
Amtfin	.00002	.00003	.00003	.00001	.00003	.00003
	(1.32)	(1.31)	(1.31)	(1.38)	(1.31)	(1.31)
Netrate	.00273	.00422	.00422	.00120	.00412	.00422
	(1.33)	(1.29)	(1.29)	(1.36)	(1.29)	(1.29)
Home	-.22864	-.36761	-.36761	-.09708	-.35701	-.36750
	(-4.55)	(-4.55)	(-4.55)	(-4.57)	(-4.55)	(-4.55)
Cosign	-.09932	-.15902	-.15902	-.04207	-.15446	-.15897
	(-1.41)	(-1.41)	(-1.41)	(-1.42)	(-1.41)	(-1.41)
Estimation Sample						
Percent Correct	58.23	58.24	58.24	58.24	58.24	58.24
Standard Deviation	.0069	.0070	.0070	.0069	.0069	.0070
% Repay. Correct	59.48	59.48	59.48	59.51	59.49	59.48
% Defaults Correct	56.99	56.99	56.99	56.97	56.99	56.99
Hold-out Sample						
Percent Correct	57.93	57.94	57.94	57.93	57.94	57.94
Standard Deviation	.0069	.0069	.0069	.0070	.0069	.0069
% Repay. Correct	59.13	59.14	59.14	59.16	59.14	59.14
% Defaults Correct	56.73	56.74	56.74	56.70	56.73	56.74

* N=5,000; balanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.11 Used Cars Data - Monte Carlo Prediction 11 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-1.3186	-2.2537	-2.2537	-0.4581	-2.1058	-2.2520
	(-10.72)	(-9.42)	(-9.42)	(-16.74)	(-10.39)	(-9.44)
Cdcmi	-.00005	-.00009	-.00009	-.00001	-.00007	-.00009
	(-1.43)	(-1.43)	(-1.43)	(-1.53)	(-1.45)	(-1.43)
Cdfpay	.00153	.00276	.00276	.00038	.00244	.00276
	(2.11)	(2.07)	(2.07)	(2.02)	(2.06)	(2.07)
Cdtdp	-.00217	-.00360	-.00360	-.00046	-.00313	-.00359
	(-1.10)	(-1.03)	(-1.03)	(-1.09)	(-1.04)	(-1.03)
Age	-.00401	-.00802	-.00802	-.00092	-.00682	-.00800
	(-1.51)	(-1.53)	(-1.53)	(-1.59)	(-1.55)	(-1.53)
Amtfin	.00002	.00004	.00004	.00000	.00003	.00004
	(1.00)	(1.04)	(1.04)	(0.99)	(1.02)	(1.03)
Netrate	.00352	.00743	.00743	.00072	.00613	.00741
	(0.97)	(1.03)	(1.03)	(0.90)	(1.01)	(1.03)
Home	-.20385	-.38935	-.38935	-.04654	-.33352	-.38865
	(-1.98)	(-1.95)	(-1.95)	(-2.12)	(-2.00)	(-1.95)
Cosign	-.10948	-.20686	-.20686	-.02452	-.17606	-.20646
	(-0.92)	(-0.91)	(-0.91)	(-1.01)	(-0.93)	(-0.91)
Estimation Sample						
Percent Correct	80.00	80.09	80.09	79.92	80.07	80.09
Standard Deviation	.0315	.0308	.0308	.0365	.0316	.0308
% Repay. Correct	88.09	88.18	88.18	87.94	88.15	88.18
% Defaults Correct	23.13	23.23	23.23	23.42	23.25	23.22
Hold-out Sample						
Percent Correct	79.65	79.72	79.72	79.54	79.69	79.72
Standard Deviation	.0308	.0299	.0299	.0360	.0310	.0299
% Repay. Correct	87.92	88.00	88.00	87.77	87.97	88.00
% Defaults Correct	21.81	21.84	21.84	22.02	21.86	21.84

* N=2,000; unbalanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.12 Used Cars Data - Monte Carlo Prediction 12 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-.21254	-.34443	-.34443	-.09048	-.33438	-.34432
	(-2.18)	(-2.18)	(-2.18)	(-2.19)	(-2.18)	(-2.18)
Cdcmi	-.00005	-.00008	-.00008	-.00002	-.00007	-.00008
	(-1.95)	(-1.96)	(-1.96)	(-1.98)	(-1.96)	(-1.96)
Cdfpay	.00194	.00317	.00317	.00082	.00307	.00317
	(3.05)	(3.10)	(3.10)	(3.11)	(3.10)	(3.10)
Cdtdp	-.00310	-.00500	-.00500	-.00130	-.00485	-.00499
	(-1.51)	(-1.51)	(-1.51)	(-1.51)	(-1.51)	(-1.51)
Age	-.00424	-.00685	-.00685	-.00180	-.00665	-.00685
	(-2.17)	(-2.15)	(-2.15)	(-2.18)	(-2.15)	(-2.15)
Amtfin	.00002	.00003	.00003	.00001	.00003	.00003
	(1.01)	(1.01)	(1.01)	(1.05)	(1.01)	(1.01)
Netrate	.00325	.00511	.00511	.00142	.00498	.00510
	(1.10)	(1.08)	(1.08)	(1.12)	(1.08)	(1.08)
Home	-.23435	-.37715	-.37715	-.09930	-.36617	-.37704
	(-2.97)	(-2.96)	(-2.96)	(-2.98)	(-2.96)	(-2.96)
Cosign	-.11369	-.18211	-.18211	-.04795	-.17681	-.18206
	(-1.14)	(-1.14)	(-1.14)	(-1.15)	(-1.14)	(-1.14)
Estimation Sample						
Percent Correct	58.36	58.37	58.37	58.37	58.37	58.37
Standard Deviation	.0108	.0108	.0108	.0107	.0108	.0108
% Repay. Correct	59.11	59.14	59.14	59.14	59.15	59.14
% Defaults Correct	57.60	57.60	57.60	57.61	57.60	57.60
Hold-out Sample						
Percent Correct	57.73	57.74	57.74	57.74	57.73	57.74
Standard Deviation	.0073	.0073	.0073	.0073	.0073	.0073
% Repay. Correct	58.51	58.52	58.52	58.52	58.52	58.52
% Defaults Correct	56.96	56.95	56.95	56.95	56.95	56.95

* N=2,000; balanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.13 Used Cars Data – Monte Carlo MSE 11 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.41	79.56	79.07	79.56	79.15	79.49	79.56
Repayments	87.75	87.96	87.25	87.96	87.37	87.87	87.96
Defaults	22.06	21.76	22.78	21.76	22.58	21.90	21.76
Hold-Out	79.61	79.76	79.25	79.76	79.33	79.69	79.76
Repayments	88.04	88.26	87.52	88.26	87.64	88.16	88.26
Defaults	21.29	20.97	22.05	20.97	21.88	21.12	20.97
YRMSE	141.44	184.77	185.56	184.77	177.93	173.59	184.64
$YRMSE_0$	70.71	92.43	92.81	92.43	89.00	86.82	92.36
PRMSE	1.02	1.06	1.32	1.06	1.36	1.02	1.06
$PRMSE_0$	0.51	0.53	0.67	0.53	0.67	0.51	0.53
$MSE(\hat{\beta})$	.00518	1.05177	1.14663	1.05177	.84861	.72540	1.04764
$RMSE(\hat{\beta})$	.07195	1.02556	1.07081	1.02556	.92120	.85170	1.02354
$VAR(\hat{\beta})$	.00518	.01895	.01592	.01895	.00025	.01373	.01888
$Bias^2(\hat{\beta})$	.00000	1.03282	1.13071	1.03282	.84836	.71167	1.02876
MSE ($\partial p / \partial x$)	.00023	.00086	.00063	.00086	.03316	.00072	.00086
RMSE ($\partial p / \partial x$)	.01525	.02936	.02511	.02936	.18209	.02684	.02932

* N=20,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table C.14 Used Cars Data – Monte Carlo MSE 12 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	89.29	89.21	88.15	89.21	89.69	89.35	89.21
Repayments	95.56	95.46	94.14	95.46	96.07	95.63	95.46
Defaults	10.46	10.65	12.88	10.65	9.47	10.33	10.64
Hold-Out	89.50	89.43	88.37	89.43	89.84	89.55	89.43
Repayments	95.80	95.71	94.38	95.71	96.23	95.86	95.71
Defaults	9.60	9.77	12.15	9.77	8.79	9.49	9.77
YRMSE	148.32	248.97	250.66	248.97	173.21	221.34	248.61
$YRMSE_0$	74.08	124.55	125.37	124.55	86.52	110.68	124.37
PRMSE	1.02	0.93	1.17	0.93	1.63	1.05	0.93
$PRMSE_0$	0.53	0.47	0.59	0.47	0.82	0.54	0.47
$MSE(\hat{\beta})$	.09771	2.70755	2.97512	2.70755	.77347	1.73941	2.69359
$RMSE(\hat{\beta})$	.31259	1.64546	1.72485	1.64546	.87947	1.31887	1.64122
$VAR(\hat{\beta})$	.00717	.03110	.02593	.03110	.00016	.01702	.03085
$Bias^2(\hat{\beta})$	.09054	2.67645	2.94918	2.67645	.77330	1.72239	2.66274
MSE ($\partial p / \partial x$)	.00352	.00794	.00734	.00794	.03024	.00556	.00791
RMSE ($\partial p / \partial x$)	.05929	.08910	.08569	.08910	.17391	.07454	.08891

* N=20,000; unbalanced; unrestricted; errors drawn from t- distribution.

Table C.15 Used Cars Data – Monte Carlo MSE 13 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	83.47	83.44	82.87	83.44	83.69	83.49	83.44
Repayments	93.96	93.91	93.09	93.91	94.28	93.98	93.91
Defaults	10.45	10.54	11.72	10.54	9.99	10.43	10.54
Hold-Out	83.61	83.58	83.00	83.58	83.85	83.63	83.58
Repayments	94.13	94.08	93.24	94.08	94.47	94.16	94.08
Defaults	9.87	9.96	11.17	9.96	9.39	9.83	9.96
YRMSE	141.72	181.98	182.29	181.98	178.11	172.01	181.86
$YRMSE_0$	70.84	91.02	91.18	91.02	89.06	86.02	90.96
PRMSE	1.06	1.03	1.12	1.03	1.31	1.05	1.03
$PRMSE_0$	0.52	0.51	0.56	0.51	0.61	0.51	0.51
$MSE(\hat{\beta})$	.01287	.79170	.84622	.79170	.87912	.54670	.78864
$RMSE(\hat{\beta})$	.11344	.88977	.91990	.88977	.93761	.73939	.88805
$VAR(\hat{\beta})$	.00474	.01712	.01534	.01712	.00025	.01266	.01706
$Bias^2(\hat{\beta})$	.00813	.77457	.83089	.77457	.87888	.53404	.77157
MSE ($\partial p / \partial x$)	.00066	.00202	.00178	.00202	.03465	.00181	.00202
RMSE ($\partial p / \partial x$)	.02560	.04499	.04220	.04499	.18614	.04259	.04496

* N=20,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table C.16 Used Cars Data – Monte Carlo MSE 21 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	58.22	58.22	58.21	58.22	58.21	58.22	58.22
Repayments	59.02	59.02	59.03	59.02	59.03	59.02	59.02
Defaults	57.39	57.40	57.38	57.40	57.38	57.40	57.40
Hold-Out	58.15	58.15	58.15	58.15	58.15	58.15	58.15
Repayments	58.58	58.57	58.61	58.57	58.61	58.58	58.57
Defaults	57.72	57.73	57.69	57.73	57.69	57.72	57.72
YRMSE	141.42	143.27	143.18	143.27	143.01	143.00	143.27
$YRMSE_0$	70.73	71.68	71.63	71.68	71.50	71.54	71.68
PRMSE	1.46	1.46	1.47	1.46	1.60	1.46	1.46
$PRMSE_0$	0.73	0.73	0.73	0.73	0.76	0.73	0.73
$MSE(\hat{\beta})$	.00316	.06463	.06306	.06463	.05103	.05574	.06453
$RMSE(\hat{\beta})$	.05624	.25422	.25112	.25422	.22590	.23608	.25404
$VAR(\hat{\beta})$	.00316	.00819	.00816	.00819	.00056	.00771	.00819
$Bias^2(\hat{\beta})$	.00000	.05644	.05490	.05644	.05047	.04802	.05635
MSE ($\partial p / \partial x$)	.00050	.00051	.00051	.00051	.01300	.00049	.00051
RMSE ($\partial p / \partial x$)	.02241	.02264	.02261	.02264	.11404	.02216	.02263

* N=20,000; balanced; unrestricted; errors drawn from normal distribution.

Table C.17 Used Cars Data – Monte Carlo MSE 22 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	62.40	62.40	62.39	62.40	62.39	62.40	62.40
Repayments	63.26	63.25	63.26	63.25	63.27	63.25	63.25
Defaults	61.53	61.54	61.51	61.54	61.51	61.53	61.54
Hold-Out	62.42	62.41	62.41	62.41	62.42	62.42	62.41
Repayments	62.99	62.98	63.03	62.98	63.03	62.98	62.98
Defaults	61.86	61.86	61.80	61.86	61.82	61.86	61.86
YRMSE	142.64	151.76	151.31	151.76	141.92	150.64	151.75
$YRMSE_0$	72.05	76.60	76.37	76.60	71.66	76.04	76.59
PRMSE	1.51	1.45	1.62	1.45	2.45	1.48	1.45
$PRMSE_0$	0.75	0.72	0.81	0.72	1.13	0.74	0.72
$MSE(\hat{\beta})$	.04712	.34878	.34491	.34878	.02129	.31142	.34838
$RMSE(\hat{\beta})$	.21707	.59057	.58729	.59057	.14591	.55805	.59023
$VAR(\hat{\beta})$	.00307	.00811	.00806	.00811	.00050	.00756	.00810
$Bias^2(\hat{\beta})$	.04405	.34067	.33685	.34067	.02079	.30386	.34027
MSE ($\partial p / \partial x$)	.00742	.00812	.00799	.00812	.00882	.00675	.00811
RMSE ($\partial p / \partial x$)	.08611	.09012	.08938	.09012	.09394	.08218	.09004

* N=20,000; balanced; unrestricted; errors drawn from t- distribution.

Table C.18 Used Cars Data – Monte Carlo MSE 23 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	60.64	60.64	60.64	60.64	60.64	60.64	60.64
Repayments	82.95	82.92	83.03	82.92	83.28	82.95	82.92
Defaults	30.28	30.33	30.17	30.33	29.84	30.28	30.33
Hold-Out	60.55	60.55	60.55	60.55	60.55	60.55	60.55
Repayments	82.09	82.04	82.20	82.04	82.45	82.08	82.04
Defaults	31.52	31.57	31.36	31.57	31.02	31.52	31.57
YRMSE	144.21	150.49	150.56	150.49	143.52	149.79	150.48
$YRMSE_0$	72.14	75.22	75.26	75.22	71.79	74.88	75.22
PRMSE	4.33	4.30	4.28	4.30	4.77	4.33	4.30
$PRMSE_0$	2.13	2.11	2.10	2.11	2.24	2.12	2.11
$MSE(\hat{\beta})$	.04824	.30345	.30808	.30345	.02975	.27417	.30314
$RMSE(\hat{\beta})$	.21963	.55086	.55505	.55086	.17249	.52361	.55058
$VAR(\hat{\beta})$	.00310	.00818	.00794	.00818	.00052	.00765	.00817
$Bias^2(\hat{\beta})$	.04514	.29527	.30014	.29527	.02923	.26652	.29497
MSE ($\partial p / \partial x$)	.00731	.00753	.00771	.00753	.00982	.00656	.00752
RMSE ($\partial p / \partial x$)	.08553	.08679	.08782	.08679	.09910	.08102	.08673

* N=20,000; balanced; unrestricted; errors drawn from chi-square distribution.

Table C.19 Used Cars Data – Monte Carlo MSE 31 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	80.02	80.13	80.29	80.13	80.34	80.18	80.13
Repayments	88.18	88.33	88.57	88.33	88.63	88.39	88.33
Defaults	22.06	21.87	21.52	21.87	21.37	21.80	21.87
Hold-Out	79.62	79.71	79.83	79.71	79.87	79.74	79.71
Repayments	87.64	87.78	87.94	87.78	88.01	87.82	87.78
Defaults	22.38	22.16	21.93	22.16	21.76	22.09	22.16
YRMSE	70.76	94.16	93.79	94.16	89.95	87.71	94.08
$YRMSE_0$	70.87	94.47	94.04	94.47	90.13	87.93	94.39
PRMSE	0.98	1.01	1.08	1.01	1.12	0.98	1.01
$PRMSE_0$	0.99	1.01	1.08	1.01	1.14	0.99	1.01
$MSE(\hat{\beta})$	.02187	.94587	1.06203	.94587	.71536	.66256	.94223
$RMSE(\hat{\beta})$	.14789	.97256	1.03055	.97256	.84579	.81398	.97069
$VAR(\hat{\beta})$	.02183	.08182	.05859	.08182	.00091	.05600	.08145
$Bias^2(\hat{\beta})$	.00004	.86404	1.00344	.86404	.71445	.60656	.86078
MSE ($\partial p / \partial x$)	.00102	.00179	.00109	.00179	.02773	.00133	.00178
RMSE ($\partial p / \partial x$)	.03193	.04232	.03301	.04232	.16653	.03653	.04223

* N=5,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table C. 20 Used Cars Data – Monte Carlo MSE 32 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	89.46	89.35	89.03	89.35	90.28	89.61	89.36
Repayments	95.57	95.42	95.03	95.42	96.60	95.75	95.42
Defaults	10.80	11.15	11.71	11.15	8.76	10.53	11.14
Hold-Out	89.03	88.88	88.57	88.88	90.03	89.20	88.88
Repayments	95.09	94.89	94.51	94.89	96.35	95.30	94.89
Defaults	10.83	11.23	11.81	11.23	8.44	10.43	11.21
YRMSE	73.81	125.94	125.70	125.94	87.32	110.79	125.73
$YRMSE_0$	75.35	127.04	126.74	127.04	88.77	111.87	126.83
PRMSE	0.84	0.82	0.89	0.82	1.03	0.85	0.82
$PRMSE_0$	0.84	0.82	0.88	0.82	1.03	0.85	0.82
$MSE(\hat{\beta})$	.13291	2.54550	2.81566	2.54550	.64026	1.69161	2.53308
$RMSE(\hat{\beta})$	.36457	1.59546	1.67799	1.59546	.80017	1.30062	1.59156
$VAR(\hat{\beta})$	.03082	.13763	.09384	.13763	.00058	.06833	.13629
$Bias^2(\hat{\beta})$	.10209	2.40787	2.72182	2.40787	.63968	1.62328	2.39679
MSE ($\partial p / \partial x$)	.00327	.00701	.00622	.00701	.02478	.00473	.00697
RMSE ($\partial p / \partial x$)	.05714	.08370	.07888	.08370	.15741	.06877	.08349

* N=5,000; unbalanced; unrestricted; errors drawn from *t*- distribution.

Table C.21 Used Cars Data – Monte Carlo MSE 33 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	83.85	83.76	83.72	83.76	84.40	83.89	83.76
Repayments	94.07	93.94	93.89	93.94	94.88	94.14	93.94
Defaults	10.58	10.80	10.88	10.80	9.30	10.50	10.80
Hold-Out	83.31	83.21	83.17	83.21	83.93	83.35	83.21
Repayments	93.46	93.31	93.26	93.31	94.36	93.52	93.31
Defaults	10.84	11.07	11.12	11.07	9.47	10.75	11.06
YRMSE	70.87	91.81	91.70	91.81	90.09	86.36	91.75
$YRMSE_0$	71.13	92.21	92.06	92.15	90.29	86.70	92.14
PRMSE	0.98	0.99	1.01	0.99	1.04	0.98	0.99
$PRMSE_0$	0.99	1.00	1.01	1.00	1.05	0.99	1.00
$MSE(\hat{\beta})$	.03092	.80365	.86788	.80365	.73495	.57536	.80079
$RMSE(\hat{\beta})$	.17583	.89646	.93160	.89646	.85729	.75852	.89487
$VAR(\hat{\beta})$	.02157	.07997	.06399	.07997	.00099	.05676	.07964
$Bias^2(\hat{\beta})$	.00935	.72368	.80389	.72368	.73396	.51859	.72115
MSE ($\partial p / \partial x$)	.00143	.00246	.00211	.00246	.02871	.00215	.00245
RMSE ($\partial p / \partial x$)	.03785	.04957	.04596	.04957	.16944	.04642	.04953

* N=5,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table C.22 Used Cars Data – Monte Carlo MSE 41 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	56.70	56.70	56.70	56.70	56.70	56.70	56.70
Repayments	57.90	57.90	57.88	57.90	57.89	57.90	57.90
Defaults	55.42	55.42	55.44	55.42	55.43	55.42	55.42
Hold-Out	56.31	56.30	56.30	56.30	56.31	56.30	56.30
Repayments	59.06	59.06	59.05	59.06	59.06	59.06	59.06
Defaults	53.53	53.53	53.53	53.53	53.53	53.53	53.53
YRMSE	70.72	71.36	71.35	71.36	71.18	71.27	71.35
$YRMSE_0$	70.84	71.55	71.54	71.55	71.23	71.45	71.55
PRMSE	1.46	1.46	1.45	1.46	1.46	1.46	1.46
$PRMSE_0$	1.45	1.45	1.44	1.45	1.45	1.45	1.45
$MSE(\hat{\beta})$	.01215	.07708	.07638	.07708	.03747	.06893	.07700
$RMSE(\hat{\beta})$	.11024	.27764	.27637	.27764	.19358	.26255	.27748
$VAR(\hat{\beta})$	.01209	.03129	.03087	.03129	.00218	.02950	.03127
$Bias^2(\hat{\beta})$	.00006	.04579	.04551	.04579	.03530	.03943	.04573
MSE ($\partial p / \partial x$)	.00193	.00196	.00193	.00196	.00944	.00184	.00196
RMSE ($\partial p / \partial x$)	.04388	.04427	.04396	.04427	.09716	.04284	.04425

* N=5,000; balanced; unrestricted; errors drawn from normal distribution.

Table C.23 Used Cars Data – Monte Carlo MSE 42 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	60.05	60.05	60.05	60.05	60.05	60.05	60.05
Repayments	61.43	61.43	61.42	61.43	61.42	61.43	61.43
Defaults	58.61	58.62	58.63	58.62	58.63	58.62	58.62
Hold-Out	59.83	59.83	59.82	59.83	59.83	59.83	59.83
Repayments	62.67	62.66	62.67	62.66	62.67	62.66	62.66
Defaults	56.93	56.94	56.93	56.94	56.93	56.94	56.94
YRMSE	71.96	74.87	74.85	74.87	71.66	74.54	74.87
$YRMSE_0$	71.31	74.25	74.23	74.25	70.96	73.92	74.24
PRMSE	1.46	1.45	1.44	1.45	1.51	1.45	1.45
$PRMSE_0$	1.45	1.44	1.43	1.44	1.52	1.44	1.44
$MSE(\hat{\beta})$	.04748	.29009	.28856	.29009	.01579	.26236	.28980
$RMSE(\hat{\beta})$	.21790	.53860	.53718	.53860	.12567	.51221	.53833
$VAR(\hat{\beta})$	.01294	.03407	.03320	.03407	.00220	.03189	.03405
$Bias^2(\hat{\beta})$	.03454	.25602	.25537	.25602	.01359	.23047	.25575
MSE ($\partial p / \partial x$)	.00746	.00789	.00781	.00789	.00632	.00682	.00788
RMSE ($\partial p / \partial x$)	.08635	.08882	.08840	.08882	.07951	.08261	.08875

* N=5,000; balanced; unrestricted; errors drawn from t- distribution.

Table C.24 Used Cars Data – Monte Carlo MSE 43 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	59.59	59.60	59.60	59.60	59.60	59.60	59.60
Repayments	86.69	86.63	86.60	86.63	86.91	86.66	86.63
Defaults	22.19	22.28	22.33	22.28	21.89	22.25	22.28
Hold-Out	59.61	59.61	59.61	59.61	59.61	59.61	59.61
Repayments	87.55	87.49	87.45	87.49	87.74	87.52	87.49
Defaults	20.65	20.73	20.79	20.73	20.39	20.69	20.73
YRMSE	72.18	75.08	75.09	75.08	71.46	74.78	75.08
$YRMSE_0$	72.25	75.31	75.32	75.31	71.43	75.00	75.30
PRMSE	1.46	1.46	1.44	1.46	1.48	1.46	1.46
$PRMSE_0$	1.45	1.45	1.43	1.45	1.47	1.45	1.45
$MSE(\hat{\beta})$	.05622	.23517	.24010	.23517	.03579	.21567	.23496
$RMSE(\hat{\beta})$	.23712	.48494	.49000	.48494	.18919	.46440	.48473
$VAR(\hat{\beta})$	.01188	.03137	.02935	.03137	.00199	.02932	.03135
$Bias^2(\hat{\beta})$	.04434	.20380	.21075	.20380	.03380	.18635	.20362
MSE ($\partial p / \partial x$)	.00866	.00873	.00909	.00873	.00871	.00807	.00873
RMSE ($\partial p / \partial x$)	.09304	.09346	.09536	.09346	.09333	.08982	.09342

* N=5,000; balanced; unrestricted; errors drawn from chi-square distribution.

Table C.25 Used Cars Data - Monte Carlo MSE 51 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	70.94	71.33	71.67	71.33	69.66	71.04	71.33
Repayments	76.02	76.59	77.12	76.59	74.14	76.17	76.58
Defaults	40.32	39.65	38.89	39.65	42.54	40.14	39.65
Hold-Out	70.73	71.14	71.47	71.14	69.34	70.84	71.14
Repayments	76.35	76.96	77.42	76.96	74.31	76.50	76.95
Defaults	37.24	36.48	36.02	36.48	39.71	37.06	36.48
YRMSE	44.79	57.62	57.51	57.62	54.99	54.31	57.58
$YRMSE_0$	71.18	91.77	91.40	91.77	87.08	86.44	91.71
PRMSE	1.04	1.05	1.10	1.05	1.06	1.03	1.05
$PRMSE_0$	1.71	1.74	1.82	1.74	1.75	1.70	1.74
$MSE(\hat{\beta})$	.05175	1.18747	1.23847	1.18747	.81453	.83708	1.18303
$RMSE(\hat{\beta})$	.22748	1.08971	1.11287	1.08971	.90251	.91492	1.08767
$VAR(\hat{\beta})$	.05159	.17861	.16025	.17861	.00304	.13507	.17803
$Bias^2(\hat{\beta})$	.00016	1.00886	1.07823	1.00886	.81148	.70201	1.00500
MSE ($\partial p / \partial x$)	.00278	.00335	.00284	.00335	.04284	.00299	.00335
RMSE ($\partial p / \partial x$)	.05276	.05789	.05328	.05789	.20697	.05470	.05785

*N=2,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table C.26 Used Cars Data - Monte Carlo MSE 52 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	81.40	81.66	81.52	81.66	80.91	81.46	81.66
Repayments	86.29	86.62	86.46	86.62	85.67	86.37	86.62
Defaults	28.15	27.81	27.84	27.81	28.88	28.12	27.82
Hold-Out	81.40	81.68	81.37	81.68	80.73	81.44	81.67
Repayments	86.66	87.00	86.61	87.00	85.82	86.71	87.00
Defaults	24.70	24.28	24.86	24.28	25.79	24.67	24.29
YRMSE	47.21	78.16	78.15	78.16	53.43	69.59	78.05
$YRMSE_0$	75.02	124.23	123.91	124.23	84.57	110.47	124.04
PRMSE	0.89	0.89	0.98	0.89	0.95	0.88	0.89
$PRMSE_0$	1.49	1.50	1.66	1.50	1.61	1.48	1.50
$MSE(\hat{\beta})$	.18802	3.09602	3.25499	3.09602	.73792	1.98593	3.07943
$RMSE(\hat{\beta})$	.43362	1.75955	1.80416	1.75955	.85903	1.40923	1.75483
$VAR(\hat{\beta})$	.07841	.32656	.27766	.32656	.00218	.18594	.32393
$Bias^2(\hat{\beta})$	.10962	2.76946	2.97733	2.76946	.73575	1.79998	2.75550
MSE ($\partial p / \partial x$)	.00517	.00926	.00870	.00926	.03935	.00718	.00922
RMSE ($\partial p / \partial x$)	.07189	.09621	.09328	.09621	.19837	.08475	.09604

*N=2,000; unbalanced; unrestricted; errors drawn from *t*- distribution.

Table C.27 Used Cars Data - Monte Carlo MSE 53 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	75.69	75.86	76.01	75.86	75.42	75.78	75.86
Repayments	83.41	83.64	83.87	83.64	83.01	83.53	83.64
Defaults	26.84	26.63	26.27	26.63	27.35	26.74	26.63
Hold-Out	75.42	75.59	75.67	75.59	75.05	75.50	75.59
Repayments	83.61	83.85	83.96	83.85	83.09	83.71	83.85
Defaults	23.83	23.55	23.45	23.55	24.46	23.71	23.55
YRMSE	44.88	57.60	57.52	57.60	54.95	54.51	57.56
$YRMSE_0$	71.38	91.81	91.54	91.81	87.00	86.83	91.75
PRMSE	1.04	1.04	1.05	1.04	1.04	1.03	1.04
$PRMSE_0$	1.71	1.73	1.76	1.73	1.72	1.71	1.73
$MSE(\hat{\beta})$	.06696	.95354	.98206	.95354	.84966	.67403	.95001
$RMSE(\hat{\beta})$	.25877	.97650	.99099	.97650	.92177	.82099	.97469
$VAR(\hat{\beta})$	.05679	.20074	.17854	.20074	.00324	.15043	.20006
$Bias^2(\hat{\beta})$	.01018	.75281	.80352	.75281	.84642	.52360	.74995
MSE ($\partial p / \partial x$)	.00387	.00551	.00503	.00551	.04488	.00514	.00551
RMSE ($\partial p / \partial x$)	.06222	.07425	.07092	.07425	.21185	.07171	.07421

*N=2,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table C.28 Used Cars Data - Monte Carlo MSE 61 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	57.98	57.99	57.99	57.99	57.99	57.99	57.99
Repayments	60.10	60.08	60.19	60.08	60.18	60.09	60.08
Defaults	55.70	55.72	55.61	55.72	55.62	55.71	55.72
Hold-Out	57.23	57.23	57.23	57.23	57.23	57.23	57.23
Repayments	60.42	60.40	60.53	60.40	60.52	60.41	60.40
Defaults	54.01	54.03	53.90	54.03	53.91	54.02	54.03
YRMSE	44.68	45.32	45.31	45.32	45.09	45.23	45.32
$YRMSE_0$	70.99	72.26	72.24	72.26	71.44	72.10	72.26
PRMSE	1.48	1.48	1.47	1.48	1.48	1.48	1.48
$PRMSE_0$	2.40	2.40	2.38	2.40	2.41	2.40	2.40
$MSE(\hat{\beta})$	.03787	.13071	.13044	.13071	.03226	.12035	.13060
$RMSE(\hat{\beta})$	.19461	.36155	.36116	.36155	.17962	.34691	.36139
$VAR(\hat{\beta})$	.03781	.09821	.09707	.09821	.00671	.09245	.09815
$Bias^2(\hat{\beta})$	.00006	.03250	.03337	.03250	.02555	.02789	.03245
MSE ($\partial p / \partial x$)	.00601	.00612	.00606	.00612	.00712	.00576	.00612
RMSE ($\partial p / \partial x$)	.07750	.07826	.07784	.07826	.08439	.07589	.07823

*N=2,000; balanced; unrestricted; errors drawn from normal distribution.

Table C.29 Used Cars Data - Monte Carlo MSE 62 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	61.81	61.82	61.81	61.82	61.81	61.82	61.82
Repayments	64.83	64.79	65.07	64.79	65.04	64.82	64.79
Defaults	58.65	58.72	58.42	58.72	58.46	58.70	58.72
Hold-Out	61.28	61.28	61.29	61.28	61.29	61.28	61.28
Repayments	65.35	65.28	65.58	65.28	65.55	65.31	65.28
Defaults	57.14	57.20	56.90	57.20	56.94	57.18	57.20
YRMSE	44.61	47.34	47.30	47.34	44.34	47.01	47.34
$YRMSE_0$	71.98	76.47	76.42	76.47	71.38	75.94	76.47
PRMSE	1.44	1.44	1.43	1.44	1.48	1.44	1.44
$PRMSE_0$	2.34	2.34	2.32	2.34	2.47	2.34	2.34
$MSE(\hat{\beta})$	.06304	.28622	.29334	.28622	.01627	.26045	.28594
$RMSE(\hat{\beta})$	.25109	.53499	.54161	.53499	.12755	.51034	.53474
$VAR(\hat{\beta})$	.03825	.10120	.09974	.10120	.00634	.09443	.10113
$Bias^2(\hat{\beta})$	.02479	.18502	.19360	.18502	.00993	.16601	.18481
MSE ($\partial p / \partial x$)	.00995	.01050	.01075	.01050	.00486	.00939	.01049
RMSE ($\partial p / \partial x$)	.09977	.10249	.10370	.10249	.06972	.09690	.10243

*N=2,000; balanced; unrestricted; errors drawn from *t*- distribution.

Table C.30 Used Cars Data - Monte Carlo MSE 63 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	60.76	60.76	60.76	60.76	60.76	60.76	60.76
Repayments	84.35	84.31	84.29	84.31	84.51	84.33	84.31
Defaults	28.43	28.49	28.54	28.49	28.22	28.46	28.48
Hold-Out	60.34	60.34	60.35	60.34	60.35	60.34	60.34
Repayments	83.80	83.76	83.71	83.76	83.95	83.78	83.76
Defaults	28.06	28.10	28.19	28.10	27.87	28.08	28.10
YRMSE	45.62	47.68	47.73	47.68	45.28	47.46	47.68
$YRMSE_0$	72.43	75.98	76.05	75.98	71.69	75.60	75.98
PRMSE	1.47	1.47	1.46	1.47	1.48	1.47	1.47
$PRMSE_0$	2.39	2.39	2.37	2.39	2.41	2.39	2.39
$MSE(\hat{\beta})$	.08636	.33030	.34512	.33030	.02296	.30469	.33003
$RMSE(\hat{\beta})$	.29387	.57471	.58747	.57471	.15152	.55199	.57448
$VAR(\hat{\beta})$	.03755	.09921	.09475	.09921	.00626	.09267	.09915
$Bias^2(\hat{\beta})$	.04881	.23108	.25037	.23108	.01670	.21202	.23088
MSE ($\partial p / \partial x$)	.01323	.01345	.01409	.01345	.00539	.01236	.01344
RMSE ($\partial p / \partial x$)	.11503	.11598	.11869	.11598	.07342	.11118	.11593

*N=2,000; balanced; unrestricted; errors drawn from chi-square distribution.

Table C.31 Used Cars Data - Monte Carlo MSE 71 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	79.41	79.56	79.56	79.15	79.49	79.56
Repayments	87.75	87.96	87.96	87.37	87.86	87.96
Defaults	22.06	21.76	21.76	22.59	21.90	21.76
Hold-Out	79.61	79.76	79.76	79.33	79.69	79.76
Repayments	88.04	88.26	88.26	87.64	88.15	88.26
Defaults	21.29	20.96	20.96	21.88	21.12	20.96
YRMSE	141.44	184.77	184.77	177.93	173.59	184.63
$YRMSE_0$	70.71	92.42	92.42	89.00	86.82	92.36
PRMSE	1.01	1.06	1.06	1.35	1.01	1.06
$PRMSE_0$	0.51	0.53	0.53	0.66	0.51	0.53
$MSE(\hat{\beta})$	.00511	1.05146	1.05146	.84861	.72516	1.04733
$RMSE(\hat{\beta})$	.07146	1.02541	1.02541	.92120	.85157	1.02339
$VAR(\hat{\beta})$	.00510	.01868	.01868	.00025	.01354	.01861
$Bias^2(\hat{\beta})$	.00000	1.03278	1.03278	.84836	.71162	1.02872
MSE ($\partial p / \partial x$)	.00023	.00086	.00086	.03316	.00072	.00086
RMSE ($\partial p / \partial x$)	.01514	.02930	.02930	.18209	.02679	.02926

*N=20,000; unbalanced; restricted; errors drawn from normal distribution.

Table C.32 Used Cars Data - Monte Carlo MSE 72 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	89.29	89.21	89.21	89.69	89.35	89.21
Repayments	95.56	95.46	95.46	96.07	95.63	95.46
Defaults	10.45	10.64	10.64	9.46	10.32	10.64
Hold-Out	89.50	89.43	89.43	89.85	89.55	89.43
Repayments	95.80	95.71	95.71	96.24	95.87	95.71
Defaults	9.60	9.78	9.78	8.76	9.50	9.78
YRMSE	148.32	248.97	248.97	173.21	221.34	248.60
$YRMSE_0$	74.07	124.55	124.55	86.52	110.68	124.37
PRMSE	1.01	0.92	0.92	1.62	1.04	0.92
$PRMSE_0$	0.53	0.47	0.47	0.81	0.54	0.47
$MSE(\hat{\beta})$	.09754	2.70693	2.70694	.77341	1.73914	2.69299
$RMSE(\hat{\beta})$	.31232	1.64528	1.64528	.87944	1.31876	1.64103
$VAR(\hat{\beta})$	.00702	.03046	.03046	.00016	.01667	.03022
$Bias^2(\hat{\beta})$	.09052	2.67647	2.67647	.77326	1.72247	2.66276
MSE ($\partial p / \partial x$)	.00351	.00793	.00793	.03024	.00555	.00790
RMSE ($\partial p / \partial x$)	.05924	.08907	.08907	.17390	.07450	.08888

*N=20,000; unbalanced; restricted; errors drawn from *t*- distribution.

Table C.33 Used Cars Data - Monte Carlo MSE 73 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	83.47	83.44	83.44	83.69	83.49	83.44
Repayments	93.96	93.91	93.91	94.27	93.99	93.91
Defaults	10.44	10.54	10.54	9.99	10.42	10.53
Hold-Out	83.62	83.58	83.58	83.85	83.63	83.59
Repayments	94.14	94.09	94.09	94.47	94.16	94.09
Defaults	9.87	9.96	9.96	9.39	9.83	9.96
YRMSE	141.72	181.97	181.97	178.10	172.01	181.85
$YRMSE_0$	70.84	91.02	91.02	89.06	86.02	90.96
PRMSE	1.03	1.01	1.01	1.29	1.03	1.01
$PRMSE_0$	0.50	0.49	0.49	0.60	0.50	0.49
$MSE(\hat{\beta})$	.01272	.79093	.79094	.87921	.54607	.78788
$RMSE(\hat{\beta})$	.11278	.88934	.88935	.93766	.73897	.88763
$VAR(\hat{\beta})$	.00460	.01663	.01663	.00024	.01230	.01658
$Bias^2(\hat{\beta})$	.00812	.77430	.77430	.87897	.53378	.77130
MSE ($\partial p / \partial x$)	.00065	.00202	.00202	.03465	.00181	.00201
RMSE ($\partial p / \partial x$)	.02546	.04490	.04490	.18615	.04252	.04487

*N=20,000; unbalanced; restricted; errors drawn from chi-square distribution.

Table C.34 Used Cars Data - Monte Carlo MSE 81 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	58.21	58.22	58.22	58.21	58.22	58.22
Repayments	59.02	59.02	59.02	59.03	59.02	59.02
Defaults	57.39	57.39	57.39	57.38	57.39	57.39
Hold-Out	58.15	58.15	58.15	58.15	58.15	58.15
Repayments	58.59	58.59	58.59	58.61	58.59	58.59
Defaults	57.71	57.72	57.72	57.69	57.71	57.72
YRMSE	141.42	143.27	143.27	143.01	142.99	143.27
$YRMSE_0$	70.73	71.68	71.68	71.50	71.54	71.68
PRMSE	1.43	1.44	1.44	1.59	1.43	1.44
$PRMSE_0$	0.72	0.72	0.72	0.75	0.72	0.72
$MSE(\hat{\beta})$	.00299	.06408	.06408	.05100	.05524	.06398
$RMSE(\hat{\beta})$	.05472	.25314	.25314	.22583	.23503	.25295
$VAR(\hat{\beta})$	.00299	.00775	.00775	.00053	.00730	.00775
$Bias^2(\hat{\beta})$	.00000	.05633	.05633	.05046	.04794	.05624
MSE ($\partial p / \partial x$)	.00048	.00048	.00048	.01300	.00047	.00048
RMSE ($\partial p / \partial x$)	.02180	.02202	.02202	.11403	.02157	.02201

*N=20,000; balanced; restricted; errors drawn from normal distribution.

Table C.35 Used Cars Data - Monte Carlo MSE 82 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	62.40	62.40	62.40	62.39	62.40	62.40
Repayments	63.26	63.25	63.25	63.27	63.25	63.25
Defaults	61.53	61.53	61.53	61.51	61.53	61.54
Hold-Out	62.42	62.42	62.42	62.42	62.42	62.42
Repayments	62.99	62.98	62.98	63.03	62.98	62.98
Defaults	61.86	61.86	61.86	61.82	61.86	61.86
YRMSE	142.64	151.76	151.76	141.92	150.63	151.75
$YRMSE_0$	72.05	76.60	76.60	71.66	76.04	76.59
PRMSE	1.50	1.44	1.44	2.44	1.47	1.44
$PRMSE_0$	0.75	0.72	0.72	1.13	0.73	0.72
$MSE(\hat{\beta})$	.04708	.34861	.34861	.02129	.31129	.34821
$RMSE(\hat{\beta})$	.21698	.59043	.59044	.14589	.55793	.59010
$VAR(\hat{\beta})$	.00304	.00804	.00804	.00050	.00749	.00803
$Bias^2(\hat{\beta})$	.04404	.34058	.34058	.02079	.30380	.34018
MSE ($\partial p / \partial x$)	.00741	.00811	.00811	.00882	.00675	.00810
RMSE ($\partial p / \partial x$)	.08607	.09008	.09008	.09394	.08214	.09000

*N=20,000; balanced; restricted; errors drawn from *t*- distribution.

Table C.36 Used Cars Data - Monte Carlo MSE 83 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	60.64	60.64	60.64	60.64	60.64	60.64
Repayments	82.94	82.91	82.91	83.27	82.94	82.91
Defaults	30.29	30.34	30.34	29.85	30.29	30.34
Hold-Out	60.55	60.55	60.55	60.55	60.55	60.55
Repayments	82.08	82.04	82.04	82.45	82.08	82.04
Defaults	31.53	31.58	31.58	31.03	31.53	31.58
YRMSE	144.21	150.48	150.48	143.52	149.79	150.48
$YRMSE_0$	72.14	75.22	75.22	71.79	74.88	75.22
PRMSE	4.32	4.28	4.28	4.76	4.32	4.28
$PRMSE_0$	2.12	2.11	2.11	2.24	2.12	2.11
$MSE(\hat{\beta})$	.04787	.30227	.30227	.02973	.27312	.30196
$RMSE(\hat{\beta})$	.21879	.54979	.54979	.17242	.52261	.54951
$VAR(\hat{\beta})$	.00296	.00783	.00783	.00050	.00732	.00782
$Bias^2(\hat{\beta})$	.04491	.29444	.29444	.02923	.26580	.29414
MSE ($\partial p / \partial x$)	.00726	.00747	.00747	.00982	.00651	.00746
RMSE ($\partial p / \partial x$)	.08520	.08645	.08645	.09910	.08070	.08639

***N=20,000; balanced; restricted; errors drawn from chi-square distribution.**

Table C.37 Used Cars Data - Monte Carlo MSE 91 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	80.00	80.10	80.10	80.32	80.15	80.11
Repayments	88.14	88.30	88.30	88.60	88.35	88.30
Defaults	22.08	21.89	21.89	21.39	21.82	21.89
Hold-Out	79.61	79.70	79.70	79.87	79.73	79.70
Repayments	87.61	87.76	87.76	88.00	87.80	87.76
Defaults	22.47	22.25	22.25	21.86	22.19	22.25
YRMSE	70.76	94.15	94.15	89.95	87.70	94.07
$YRMSE_0$	70.86	94.45	94.45	90.13	87.91	94.37
PRMSE	0.93	0.96	0.96	1.07	0.93	0.96
$PRMSE_0$	0.93	0.96	0.96	1.09	0.93	0.96
$MSE(\hat{\beta})$	.02112	.94822	.94822	.71501	.66387	.94457
$RMSE(\hat{\beta})$	.14531	.97376	.97376	.84558	.81478	.97189
$VAR(\hat{\beta})$	.02105	.07897	.07897	.00087	.05399	.07860
$Bias^2(\hat{\beta})$	.00006	.86925	.86925	.71414	.60988	.86596
MSE ($\partial p / \partial x$)	.00099	.00174	.00174	.02771	.00129	.00173
RMSE ($\partial p / \partial x$)	.03139	.04170	.04170	.16647	.03591	.04161

*N=5,000; unbalanced; restricted; errors drawn from normal distribution.

Table C.38 Used Cars Data - Monte Carlo MSE 92 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	89.49	89.36	89.36	90.33	89.63	89.37
Repayments	95.60	95.44	95.44	96.66	95.78	95.44
Defaults	10.73	11.08	11.08	8.65	10.45	11.07
Hold-Out	89.07	88.90	88.90	90.11	89.24	88.90
Repayments	95.13	94.91	94.91	96.44	95.34	94.92
Defaults	10.84	11.27	11.27	8.36	10.44	11.26
YRMSE	73.79	125.89	125.89	87.32	110.75	125.68
$YRMSE_0$	75.32	127.00	127.00	88.76	111.84	126.79
PRMSE	0.78	0.77	0.77	0.99	0.80	0.77
$PRMSE_0$	0.79	0.77	0.77	0.99	0.80	0.77
$MSE(\hat{\beta})$	.13058	2.55058	2.55058	.63971	1.69335	2.53811
$RMSE(\hat{\beta})$	.36136	1.59705	1.59705	.79982	1.30129	1.59314
$VAR(\hat{\beta})$	.02789	.12519	.12519	.00052	.06180	.12395
$Bias^2(\hat{\beta})$	.10269	2.42539	2.42539	.63919	1.63155	2.41416
MSE ($\partial p / \partial x$)	.00313	.00682	.00682	.02475	.00459	.00679
RMSE ($\partial p / \partial x$)	.05591	.08261	.08261	.15732	.06778	.08240

*N=5,000; unbalanced; restricted; errors drawn from *t*- distribution.

Table C.39 Used Cars Data - Monte Carlo MSE 93 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	83.88	83.79	83.79	84.45	83.92	83.79
Repayments	94.13	93.99	93.99	94.95	94.18	93.99
Defaults	10.47	10.71	10.71	9.17	10.41	10.71
Hold-Out	83.36	83.25	83.25	84.00	83.40	83.25
Repayments	93.51	93.35	93.35	94.45	93.57	93.35
Defaults	10.84	11.09	11.09	9.40	10.77	11.08
YRMSE	70.86	91.77	91.77	90.09	86.33	91.70
$YRMSE_0$	71.10	92.16	92.16	90.28	86.66	92.09
PRMSE	0.91	0.91	0.91	0.97	0.91	0.91
$PRMSE_0$	0.91	0.92	0.92	0.98	0.91	0.92
$MSE(\hat{\beta})$	.02756	.80040	.80040	.73421	.57197	.79754
$RMSE(\hat{\beta})$	.16602	.89465	.89465	.85686	.75629	.89305
$VAR(\hat{\beta})$	.01910	.07111	.07111	.00087	.05032	.07082
$Bias^2(\hat{\beta})$	.00846	.72929	.72929	.73334	.52165	.72673
MSE ($\partial p / \partial x$)	.00127	.00227	.00227	.02867	.00199	.00227
RMSE ($\partial p / \partial x$)	.03570	.04765	.04765	.16933	.04462	.04761

*N=5,000; unbalanced; restricted; errors drawn from chi-square distribution.

Table C.40 Used Cars Data - Monte Carlo MSE 101 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	56.69	56.69	56.69	56.69	56.69	56.69
Repayments	57.70	57.71	57.71	57.70	57.71	57.71
Defaults	55.60	55.59	55.59	55.60	55.60	55.59
Hold-Out	56.32	56.32	56.32	56.32	56.32	56.32
Repayments	58.89	58.89	58.89	58.90	58.89	58.89
Defaults	53.73	53.72	53.72	53.73	53.72	53.72
YRMSE	70.72	71.35	71.35	71.19	71.26	71.35
$YRMSE_0$	70.83	71.52	71.52	71.23	71.43	71.52
PRMSE	1.36	1.36	1.36	1.36	1.36	1.36
$PRMSE_0$	1.35	1.35	1.35	1.36	1.35	1.35
$MSE(\hat{\beta})$	.01187	.07751	.07751	.03715	.06929	.07743
$RMSE(\hat{\beta})$	.10895	.27841	.27841	.19274	.26322	.27825
$VAR(\hat{\beta})$	.01175	.03042	.03042	.00211	.02868	.03040
$Bias^2(\hat{\beta})$	.00012	.04710	.04710	.03504	.04061	.04703
MSE ($\partial p / \partial x$)	.00188	.00191	.00191	.00940	.00179	.00191
RMSE ($\partial p / \partial x$)	.04337	.04376	.04376	.09696	.04228	.04374

*N=5,000; balanced; restricted; errors drawn from normal distribution.

Table C.41 Used Cars Data - Monte Carlo MSE 102 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	60.04	60.05	60.05	60.05	60.05	60.05
Repayments	61.32	61.32	61.32	61.32	61.32	61.32
Defaults	58.71	58.72	58.72	58.73	58.72	58.72
Hold-Out	59.85	59.84	59.84	59.84	59.84	59.84
Repayments	62.57	62.57	62.57	62.58	62.57	62.57
Defaults	57.06	57.06	57.06	57.05	57.06	57.06
YRMSE	71.96	74.86	74.86	71.66	74.53	74.86
$YRMSE_0$	71.30	74.22	74.22	70.95	73.89	74.22
PRMSE	1.37	1.37	1.37	1.43	1.37	1.37
$PRMSE_0$	1.37	1.36	1.36	1.45	1.36	1.36
$MSE(\hat{\beta})$	.04781	.29217	.29217	.01562	.26421	.29187
$RMSE(\hat{\beta})$	.21866	.54052	.54052	.12500	.51401	.54025
$VAR(\hat{\beta})$	.01261	.03321	.03321	.00214	.03109	.03319
$Bias^2(\hat{\beta})$	.03520	.25895	.25895	.01348	.23312	.25868
MSE ($\partial p / \partial x$)	.00751	.00795	.00795	.00630	.00687	.00793
RMSE ($\partial p / \partial x$)	.08665	.08914	.08914	.07936	.08290	.08907

*N=5,000; balanced; restricted; errors drawn from *t*- distribution.

Table C.42 Used Cars Data - Monte Carlo MSE 103 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	59.59	59.59	59.59	59.59	59.59	59.59
Repayments	86.74	86.68	86.68	86.96	86.70	86.68
Defaults	22.11	22.20	22.20	21.81	22.16	22.20
Hold-Out	59.62	59.62	59.62	59.62	59.62	59.62
Repayments	87.63	87.58	87.58	87.82	87.60	87.58
Defaults	20.56	20.64	20.64	20.29	20.61	20.64
YRMSE	72.18	75.07	75.07	71.46	74.78	75.07
$YRMSE_0$	72.24	75.27	75.27	71.42	74.96	75.27
PRMSE	1.37	1.36	1.36	1.39	1.36	1.36
$PRMSE_0$	1.36	1.35	1.35	1.38	1.35	1.35
$MSE(\hat{\beta})$	.05842	.24231	.24231	.03584	.22234	.24210
$RMSE(\hat{\beta})$	.24170	.49225	.49225	.18932	.47153	.49204
$VAR(\hat{\beta})$	.01146	.03027	.03027	.00192	.02828	.03024
$Bias^2(\hat{\beta})$	.04696	.21205	.21205	.03393	.19406	.21186
MSE ($\partial p / \partial x$)	.00900	.00908	.00908	.00869	.00839	.00907
RMSE ($\partial p / \partial x$)	.09485	.09529	.09529	.09323	.09158	.09525

*N=5,000; balanced; restricted; errors drawn from chi-square distribution.

Table C.43 Used Cars Data - Monte Carlo MSE 111 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	71.03	71.42	71.42	69.73	71.13	71.41
Repayments	76.23	76.80	76.80	74.31	76.36	76.79
Defaults	39.70	39.02	39.02	42.00	39.57	39.02
Hold-Out	71.12	71.54	71.54	69.72	71.22	71.53
Repayments	76.86	77.47	77.47	74.81	77.02	77.47
Defaults	36.90	36.17	36.17	39.38	36.72	36.17
YRMSE	44.77	57.45	57.45	55.00	54.20	57.41
$YRMSE_0$	71.09	91.50	91.50	87.07	86.25	91.43
PRMSE	0.92	0.93	0.93	0.95	0.91	0.93
$PRMSE_0$	1.50	1.53	1.53	1.56	1.50	1.53
$MSE(\hat{\beta})$	.03411	1.07005	1.07005	.82105	.74712	1.06595
$RMSE(\hat{\beta})$	.18469	1.03443	1.03443	.90612	.86436	1.03245
$VAR(\hat{\beta})$	.03141	.11130	.11130	.00170	.08222	.11091
$Bias^2(\hat{\beta})$	.00270	.95875	.95875	.81935	.66490	.95504
MSE ($\partial p / \partial x$)	.00185	.00256	.00256	.04319	.00230	.00256
RMSE ($\partial p / \partial x$)	.04297	.05063	.05063	.20782	.04797	.05059

*N=2,000; unbalanced; restricted; errors drawn from normal distribution.

Table C.44 Used Cars Data - Monte Carlo MSE 112 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	81.59	81.82	81.82	81.11	81.64	81.82
Repayments	86.57	86.86	86.86	85.96	86.63	86.86
Defaults	27.44	27.12	27.12	28.18	27.41	27.12
Hold-Out	81.96	82.22	82.22	81.28	81.99	82.21
Repayments	87.31	87.64	87.64	86.47	87.35	87.63
Defaults	24.19	23.75	23.75	25.27	24.16	23.76
YRMSE	47.15	77.84	77.84	53.43	69.41	77.72
$YRMSE_0$	74.87	123.77	123.77	84.55	110.24	123.59
PRMSE	0.77	0.77	0.77	0.85	0.77	0.77
$PRMSE_0$	1.29	1.29	1.29	1.45	1.29	1.29
$MSE(\hat{\beta})$	.14782	2.85419	2.85419	.74372	1.82900	2.83888
$RMSE(\hat{\beta})$	.38447	1.68943	1.68944	.86239	1.35241	1.68490
$VAR(\hat{\beta})$	.04728	.20574	.20575	.00112	.10904	.20381
$Bias^2(\hat{\beta})$	.10054	2.64844	2.64845	.74261	1.71996	2.63507
MSE ($\partial p / \partial x$)	.00471	.00890	.00890	.03966	.00691	.00887
RMSE ($\partial p / \partial x$)	.06862	.09433	.09433	.19916	.08315	.09416

*N=2,000; unbalanced; restricted; errors drawn from t- distribution.

Table C.45 Used Cars Data - Monte Carlo MSE 113 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	75.93	76.07	76.07	75.72	76.02	76.07
Repayments	83.84	84.04	84.04	83.51	83.96	84.04
Defaults	25.86	25.66	25.66	26.34	25.78	25.66
Hold-Out	76.15	76.31	76.31	75.81	76.22	76.31
Repayments	84.58	84.81	84.81	84.10	84.69	84.81
Defaults	23.00	22.72	22.72	23.58	22.88	22.72
YRMSE	44.87	57.43	57.43	54.95	54.40	57.40
$YRMSE_0$	71.27	91.51	91.51	86.98	86.61	91.45
PRMSE	0.89	0.90	0.90	0.90	0.89	0.90
$PRMSE_0$	1.46	1.47	1.47	1.49	1.46	1.47
$MSE(\hat{\beta})$	.05003	.84279	.84279	.85549	.58913	.83959
$RMSE(\hat{\beta})$	.22367	.91803	.91804	.92493	.76755	.91629
$VAR(\hat{\beta})$	.03403	.12389	.12389	.00177	.09056	.12344
$Bias^2(\hat{\beta})$	.01599	.71890	.71890	.85372	.49857	.71615
MSE ($\partial p / \partial x$)	.00307	.00483	.00483	.04519	.00455	.00482
RMSE ($\partial p / \partial x$)	.05539	.06949	.06949	.21259	.06742	.06946

*N=2,000; unbalanced; restricted; errors drawn from chi-square distribution.

Table C.46 Used Cars Data - Monte Carlo MSE 121 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	57.85	57.85	57.85	57.85	57.85	57.85
Repayments	59.31	59.30	59.30	59.39	59.31	59.30
Defaults	56.22	56.23	56.23	56.16	56.23	56.23
Hold-Out	57.28	57.28	57.28	57.28	57.28	57.28
Repayments	59.99	59.98	59.98	60.06	59.99	59.98
Defaults	54.55	54.56	54.56	54.48	54.55	54.56
YRMSE	44.71	45.30	45.30	45.12	45.22	45.30
$YRMSE_0$	70.97	72.13	72.13	71.45	71.98	72.13
PRMSE	1.42	1.42	1.42	1.42	1.42	1.42
$PRMSE_0$	2.30	2.31	2.31	2.31	2.31	2.31
$MSE(\hat{\beta})$	.03013	.12302	.12302	.02775	.11258	.12291
$RMSE(\hat{\beta})$	.17358	.35074	.35074	.16658	.33553	.35058
$VAR(\hat{\beta})$	.02890	.07506	.07506	.00513	.07066	.07501
$Bias^2(\hat{\beta})$	.00123	.04796	.04796	.02262	.04192	.04789
MSE ($\partial p / \partial x$)	.00478	.00488	.00488	.00664	.00453	.00488
RMSE ($\partial p / \partial x$)	.06912	.06986	.06986	.08146	.06732	.06983

*N=2,000; balanced; restricted; errors drawn from normal distribution.

Table C.47 Used Cars Data - Monte Carlo MSE 122 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	61.63	61.64	61.64	61.65	61.64	61.64
Repayments	63.82	63.79	63.79	63.99	63.81	63.79
Defaults	59.33	59.38	59.38	59.19	59.36	59.38
Hold-Out	61.24	61.24	61.24	61.25	61.25	61.24
Repayments	64.64	64.60	64.60	64.82	64.62	64.60
Defaults	57.78	57.82	57.82	57.60	57.80	57.82
YRMSE	44.61	47.22	47.22	44.36	46.91	47.22
$YRMSE_0$	71.95	76.27	76.27	71.40	75.77	76.27
PRMSE	1.52	1.52	1.52	1.54	1.52	1.52
$PRMSE_0$	2.47	2.47	2.47	2.55	2.47	2.47
$MSE(\hat{\beta})$	.06892	.31618	.31618	.01370	.28793	.31588
$RMSE(\hat{\beta})$	.26253	.56230	.56230	.11706	.53659	.56203
$VAR(\hat{\beta})$	.03395	.08983	.08983	.00563	.08384	.08977
$Bias^2(\hat{\beta})$	.03497	.22635	.22635	.00807	.20409	.22612
MSE ($\partial p / \partial x$)	.01088	.01146	.01146	.00447	.01022	.01145
RMSE ($\partial p / \partial x$)	.10431	.10705	.10705	.06687	.10109	.10698

*N=2,000; balanced; restricted; errors drawn from *t*- distribution.

Table C.48 Used Cars Data - Monte Carlo MSE 123 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	60.66	60.66	60.66	60.67	60.66	60.66
Repayments	84.14	84.11	84.11	84.31	84.13	84.11
Defaults	28.48	28.52	28.52	28.27	28.49	28.52
Hold-Out	60.37	60.36	60.36	60.37	60.37	60.36
Repayments	83.74	83.71	83.71	83.89	83.72	83.71
Defaults	28.21	28.25	28.25	28.01	28.23	28.25
YRMSE	45.64	47.66	47.66	45.31	47.44	47.66
$YRMSE_0$	72.43	75.92	75.92	71.70	75.55	75.92
PRMSE	1.42	1.41	1.41	1.43	1.42	1.41
$PRMSE_0$	2.31	2.31	2.31	2.33	2.31	2.31
$MSE(\hat{\beta})$	.09080	.35471	.35471	.02067	.32716	.35442
$RMSE(\hat{\beta})$	.30133	.59558	.59558	.14377	.57198	.59533
$VAR(\hat{\beta})$	.02874	.07604	.07604	.00476	.07097	.07599
$Bias^2(\hat{\beta})$	.06206	.27867	.27867	.01591	.25619	.27843
MSE ($\partial p / \partial x$)	.01390	.01414	.01414	.00504	.01296	.01413
RMSE ($\partial p / \partial x$)	.11791	.11893	.11893	.07103	.11386	.11888

*N=2,000; balanced; restricted; errors drawn from chi-square distribution.

Table C.49 PSE Data - Monte Carlo Prediction 1 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.5266	-2.7307	-2.9629	-2.7307	-0.4169	-2.4158	-2.7267
	(-14.80)	(-15.05)	(-13.86)	(-15.05)	(-12.80)	(-15.18)	(-15.05)
Black	0.3366	0.6024	0.6110	0.6024	0.0965	0.5367	0.6016
	(4.46)	(4.63)	(4.07)	(4.63)	(4.12)	(4.54)	(4.63)
Other Race	0.1981	0.3878	0.3693	0.3878	0.0589	0.3437	0.3872
	(3.30)	(3.68)	(3.69)	(3.68)	(3.74)	(3.70)	(3.68)
Catholic HS	0.5466	1.0180	0.7114	1.0180	0.1156	0.8438	1.0156
	(4.70)	(4.58)	(6.19)	(4.58)	(6.13)	(5.00)	(4.59)
Income1	0.2634	0.4496	0.6250	0.4496	0.0959	0.4203	0.4493
	(4.54)	(4.50)	(4.74)	(4.50)	(4.78)	(4.58)	(4.50)
Income2	0.5857	1.0144	1.1458	1.0144	0.1792	0.9209	1.0132
	(7.76)	(7.51)	(7.98)	(7.51)	(8.20)	(7.77)	(7.52)
Income3	0.6770	1.2526	1.1065	1.2526	0.1743	1.0530	1.2497
	(6.76)	(6.81)	(7.61)	(6.81)	(7.76)	(7.31)	(6.82)
GPA	0.8602	1.5250	1.4910	1.5250	0.2360	1.3502	1.5227
	(22.03)	(21.74)	(23.01)	(21.74)	(25.99)	(23.48)	(21.77)
Parent HS	0.2777	0.4775	0.7305	0.4775	0.1108	0.4517	0.4772
	(3.53)	(3.56)	(3.78)	(3.56)	(3.81)	(3.63)	(3.56)
Parent College	0.7257	1.3106	1.2666	1.3106	0.1984	1.1343	1.3082
	(7.81)	(7.96)	(6.38)	(7.96)	(6.62)	(7.86)	(7.96)
Estimation Sample							
Percent Correct	80.52	80.55	80.44	80.55	80.29	80.55	80.55
Standard Deviation	.0056	.0056	.0056	.0056	.0053	.0056	.0055
% Repay. Correct	93.96	93.54	93.73	93.54	96.46	94.03	93.55
% Defaults Correct	34.34	35.92	34.77	35.92	24.75	34.25	35.90
Hold-out Sample							
Percent Correct	80.38	80.41	80.25	80.41	80.15	80.40	80.41
Standard Deviation	.0140	.0137	.0138	.0137	.0137	.0138	.0137
% Repay. Correct	93.92	93.51	93.65	93.51	96.41	93.98	93.52
% Defaults Correct	33.95	35.49	34.31	35.49	24.37	33.83	35.47

* N=5,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.50 PSE Data - Monte Carlo Prediction 2 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-2.3216	-3.9639	-4.1999	-3.9639	-0.8266	-3.7298	-3.9613
	(-25.06)	(-23.85)	(-24.60)	(-23.85)	(-31.16)	(-24.87)	(-23.86)
Black	0.3400	0.5847	0.5913	0.5847	0.1169	0.5471	0.5842
	(5.27)	(5.31)	(5.20)	(5.31)	(5.22)	(5.30)	(5.31)
Other Race	0.1731	0.3210	0.3100	0.3210	0.0618	0.2994	0.3208
	(3.14)	(3.40)	(3.36)	(3.40)	(3.38)	(3.41)	(3.40)
Catholic HS	0.5492	0.9424	0.8615	0.9424	0.1728	0.8729	0.9416
	(5.57)	(5.56)	(6.18)	(5.56)	(6.16)	(5.65)	(5.56)
Income1	0.2704	0.4637	0.5053	0.4637	0.0991	0.4378	0.4634
	(5.36)	(5.43)	(5.64)	(5.43)	(5.70)	(5.47)	(5.43)
Income2	0.6134	1.0380	1.1251	1.0380	0.2210	0.9798	1.0374
	(9.30)	(9.35)	(9.44)	(9.35)	(9.64)	(9.41)	(9.35)
Income3	0.7220	1.2699	1.2241	1.2699	0.2429	1.1787	1.2689
	(7.76)	(7.95)	(8.46)	(7.95)	(8.64)	(8.09)	(7.95)
GPA	0.8979	1.5314	1.5957	1.5314	0.3146	1.4379	1.5304
	(29.46)	(28.07)	(27.44)	(28.07)	(35.72)	(29.41)	(28.09)
Parent HS	0.2703	0.4684	0.5018	0.4684	0.0987	0.4420	0.4682
	(3.75)	(3.76)	(4.24)	(3.76)	(4.23)	(3.82)	(3.76)
Parent College	0.7608	1.3010	1.3237	1.3010	0.2615	1.2177	1.3001
	(8.75)	(8.70)	(8.91)	(8.70)	(9.16)	(8.78)	(8.70)
Estimation Sample							
Percent Correct	75.56	75.58	75.51	75.58	75.53	75.58	75.58
Standard Deviation	.0060	.0060	.0059	.0060	.0059	.0059	.0060
% Repay. Correct	71.88	72.11	70.74	72.11	70.89	71.96	72.11
% Defaults Correct	79.24	79.05	80.28	79.05	80.17	79.21	79.05
Hold-out Sample							
Percent Correct	75.50	75.50	75.47	75.50	75.48	75.51	75.50
Standard Deviation	.0138	.0137	.0142	.0137	.0141	.0138	.0137
% Repay. Correct	71.72	71.93	70.59	71.93	70.73	71.79	71.93
% Defaults Correct	79.28	79.08	80.35	79.08	80.23	79.23	79.08

* N=5,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.51 PSE Data - Monte Carlo Prediction 3 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.5223	-2.7245	-2.9409	-2.7245	-0.4120	-2.4051	-2.7204
	(-9.11)	(-9.35)	(-8.68)	(-9.35)	(-7.98)	(-9.45)	(-9.36)
Black	0.3416	0.6104	0.6132	0.6104	0.0967	0.5426	0.6096
	(2.90)	(3.00)	(2.62)	(3.00)	(2.66)	(2.94)	(3.00)
Other Race	0.1968	0.3859	0.3600	0.3859	0.0573	0.3406	0.3853
	(1.96)	(2.20)	(2.19)	(2.20)	(2.21)	(2.21)	(2.20)
Catholic HS	0.5514	1.0264	0.7053	1.0264	0.1144	0.8438	1.0238
	(2.91)	(2.90)	(3.89)	(2.90)	(3.84)	(3.20)	(2.91)
Income1	0.2676	0.4578	0.6311	0.4578	0.0966	0.4269	0.4574
	(3.00)	(2.99)	(3.16)	(2.99)	(3.17)	(3.04)	(2.99)
Income2	0.5931	1.0279	1.1550	1.0279	0.1801	0.9310	1.0267
	(4.98)	(4.85)	(5.29)	(4.85)	(5.38)	(5.04)	(4.86)
Income3	0.6795	1.2574	1.1073	1.2574	0.1739	1.0531	1.2545
	(4.42)	(4.39)	(5.03)	(4.39)	(5.08)	(4.74)	(4.39)
GPA	0.8628	1.5298	1.4938	1.5298	0.2358	1.3526	1.5275
	(14.44)	(14.18)	(15.07)	(14.18)	(16.90)	(15.38)	(14.19)
Parent HS	0.2655	0.4572	0.6998	0.4572	0.1057	0.4321	0.4569
	(2.10)	(2.12)	(2.26)	(2.12)	(2.28)	(2.16)	(2.12)
Parent College	0.7206	1.3027	1.2446	1.3027	0.1945	1.1231	1.3002
	(4.68)	(4.80)	(3.88)	(4.80)	(4.04)	(4.76)	(4.80)
Estimation Sample							
Percent Correct	80.59	80.60	80.53	80.60	80.37	80.63	80.60
Standard Deviation	.0088	.0088	.0088	.0088	.0084	.0087	.0087
% Repay. Correct	93.96	93.54	93.77	93.54	96.43	94.04	93.55
% Defaults Correct	34.46	35.97	34.87	35.97	24.96	34.35	35.95
Hold-out Sample							
Percent Correct	80.40	80.39	80.28	80.39	80.17	80.42	80.39
Standard Deviation	.0130	.0130	.0129	.0130	.0128	.0130	.0130
% Repay. Correct	93.81	93.38	93.59	93.38	96.28	93.89	93.39
% Defaults Correct	34.01	35.44	34.23	35.44	24.44	33.84	35.43

* N=2,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.52 PSE Data - Monte Carlo Prediction 4 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-2.3203	-3.9615	-4.1927	-3.9615	-0.8243	-3.7257	-3.9589
	(-15.37)	(-14.79)	(-15.17)	(-14.79)	(-19.41)	(-15.45)	(-14.79)
Black	0.3364	0.5793	0.5834	0.5793	0.1153	0.5415	0.5789
	(3.08)	(3.11)	(3.02)	(3.11)	(3.05)	(3.11)	(3.11)
Other Race	0.1764	0.3265	0.3158	0.3265	0.0628	0.3044	0.3262
	(1.94)	(2.09)	(2.08)	(2.09)	(2.10)	(2.09)	(2.09)
Catholic HS	0.5595	0.9629	0.8720	0.9629	0.1749	0.8902	0.9620
	(3.52)	(3.48)	(3.80)	(3.48)	(3.80)	(3.53)	(3.48)
Income1	0.2679	0.4592	0.5002	0.4592	0.0980	0.4334	0.4589
	(3.35)	(3.38)	(3.54)	(3.38)	(3.59)	(3.41)	(3.38)
Income2	0.6122	1.0357	1.1230	1.0357	0.2203	0.9773	1.0350
	(5.74)	(5.76)	(5.85)	(5.76)	(6.04)	(5.81)	(5.76)
Income3	0.7208	1.2665	1.2176	1.2665	0.2414	1.1744	1.2655
	(5.03)	(5.14)	(5.48)	(5.14)	(5.57)	(5.23)	(5.14)
GPA	0.8993	1.5342	1.5976	1.5342	0.3146	1.4399	1.5332
	(16.76)	(16.13)	(15.82)	(16.13)	(20.78)	(16.92)	(16.14)
Parent HS	0.2693	0.4658	0.4967	0.4658	0.0977	0.4392	0.4655
	(2.29)	(2.31)	(2.57)	(2.31)	(2.56)	(2.34)	(2.31)
Parent College	0.7569	1.2950	1.3130	1.2950	0.2592	1.2111	1.2940
	(5.61)	(5.61)	(5.68)	(5.61)	(5.84)	(5.66)	(5.62)
Estimation Sample							
Percent Correct	75.48	75.51	75.50	75.51	75.51	75.52	75.51
Standard Deviation	.0096	.0096	.0095	.0096	.0095	.0097	.0096
% Repay. Correct	71.90	72.14	70.82	72.14	70.96	71.99	72.14
% Defaults Correct	79.06	78.88	80.18	78.88	80.05	79.05	78.88
Hold-out Sample							
Percent Correct	75.31	75.32	75.34	75.32	75.34	75.33	75.32
Standard Deviation	.0140	.0141	.0140	.0141	.0141	.0141	.0141
% Repay. Correct	71.73	71.95	70.63	71.95	70.76	71.80	71.94
% Defaults Correct	78.89	78.70	80.05	78.70	79.92	78.86	78.70

* N=2,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.53 PSE Data - Monte Carlo Prediction 5 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-1.5364	-2.7453	-2.9583	-2.7467	-0.4136	-2.4189	-2.7424
	(-6.04)	(-6.24)	(-5.75)	(-6.25)	(-5.30)	(-6.31)	(-6.25)
Black	0.3324	0.5955	0.5913	0.5960	0.0928	0.5277	0.5951
	(1.94)	(2.02)	(1.75)	(2.03)	(1.80)	(2.00)	(2.03)
Other Race	0.2107	0.4078	0.3774	0.4075	0.0598	0.3585	0.4069
	(1.46)	(1.60)	(1.60)	(1.60)	(1.61)	(1.61)	(1.60)
Catholic HS	0.5734	1.0670	0.7160	1.1371	0.1157	0.8710	1.1008
	(2.03)	(1.98)	(2.75)	(1.14)	(2.72)	(2.19)	(1.63)
Income1	0.2696	0.4623	0.6405	0.4634	0.0975	0.4315	0.4630
	(2.17)	(2.17)	(2.29)	(2.17)	(2.31)	(2.21)	(2.17)
Income2	0.6032	1.0480	1.1764	1.0485	0.1825	0.9469	1.0472
	(3.47)	(3.42)	(3.69)	(3.43)	(3.81)	(3.56)	(3.43)
Income3	0.7200	1.3397	1.1425	1.3407	0.1788	1.1019	1.3369
	(2.82)	(2.81)	(3.36)	(2.81)	(3.42)	(3.12)	(2.82)
GPA	0.8692	1.5387	1.5028	1.5391	0.2364	1.3580	1.5368
	(9.93)	(9.85)	(10.41)	(9.85)	(11.73)	(10.71)	(9.86)
Parent HS	0.2640	0.4556	0.6934	0.4554	0.1043	0.4291	0.4551
	(1.40)	(1.43)	(1.51)	(1.43)	(1.52)	(1.45)	(1.43)
Parent College	0.7188	1.2992	1.2292	1.2991	0.1913	1.1132	1.2965
	(3.09)	(3.16)	(2.54)	(3.16)	(2.63)	(3.12)	(3.16)
Estimation Sample							
Percent Correct	80.62	80.67	80.65	80.68	80.44	80.69	80.67
Standard Deviation	.0135	.0135	.0130	.0135	.0123	.0133	.0135
% Repay. Correct	93.77	93.42	93.68	93.41	96.24	93.90	93.42
% Defaults Correct	35.37	36.84	35.82	36.85	26.00	35.23	36.83
Hold-out Sample							
Percent Correct	80.15	80.14	80.08	80.16	79.98	80.19	80.16
Standard Deviation	.0125	.0124	.0127	.0125	.0129	.0126	.0125
% Repay. Correct	93.48	93.09	93.33	93.09	95.99	93.59	93.10
% Defaults Correct	34.33	35.64	34.49	35.67	24.91	34.10	35.65

* N=1,000; unbalanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.54 PSE Data - Monte Carlo Prediction 6 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	DA	ME	GME1	GME2	GME3
Intercept	-2.3507	-4.0164	-4.2390	-4.0164	-0.8287	-3.7711	-4.0137
	(-10.96)	(-10.53)	(-10.81)	(-10.53)	(-13.86)	(-11.03)	(-10.54)
Black	0.3415	0.5883	0.5939	0.5883	0.1166	0.5493	0.5879
	(2.15)	(2.16)	(2.14)	(2.16)	(2.17)	(2.17)	(2.16)
Other Race	0.1825	0.3350	0.3241	0.3350	0.0642	0.3120	0.3348
	(1.42)	(1.52)	(1.51)	(1.52)	(1.53)	(1.52)	(1.52)
Catholic HS	0.5607	0.9674	0.8657	0.9674	0.1727	0.8909	0.9665
	(2.40)	(2.37)	(2.58)	(2.37)	(2.59)	(2.41)	(2.37)
Income1	0.2752	0.4726	0.5137	0.4726	0.1001	0.4455	0.4723
	(2.36)	(2.40)	(2.50)	(2.40)	(2.52)	(2.42)	(2.40)
Income2	0.6194	1.0489	1.1346	1.0489	0.2214	0.9883	1.0482
	(4.08)	(4.07)	(4.12)	(4.07)	(4.23)	(4.11)	(4.07)
Income3	0.7343	1.2857	1.2356	1.2857	0.2436	1.1901	1.2846
	(3.65)	(3.69)	(3.95)	(3.69)	(4.01)	(3.75)	(3.69)
GPA	0.9065	1.5474	1.6082	1.5474	0.3149	1.4500	1.5463
	(12.25)	(11.72)	(11.46)	(11.72)	(15.31)	(12.34)	(11.73)
Parent HS	0.2770	0.4800	0.5049	0.4800	0.0988	0.4511	0.4797
	(1.64)	(1.66)	(1.83)	(1.66)	(1.82)	(1.68)	(1.66)
Parent College	0.7744	1.3260	1.3377	1.3260	0.2626	1.2374	1.3249
	(3.80)	(3.81)	(3.88)	(3.81)	(3.99)	(3.85)	(3.81)
Estimation Sample							
Percent Correct	75.55	75.60	75.59	75.60	75.58	75.61	75.60
Standard Deviation	.0134	.0135	.0133	.0135	.0133	.0134	.0135
% Repay. Correct	72.07	72.33	70.98	72.33	71.11	72.17	72.33
% Defaults Correct	79.03	78.87	80.20	78.87	80.05	79.05	78.88
Hold-out Sample							
Percent Correct	75.03	75.05	75.05	75.05	75.05	75.06	75.05
Standard Deviation	.0130	.0127	.0129	.0127	.0130	.0127	.0127
% Repay. Correct	71.60	71.82	70.47	71.82	70.61	71.67	71.82
% Defaults Correct	78.47	78.28	79.64	78.28	79.49	78.45	78.28

* N=1,000; balanced; unrestricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.55 PSE Data - Monte Carlo Prediction 7 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-1.5266	-2.7307	-2.7307	-0.4169	-2.4158	-2.7267
	(-14.80)	(-15.05)	(-15.05)	(-12.80)	(-15.18)	(-15.05)
Black	0.3366	0.6024	0.6024	0.0965	0.5367	0.6016
	(4.46)	(4.63)	(4.63)	(4.12)	(4.54)	(4.63)
Other Race	0.1981	0.3878	0.3878	0.0589	0.3437	0.3872
	(3.30)	(3.68)	(3.68)	(3.74)	(3.70)	(3.68)
Catholic HS	0.5466	1.0180	1.0180	0.1156	0.8438	1.0156
	(4.70)	(4.58)	(4.58)	(6.13)	(5.00)	(4.59)
Income1	0.2634	0.4496	0.4496	0.0959	0.4203	0.4493
	(4.54)	(4.50)	(4.50)	(4.78)	(4.58)	(4.50)
Income2	0.5857	1.0144	1.0144	0.1792	0.9209	1.0132
	(7.76)	(7.51)	(7.51)	(8.20)	(7.77)	(7.52)
Income3	0.6770	1.2526	1.2526	0.1743	1.0530	1.2497
	(6.76)	(6.81)	(6.81)	(7.76)	(7.31)	(6.82)
GPA	0.8602	1.5250	1.5250	0.2360	1.3502	1.5227
	(22.03)	(21.74)	(21.74)	(25.99)	(23.48)	(21.77)
Parent HS	0.2777	0.4775	0.4775	0.1108	0.4517	0.4772
	(3.53)	(3.56)	(3.56)	(3.81)	(3.63)	(3.56)
Parent College	0.7257	1.3106	1.3106	0.1984	1.1343	1.3082
	(7.81)	(7.96)	(7.96)	(6.62)	(7.86)	(7.96)
Estimation Sample						
Percent Correct	80.52	80.55	80.55	80.29	80.55	80.55
Standard Deviation	.0056	.0056	.0056	.0053	.0056	.0055
% Repay. Correct	93.96	93.54	93.54	96.46	94.03	93.55
% Defaults Correct	34.34	35.92	35.92	24.75	34.25	35.90
Hold-out Sample						
Percent Correct	80.38	80.41	80.41	80.15	80.40	80.41
Standard Deviation	.0140	.0137	.0137	.0137	.0138	.0137
% Repay. Correct	93.92	93.51	93.51	96.41	93.98	93.52
% Defaults Correct	33.95	35.49	35.49	24.37	33.83	35.47

* N=5,000; unbalanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.56 PSE Data - Monte Carlo Prediction 8 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-2.3216	-3.9639	-3.9639	-0.8266	-3.7298	-3.9613
	(-25.06)	(-23.85)	(-23.85)	(-31.16)	(-24.87)	(-23.86)
Black	0.3400	0.5847	0.5847	0.1169	0.5471	0.5842
	(5.27)	(5.31)	(5.31)	(5.22)	(5.30)	(5.31)
Other Race	0.1731	0.3210	0.3210	0.0618	0.2994	0.3208
	(3.14)	(3.40)	(3.40)	(3.38)	(3.41)	(3.40)
Catholic HS	0.5492	0.9424	0.9424	0.1728	0.8729	0.9416
	(5.57)	(5.56)	(5.56)	(6.16)	(5.65)	(5.56)
Income1	0.2704	0.4637	0.4637	0.0991	0.4378	0.4634
	(5.36)	(5.43)	(5.43)	(5.70)	(5.47)	(5.43)
Income2	0.6134	1.0380	1.0380	0.2210	0.9798	1.0374
	(9.30)	(9.35)	(9.35)	(9.64)	(9.41)	(9.35)
Income3	0.7220	1.2699	1.2699	0.2429	1.1787	1.2689
	(7.76)	(7.95)	(7.95)	(8.64)	(8.09)	(7.95)
GPA	0.8979	1.5314	1.5314	0.3146	1.4379	1.5304
	(29.46)	(28.07)	(28.07)	(35.72)	(29.41)	(28.09)
Parent HS	0.2703	0.4684	0.4684	0.0987	0.4420	0.4682
	(3.75)	(3.76)	(3.76)	(4.23)	(3.82)	(3.76)
Parent College	0.7608	1.3010	1.3010	0.2615	1.2177	1.3001
	(8.75)	(8.70)	(8.70)	(9.16)	(8.78)	(8.70)
Estimation Sample						
Percent Correct	75.56	75.58	75.58	75.53	75.58	75.58
Standard Deviation	.0060	.0060	.0060	.0059	.0059	.0060
% Repay. Correct	71.88	72.11	72.11	70.89	71.96	72.11
% Defaults Correct	79.24	79.05	79.05	80.17	79.21	79.05
Hold-out Sample						
Percent Correct	75.50	75.50	75.50	75.48	75.51	75.50
Standard Deviation	.0138	.0137	.0137	.0141	.0138	.0137
% Repay. Correct	71.72	71.93	71.93	70.73	71.79	71.93
% Defaults Correct	79.28	79.08	79.08	80.23	79.23	79.08

* N=5,000; balanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.57 PSE Data - Monte Carlo Prediction 9 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-1.5233	-2.7260	-2.7260	-0.4122	-2.4063	-2.7218
	(-9.20)	(-9.43)	(-9.43)	(-8.03)	(-9.53)	(-9.43)
Black	0.3418	0.6106	0.6106	0.0967	0.5427	0.6097
	(2.90)	(3.00)	(3.00)	(2.67)	(2.95)	(3.00)
Other Race	0.1976	0.3868	0.3868	0.0574	0.3414	0.3862
	(2.00)	(2.23)	(2.23)	(2.25)	(2.24)	(2.23)
Catholic HS	0.5514	1.0264	1.0264	0.1144	0.8437	1.0238
	(2.91)	(2.90)	(2.90)	(3.84)	(3.20)	(2.91)
Income1	0.2675	0.4576	0.4576	0.0966	0.4268	0.4572
	(3.00)	(2.99)	(2.99)	(3.17)	(3.04)	(2.99)
Income2	0.5930	1.0277	1.0277	0.1801	0.9309	1.0265
	(4.98)	(4.85)	(4.85)	(5.38)	(5.03)	(4.85)
Income3	0.6795	1.2573	1.2573	0.1739	1.0530	1.2543
	(4.42)	(4.39)	(4.39)	(5.08)	(4.74)	(4.39)
GPA	0.8627	1.5298	1.5298	0.2358	1.3526	1.5275
	(14.43)	(14.17)	(14.17)	(16.90)	(15.38)	(14.19)
Parent HS	0.2666	0.4588	0.4588	0.1059	0.4334	0.4585
	(2.14)	(2.17)	(2.17)	(2.31)	(2.20)	(2.17)
Parent College	0.7215	1.3041	1.3041	0.1947	1.1243	1.3017
	(4.74)	(4.85)	(4.85)	(4.07)	(4.81)	(4.85)
Estimation Sample						
Percent Correct	80.59	80.61	80.61	80.37	80.63	80.60
Standard Deviation	.0088	.0088	.0088	.0084	.0087	.0087
% Repay. Correct	93.96	93.54	93.54	96.43	94.04	93.55
% Defaults Correct	34.45	35.97	35.97	24.95	34.35	35.95
Hold-out Sample						
Percent Correct	80.40	80.39	80.39	80.17	80.43	80.39
Standard Deviation	.0130	.0130	.0130	.0128	.0130	.0130
% Repay. Correct	93.81	93.38	93.38	96.28	93.89	93.39
% Defaults Correct	34.01	35.45	35.45	24.44	33.83	35.43

* N=2,000; unbalanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.58 PSE Data - Monte Carlo Prediction 10 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-2.3209	-3.9624	-3.9624	-0.8244	-3.7265	-3.9597
	(-15.43)	(-14.83)	(-14.83)	(-19.48)	(-15.50)	(-14.84)
Black	0.3367	0.5796	0.5796	0.1153	0.5418	0.5791
	(3.09)	(3.12)	(3.12)	(3.05)	(3.11)	(3.12)
Other Race	0.1774	0.3278	0.3278	0.0631	0.3056	0.3276
	(2.00)	(2.14)	(2.14)	(2.15)	(2.14)	(2.14)
Catholic HS	0.5595	0.9629	0.9629	0.1749	0.8903	0.9620
	(3.52)	(3.48)	(3.48)	(3.80)	(3.53)	(3.48)
Income1	0.2679	0.4592	0.4592	0.0980	0.4334	0.4589
	(3.35)	(3.38)	(3.38)	(3.59)	(3.41)	(3.38)
Income2	0.6123	1.0358	1.0358	0.2203	0.9774	1.0351
	(5.74)	(5.76)	(5.76)	(6.04)	(5.81)	(5.76)
Income3	0.7209	1.2665	1.2666	0.2414	1.1744	1.2655
	(5.03)	(5.14)	(5.14)	(5.57)	(5.23)	(5.14)
GPA	0.8992	1.5342	1.5342	0.3146	1.4399	1.5331
	(16.76)	(16.13)	(16.13)	(20.78)	(16.92)	(16.14)
Parent HS	0.2697	0.4665	0.4665	0.0978	0.4398	0.4662
	(2.31)	(2.32)	(2.32)	(2.57)	(2.36)	(2.32)
Parent College	0.7572	1.2955	1.2955	0.2593	1.2115	1.2946
	(5.63)	(5.63)	(5.63)	(5.84)	(5.67)	(5.63)
Estimation Sample						
Percent Correct	75.48	75.51	75.51	75.51	75.52	75.51
Standard Deviation	.0096	.0096	.0096	.0095	.0097	.0096
% Repay. Correct	71.90	72.14	72.14	70.96	71.99	72.14
% Defaults Correct	79.06	78.88	78.88	80.05	79.05	78.88
Hold-out Sample						
Percent Correct	75.31	75.32	75.32	75.34	75.33	75.32
Standard Deviation	.0140	.0141	.0141	.0141	.0141	.0141
% Repay. Correct	71.73	71.95	71.95	70.76	71.80	71.95
% Defaults Correct	78.89	78.70	78.70	79.92	78.86	78.70

*N=2,000; balanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.59 PSE Data - Monte Carlo Prediction 11 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-1.5463	-2.7612	-2.7612	-0.4162	-2.4320	-2.7569
	(-6.34)	(-6.53)	(-6.53)	(-5.56)	(-6.61)	(-6.53)
Black	0.3355	0.5998	0.5998	0.0939	0.5313	0.5989
	(2.02)	(2.08)	(2.08)	(1.88)	(2.05)	(2.08)
Other Race	0.2166	0.4161	0.4161	0.0613	0.3661	0.4154
	(1.60)	(1.73)	(1.73)	(1.77)	(1.74)	(1.73)
Catholic HS	0.6013	1.1475	1.1369	0.1157	0.8709	1.1006
	(1.39)	(1.03)	(1.14)	(2.73)	(2.19)	(1.63)
Income1	0.2711	0.4648	0.4648	0.0977	0.4328	0.4644
	(2.26)	(2.26)	(2.26)	(2.37)	(2.30)	(2.26)
Income2	0.6041	1.0492	1.0492	0.1826	0.9475	1.0479
	(3.52)	(3.46)	(3.46)	(3.87)	(3.60)	(3.46)
Income3	0.7214	1.3420	1.3420	0.1790	1.1031	1.3382
	(2.85)	(2.83)	(2.83)	(3.47)	(3.16)	(2.84)
GPA	0.8693	1.5389	1.5389	0.2364	1.3579	1.5365
	(9.92)	(9.84)	(9.84)	(11.72)	(10.71)	(9.86)
Parent HS	0.2723	0.4689	0.4689	0.1066	0.4411	0.4685
	(1.58)	(1.60)	(1.60)	(1.67)	(1.62)	(1.60)
Parent College	0.7260	1.3108	1.3108	0.1934	1.1237	1.3082
	(3.27)	(3.33)	(3.33)	(2.81)	(3.30)	(3.33)
Estimation Sample						
Percent Correct	80.62	80.68	80.68	80.43	80.69	80.67
Standard Deviation	.0134	.0134	.0134	.0123	.0134	.0134
% Repay. Correct	93.77	93.42	93.42	96.25	93.90	93.42
% Defaults Correct	35.37	36.84	36.84	25.97	35.22	36.82
Hold-out Sample						
Percent Correct	80.18	80.17	80.17	79.98	80.20	80.17
Standard Deviation	.0125	.0124	.0124	.0130	.0126	.0124
% Repay. Correct	93.50	93.10	93.10	95.99	93.60	93.11
% Defaults Correct	34.37	35.68	35.68	24.91	34.12	35.66

* N=1,000; unbalanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C.60 PSE Data - Monte Carlo Prediction 12 (See Table 2.3 for design)

Variable\Estimator	Probit	Logit	ME	GME1	GME2	GME3
Intercept	-2.3556	-4.0239	-4.0239	-0.8301	-3.7781	-4.0212
	(-11.25)	(-10.78)	(-10.78)	(-14.30)	(-11.30)	(-10.78)
Black	0.3435	0.5913	0.5913	0.1172	0.5521	0.5908
	(2.20)	(2.20)	(2.20)	(2.21)	(2.21)	(2.20)
Other Race	0.1884	0.3436	0.3436	0.0658	0.3199	0.3433
	(1.60)	(1.68)	(1.68)	(1.69)	(1.68)	(1.68)
Catholic HS	0.5613	0.9684	0.9684	0.1729	0.8919	0.9676
	(2.42)	(2.39)	(2.39)	(2.61)	(2.43)	(2.39)
Income1	0.2755	0.4729	0.4729	0.1002	0.4458	0.4726
	(2.39)	(2.42)	(2.42)	(2.53)	(2.44)	(2.42)
Income2	0.6197	1.0491	1.0491	0.2215	0.9885	1.0485
	(4.10)	(4.09)	(4.09)	(4.24)	(4.12)	(4.09)
Income3	0.7345	1.2862	1.2862	0.2437	1.1906	1.2851
	(3.66)	(3.69)	(3.69)	(4.02)	(3.76)	(3.69)
GPA	0.9064	1.5473	1.5473	0.3149	1.4499	1.5462
	(12.24)	(11.72)	(11.72)	(15.31)	(12.34)	(11.73)
Parent HS	0.2812	0.4865	0.4865	0.0999	0.4571	0.4862
	(1.73)	(1.75)	(1.75)	(1.91)	(1.77)	(1.75)
Parent College	0.7779	1.3314	1.3314	0.2635	1.2424	1.3304
	(3.92)	(3.92)	(3.92)	(4.09)	(3.95)	(3.92)
Estimation Sample						
Percent Correct	75.56	75.60	75.60	75.58	75.61	75.60
Standard Deviation	.0134	.0135	.0135	.0134	.0134	.0135
% Repay. Correct	72.07	72.32	72.32	71.12	72.17	72.32
% Defaults Correct	79.04	78.88	78.88	80.05	79.06	78.88
Hold-out Sample						
Percent Correct	75.05	75.07	75.07	75.06	75.08	75.07
Standard Deviation	.0130	.0128	.0128	.0130	.0128	.0128
% Repay. Correct	71.60	71.83	71.83	70.62	71.68	71.83
% Defaults Correct	78.50	78.31	78.31	79.51	78.48	78.31

* N=1,000; balanced; restricted; t-statistics (in parentheses) are computed from sample standard errors.

Table C. 61 PSE Data - Monte Carlo MSE 11 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.78	79.79	79.72	79.79	79.55	79.79	79.79
PSE	93.53	93.30	93.09	93.30	95.74	93.65	93.30
Non-PSE	34.64	35.42	35.80	35.42	26.36	34.29	35.41
Hold-Out	80.27	80.27	80.14	80.27	79.99	80.28	80.27
PSE	93.31	93.03	92.85	93.03	95.69	93.44	93.04
Non-PSE	36.16	37.09	37.17	37.09	26.87	35.75	37.07
YRMSE	70.78	97.51	93.31	97.51	94.30	87.03	97.36
$YRMSE_0$	31.68	44.07	41.95	44.07	42.63	39.19	44.00
PRMSE	1.28	1.43	2.16	1.43	3.99	1.30	1.43
$PRMSE_0$	0.58	0.65	0.97	0.65	1.79	0.58	0.65
$MSE(\hat{\beta})$	.06564	2.89661	3.79238	2.89661	2.43038	1.65363	2.87848
$RMSE(\hat{\beta})$	.25621	1.70194	1.94740	1.70194	1.55897	1.28593	1.69661
$VAR(\hat{\beta})$	.06527	.20491	.21316	.20491	.00535	.15217	.20410
$Bias^2(\hat{\beta})$	.00037	2.69170	3.57922	2.69170	2.42503	1.50146	2.67437
MSE ($\partial p / \partial x$)	.00973	.01434	.03662	.01434	.42161	.00974	.01419
RMSE ($\partial p / \partial x$)	.09863	.11973	.19136	.11973	.64931	.09869	.11911

*N=5,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table C. 62 PSE Data - Monte Carlo MSE 12 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	84.17	84.20	83.89	84.20	83.38	84.16	84.20
PSE	95.31	94.92	94.72	94.92	97.48	95.40	94.93
Non-PSE	40.18	41.88	41.11	41.88	27.70	39.78	41.86
Hold-Out	84.44	84.47	84.13	84.47	83.64	84.42	84.47
PSE	95.23	94.76	94.47	94.76	97.49	95.33	94.77
Non-PSE	41.07	43.13	42.61	43.13	27.98	40.58	43.09
YRMSE	72.95	128.35	117.68	128.35	92.63	105.17	127.99
$YRMSE_0$	32.69	58.21	52.92	58.21	41.88	47.48	58.04
PRMSE	2.19	1.57	3.17	1.57	6.46	2.39	1.58
$PRMSE_0$	0.99	0.71	1.43	0.71	2.92	1.08	0.71
$MSE(\hat{\beta})$	.27597	7.05055	8.70149	7.05055	2.32254	3.67324	6.99342
$RMSE(\hat{\beta})$	.52533	2.65529	2.94983	2.65529	1.52399	1.91657	2.64451
$VAR(\hat{\beta})$	.08010	.26376	.25734	.26376	.00473	.16427	.26183
$Bias^2(\hat{\beta})$	.19586	6.78678	8.44414	6.78678	2.31780	3.50897	6.73160
MSE ($\partial p / \partial x$)	.02487	.04698	.12674	.04698	.41052	.01449	.04632
RMSE ($\partial p / \partial x$)	.15770	.21676	.35600	.21676	.64072	.12037	.21523

*N=5,000; unbalanced; unrestricted; errors drawn from t- distribution.

Table C. 63 PSE Data - Monte Carlo MSE 13 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	81.53	81.55	81.44	81.55	81.18	81.54	81.55
PSE	91.06	90.93	91.29	90.93	93.44	91.34	90.94
Non-PSE	51.28	51.74	50.16	51.74	42.27	50.41	51.73
Hold-Out	82.01	82.02	81.88	82.02	81.67	82.01	82.01
PSE	91.18	91.06	91.07	91.06	93.15	91.38	91.06
Non-PSE	52.09	52.49	51.86	52.49	44.13	51.39	52.47
YRMSE	76.90	134.27	117.46	134.27	93.05	108.07	133.86
$YRMSE_0$	34.64	61.25	53.18	61.25	42.06	49.10	61.05
PRMSE	69.00	68.82	68.60	68.82	71.39	69.30	68.83
$PRMSE_0$	31.74	31.66	31.55	31.66	32.86	31.89	31.67
$MSE(\hat{\beta})$	1.19691	11.7649	12.4019	11.7649	1.92572	6.60128	11.6775
$RMSE(\hat{\beta})$	1.09404	3.43000	3.52163	3.43000	1.38770	2.56930	3.41723
$VAR(\hat{\beta})$	.07910	.25722	.24808	.25722	.00492	.16276	.25537
$Bias^2(\hat{\beta})$	1.11781	11.5077	12.1538	11.5077	1.92080	6.43852	11.4221
MSE ($\partial p / \partial x$)	.17803	.23877	.31962	.23877	.36739	.09986	.23630
RMSE ($\partial p / \partial x$)	.42193	.48864	.56535	.48864	.60613	.31600	.48611

*N=5,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table C. 64 PSE Data - Monte Carlo MSE 21 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	73.53	73.54	73.52	73.54	73.52	73.53	73.54
PSE	70.27	70.43	69.22	70.43	69.35	70.29	70.43
Non-PSE	76.77	76.63	77.80	76.63	77.68	76.75	76.63
Hold-Out	72.73	72.73	72.73	72.73	72.73	72.74	72.73
PSE	68.55	68.73	67.37	68.73	67.50	68.58	68.73
Non-PSE	76.86	76.69	78.03	76.69	77.89	76.85	76.69
YRMSE	70.71	82.66	84.13	82.66	81.18	79.57	82.62
$YRMSE_0$	31.72	36.74	37.41	36.74	36.05	35.45	36.73
PRMSE	1.38	1.46	1.74	1.46	2.79	1.39	1.46
$PRMSE_0$	0.63	0.66	0.79	0.66	1.19	0.63	0.66
$MSE(\hat{\beta})$	.05349	4.46490	5.32587	4.46490	3.79778	3.29101	4.45082
$RMSE(\hat{\beta})$	.23129	2.11303	2.30779	2.11303	1.94879	1.81412	2.10970
$VAR(\hat{\beta})$	.05324	.15664	.14616	.15664	.00520	.13313	.15636
$Bias^2(\hat{\beta})$	.00026	4.30826	5.17972	4.30826	3.79258	3.15789	4.29446
MSE ($\partial p / \partial x$)	.00669	.00759	.00728	.00759	.64692	.00752	.00756
RMSE ($\partial p / \partial x$)	.08182	.08713	.08534	.08713	.80431	.08673	.08698

*N=5,000; balanced; unrestricted; errors drawn from normal distribution.

Table C. 65 PSE Data - Monte Carlo MSE 22 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.26	79.26	79.21	79.26	79.22	79.26	79.26
PSE	75.96	76.20	74.41	76.20	74.61	75.98	76.20
Non-PSE	82.44	82.22	83.84	82.22	83.67	82.43	82.22
Hold-Out	78.44	78.44	78.36	78.44	78.37	78.43	78.44
PSE	74.18	74.48	72.41	74.48	72.62	74.20	74.48
Non-PSE	82.50	82.23	84.03	82.23	83.85	82.48	82.23
YRMSE	73.18	107.51	109.82	107.51	79.08	98.14	107.39
$YRMSE_0$	32.66	47.20	48.31	47.20	35.06	43.23	47.15
PRMSE	1.88	1.50	2.03	1.50	5.36	1.86	1.50
$PRMSE_0$	0.84	0.68	0.91	0.68	2.29	0.82	0.68
$MSE(\hat{\beta})$	1.04610	15.9629	17.8376	15.9629	3.09457	11.4317	15.9039
$RMSE(\hat{\beta})$	1.02279	3.99536	4.22346	3.99536	1.75914	3.38107	3.98796
$VAR(\hat{\beta})$	.06646	.20404	.17101	.20404	.00432	.15561	.20337
$Bias^2(\hat{\beta})$	.97964	15.7589	17.6666	15.7589	3.09025	11.2760	15.7005
MSE ($\partial p / \partial x$)	.05477	.05759	.03267	.05759	.57529	.03196	.05725
RMSE ($\partial p / \partial x$)	.23402	.23997	.18075	.23997	.75848	.17878	.23927

*N=5,000; balanced; unrestricted; errors drawn from t- distribution.

Table C. 66 PSE Data - Monte Carlo MSE 23 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	75.45	75.47	75.57	75.47	75.56	75.48	75.47
PSE	69.71	69.54	68.23	69.54	68.38	69.39	69.54
Non-PSE	80.49	80.69	82.02	80.69	81.88	80.83	80.69
Hold-Out	74.68	74.70	74.81	74.70	74.80	74.72	74.70
PSE	67.81	67.60	66.31	67.60	66.46	67.47	67.60
Non-PSE	80.64	80.86	82.17	80.86	82.03	81.01	80.87
YRMSE	71.41	89.98	93.92	89.98	80.60	85.42	89.93
$YRMSE_0$	32.01	39.89	41.66	39.89	35.74	37.97	39.87
PRMSE	33.18	33.00	33.04	33.00	34.71	33.20	33.00
$PRMSE_0$	14.15	14.07	14.09	14.07	14.75	14.15	14.07
$MSE(\hat{\beta})$	.14525	6.80643	8.32870	6.80643	3.61119	5.00014	6.78433
$RMSE(\hat{\beta})$	.38112	2.60891	2.88595	2.60891	1.90031	2.23610	2.60467
$VAR(\hat{\beta})$	.05470	.16849	.15567	.16849	.00484	.13920	.16813
$Bias^2(\hat{\beta})$	.09055	6.63794	8.17303	6.63794	3.60635	4.86094	6.61620
MSE ($\partial p / \partial x$)	.00754	.00788	.00821	.00788	.62967	.00900	.00788
RMSE ($\partial p / \partial x$)	.08685	.08878	.09062	.08878	.79352	.09487	.08874

*N=5,000; balanced; unrestricted; errors drawn from chi-square distribution.

Table C. 67 PSE Data - Monte Carlo MSE 31 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	80.21	80.21	80.14	80.21	79.98	80.23	80.21
PSE	92.79	92.57	92.65	92.57	95.08	92.99	92.58
Non-PSE	39.51	40.23	39.68	40.23	31.09	38.93	40.22
Hold-Out	80.20	80.20	80.03	80.20	79.89	80.18	80.20
PSE	92.94	92.75	92.48	92.75	94.98	93.07	92.75
Non-PSE	37.31	37.95	38.09	37.95	29.06	36.79	37.93
YRMSE	44.84	64.31	59.55	64.31	61.22	55.78	64.18
$YRMSE_0$	31.92	46.33	42.73	46.33	43.95	40.02	46.23
PRMSE	1.26	1.34	1.82	1.34	2.87	1.26	1.34
$PRMSE_0$	0.91	0.98	1.32	0.98	2.06	0.91	0.98
$MSE(\hat{\beta})$	.18510	4.01678	4.84180	4.01678	3.04673	2.14615	3.98583
$RMSE(\hat{\beta})$	.43023	2.00419	2.20041	2.00419	1.74549	1.46498	1.99645
$VAR(\hat{\beta})$	.18354	.60446	.51166	.60446	.01221	.39419	.60029
$Bias^2(\hat{\beta})$	.00156	3.41232	4.33014	3.41232	3.03451	1.75197	3.38554
MSE ($\partial p / \partial x$)	.02603	.03483	.05187	.03483	.48525	.02415	.03447
RMSE ($\partial p / \partial x$)	.16133	.18662	.22775	.18662	.69660	.15539	.18567

*N=2,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table C. 68 PSE Data - Monte Carlo MSE 32 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	84.50	84.54	84.20	84.54	83.66	84.50	84.55
PSE	94.63	94.30	94.38	94.30	96.83	94.78	94.31
Non-PSE	45.90	47.34	45.37	47.34	33.41	45.28	47.32
Hold-Out	84.44	84.48	83.99	84.48	83.44	84.43	84.48
PSE	94.71	94.45	94.09	94.45	96.78	94.81	94.45
Non-PSE	43.89	45.13	44.15	45.13	30.74	43.42	45.11
YRMSE	46.55	84.87	75.13	84.87	60.33	67.16	84.57
$YRMSE_0$	33.16	61.29	54.00	61.29	43.27	48.27	61.06
PRMSE	1.71	1.39	2.42	1.39	4.41	1.83	1.39
$PRMSE_0$	1.23	1.03	1.79	1.03	3.12	1.32	1.03
$MSE(\hat{\beta})$	.51075	9.39570	11.1590	9.39570	2.88271	4.63806	9.30255
$RMSE(\hat{\beta})$	.71467	3.06524	3.34051	3.06524	1.69785	2.15362	3.05001
$VAR(\hat{\beta})$	.23773	.80232	.57568	.80232	.01042	.40976	.79103
$Bias^2(\hat{\beta})$	.27302	8.59337	10.5833	8.59337	2.87229	4.22831	8.51152
MSE ($\partial p / \partial x$)	.04242	.05970	.10658	.05970	.46920	.02618	.05874
RMSE ($\partial p / \partial x$)	.20596	.24433	.32647	.24433	.68498	.16179	.24237

*N=2,000; unbalanced; unrestricted; errors drawn from *t*- distribution.

Table C. 69 PSE Data - Monte Carlo MSE 33 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	81.95	81.97	81.88	81.97	81.64	81.95	81.97
PSE	91.17	91.08	91.15	91.08	93.10	91.38	91.08
Non-PSE	53.17	53.55	52.92	53.55	45.83	52.51	53.53
Hold-Out	81.85	81.86	81.76	81.86	81.57	81.86	81.86
PSE	91.62	91.52	91.10	91.52	93.09	91.72	91.53
Non-PSE	50.35	50.69	51.63	50.68	44.41	50.05	50.67
YRMSE	49.40	89.68	74.02	89.53	60.47	68.68	88.84
$YRMSE_0$	35.23	64.89	53.31	64.78	43.32	49.48	64.29
PRMSE	45.68	45.57	45.38	45.57	47.08	45.87	45.57
$PRMSE_0$	33.23	33.15	33.01	33.15	34.34	33.38	33.15
$MSE(\hat{\beta})$	1.55804	14.9054	13.7407	14.7082	2.51781	7.20312	14.2095
$RMSE(\hat{\beta})$	1.24822	3.86075	3.70684	3.83512	1.58676	2.68386	3.76955
$VAR(\hat{\beta})$	.31635	1.42105	.58710	1.25666	.01140	.43034	.94089
$Bias^2(\hat{\beta})$	1.24170	13.4843	13.1536	13.4515	2.50642	6.77277	13.2686
MSE ($\partial p / \partial x$)	.20873	.28481	.25742	.27600	.43096	.09191	.25661
RMSE ($\partial p / \partial x$)	.45687	.53368	.50737	.52535	.65647	.30317	.50656

*N=2,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table C. 70 PSE Data - Monte Carlo MSE 41 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	73.63	73.64	73.61	73.64	73.61	73.64	73.64
PSE	70.63	70.79	69.39	70.79	69.55	70.64	70.79
Non-PSE	76.58	76.45	77.78	76.45	77.62	76.60	76.45
Hold-Out	73.07	73.06	73.07	73.06	73.08	73.07	73.06
PSE	70.03	70.17	68.93	70.17	69.07	70.03	70.17
Non-PSE	76.11	75.96	77.20	75.96	77.08	76.11	75.96
YRMSE	44.74	52.70	53.36	52.70	51.49	50.60	52.67
$YRMSE_0$	31.79	37.16	37.64	37.16	36.17	35.76	37.15
PRMSE	1.38	1.42	1.52	1.42	2.07	1.38	1.42
$PRMSE_0$	0.96	0.99	1.07	0.99	1.44	0.96	0.99
$MSE(\hat{\beta})$	.13911	4.76300	5.65469	4.76300	3.82088	3.52324	4.74805
$RMSE(\hat{\beta})$	.37298	2.18243	2.37796	2.18243	1.95471	1.87703	2.17900
$VAR(\hat{\beta})$	.13858	.40895	.37531	.40895	.01351	.34685	.40820
$Bias^2(\hat{\beta})$	.00053	4.35405	5.27938	4.35405	3.80737	3.17639	4.33985
MSE ($\partial p / \partial x$)	.01581	.01702	.01586	.01702	.57491	.01640	.01699
RMSE ($\partial p / \partial x$)	.12573	.13044	.12594	.13044	.75823	.12806	.13035

*N=2,000; balanced; unrestricted; errors drawn from normal distribution.

Table C. 71 PSE Data - Monte Carlo MSE 42 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.22	79.25	79.15	79.25	79.17	79.25	79.25
PSE	76.17	76.46	74.38	76.46	74.60	76.20	76.45
Non-PSE	82.13	81.90	83.69	81.90	83.52	82.15	81.91
Hold-Out	78.91	78.92	78.84	78.92	78.85	78.92	78.93
PSE	75.52	75.74	74.02	75.74	74.18	75.51	75.73
Non-PSE	82.21	82.04	83.52	82.04	83.39	82.25	82.04
YRMSE	46.21	68.48	69.34	68.48	50.06	62.26	68.40
$YRMSE_0$	33.11	47.96	48.62	47.96	35.54	43.79	47.90
PRMSE	1.57	1.40	1.70	1.40	3.55	1.55	1.40
$PRMSE_0$	1.11	0.98	1.18	0.98	2.49	1.09	0.98
$MSE(\hat{\beta})$	1.19719	16.6062	18.6658	16.6062	3.08510	11.8753	16.5441
$RMSE(\hat{\beta})$	1.09416	4.07507	4.32039	4.07507	1.75644	3.44606	4.06744
$VAR(\hat{\beta})$	.16205	.49526	.42072	.49526	.01070	.37826	.49364
$Bias^2(\hat{\beta})$	1.03514	16.1109	18.2451	16.1109	3.07440	11.4971	16.0505
MSE ($\partial p / \partial x$)	.04240	.03396	.02148	.03396	.50539	.02301	.03382
RMSE ($\partial p / \partial x$)	.20592	.18428	.14656	.18428	.71091	.15170	.18391

*N=2,000; balanced; unrestricted; errors drawn from *t*- distribution.

Table C. 72 PSE Data - Monte Carlo MSE 43 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	75.49	75.54	75.65	75.54	75.65	75.56	75.54
PSE	69.87	69.66	67.94	69.66	68.13	69.47	69.66
Non-PSE	80.40	80.67	82.38	80.67	82.21	80.87	80.67
Hold-Out	75.00	75.02	75.12	75.02	75.10	75.03	75.02
PSE	69.64	69.44	67.88	69.44	68.03	69.25	69.43
Non-PSE	79.68	79.90	81.45	79.90	81.29	80.09	79.91
YRMSE	45.16	57.38	59.60	57.38	51.07	54.32	57.35
$YRMSE_0$	32.10	40.35	41.91	40.35	35.93	38.29	40.32
PRMSE	20.92	20.82	20.83	20.82	21.92	20.94	20.82
$PRMSE_0$	13.48	13.41	13.41	13.41	14.12	13.49	13.41
$MSE(\hat{\beta})$	.25136	7.36610	8.94426	7.36610	3.59946	5.42454	7.34222
$RMSE(\hat{\beta})$	.50136	2.71406	2.99069	2.71406	1.89722	2.32906	2.70965
$VAR(\hat{\beta})$	.14251	.43069	.39931	.43069	.01248	.35614	.42976
$Bias^2(\hat{\beta})$	.10886	6.93541	8.54495	6.93541	3.58698	5.06840	6.91246
MSE ($\partial p / \partial x$)	.01535	.01796	.01858	.01796	.55583	.01868	.01795
RMSE ($\partial p / \partial x$)	.12390	.13401	.13629	.13401	.74554	.13666	.13399

•N=2,000; balanced; unrestricted; errors drawn from chi-square distribution.

Table C. 73 PSE Data - Monte Carlo MSE 51 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	80.51	80.54	80.43	80.54	80.20	80.55	80.54
PSE	93.24	93.05	93.29	93.04	95.84	93.52	93.05
Non-PSE	38.01	38.76	37.46	38.79	27.96	37.22	38.77
Hold-Out	80.05	80.06	79.92	80.05	79.79	80.05	80.05
PSE	93.07	92.86	93.08	92.86	95.41	93.30	92.87
Non-PSE	36.84	37.56	36.25	37.53	27.90	36.06	37.51
YRMSE	32.21	50.17	41.93	49.44	44.03	39.99	47.67
$YRMSE_0$	32.68	51.76	42.27	50.90	44.65	40.63	48.83
PRMSE	1.27	1.31	1.65	1.32	2.30	1.26	1.32
$PRMSE_0$	1.30	1.33	1.65	1.35	2.39	1.28	1.35
$MSE(\hat{\beta})$	.77778	9.51921	5.08196	8.49106	3.67120	2.84953	6.26449
$RMSE(\hat{\beta})$	.88192	3.08532	2.25432	2.91394	1.91604	1.68806	2.50290
$VAR(\hat{\beta})$	.74910	4.58188	1.16433	3.69433	.02751	.86623	1.80963
$Bias^2(\hat{\beta})$	.02867	4.93733	3.91762	4.79672	3.64369	1.98331	4.45486
MSE ($\partial p / \partial x$)	.11330	.27190	.11064	.21973	.64153	.05564	.11735
RMSE ($\partial p / \partial x$)	.33660	.52144	.33263	.46876	.80095	.23588	.34256

*N=1,000; unbalanced; unrestricted; errors drawn from normal distribution.

Table C. 74 PSE Data - Monte Carlo MSE 52 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	84.67	84.76	84.37	84.75	83.86	84.68	84.75
PSE	94.99	94.69	94.83	94.69	97.39	95.24	94.70
Non-PSE	43.71	45.32	42.81	45.31	30.10	42.74	45.28
Hold-Out	84.32	84.36	84.12	84.36	83.66	84.31	84.36
PSE	94.82	94.53	94.71	94.53	97.07	95.03	94.54
Non-PSE	43.07	44.39	42.48	44.38	30.92	42.14	44.35
YRMSE	34.98	70.37	52.81	68.26	43.55	48.20	63.72
$YRMSE_0$	35.68	73.02	52.90	70.54	44.13	48.74	65.19
PRMSE	1.45	1.30	1.90	1.30	3.25	1.51	1.30
$PRMSE_0$	1.50	1.34	1.88	1.34	3.35	1.57	1.34
$MSE(\hat{\beta})$	2.22803	25.1001	10.5712	21.1661	3.55025	5.35358	13.7702
$RMSE(\hat{\beta})$	1.49266	5.01000	3.25134	4.60066	1.88421	2.31378	3.71083
$VAR(\hat{\beta})$	1.69304	11.8908	1.28182	8.65186	.02212	.88811	2.95461
$Bias^2(\hat{\beta})$	.53499	13.2093	9.28939	12.5142	3.52814	4.46546	10.8156
MSE ($\partial p / \partial x$)	.26820	.68536	.17095	.52292	.62929	.06100	.22943
RMSE ($\partial p / \partial x$)	.51788	.82786	.41346	.72313	.79328	.24698	.47899

*N=1,000; unbalanced; unrestricted; errors drawn from *t*- distribution.

Table C. 75 PSE Data - Monte Carlo MSE 53 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	82.29	82.34	82.12	82.34	81.78	82.30	82.34
PSE	91.49	91.38	91.73	91.38	93.83	91.80	91.39
Non-PSE	52.49	53.02	50.99	53.02	42.67	51.52	53.00
Hold-Out	81.58	81.56	81.38	81.57	81.08	81.53	81.57
PSE	91.44	91.28	91.83	91.29	93.73	91.76	91.30
Non-PSE	50.41	50.85	48.38	50.85	41.09	49.22	50.82
YRMSE	37.28	74.95	51.73	72.88	43.56	49.59	68.40
$YRMSE_0$	38.17	78.13	52.33	75.72	44.04	50.61	70.50
PRMSE	32.39	32.32	32.14	32.32	33.35	32.54	32.33
$PRMSE_0$	32.34	32.28	32.10	32.28	33.45	32.52	32.28
$MSE(\hat{\beta})$	3.52909	31.7908	13.6979	27.6519	3.11607	8.29979	19.8938
$RMSE(\hat{\beta})$	1.87859	5.63833	3.70107	5.25851	1.76524	2.88094	4.46025
$VAR(\hat{\beta})$	1.58709	11.7168	1.31770	8.45424	.02496	.88208	2.85942
$Bias^2(\hat{\beta})$	1.94201	20.0740	12.3802	19.1977	3.09112	7.41771	17.0344
MSE ($\partial p / \partial x$)	.52187	1.12499	.32255	.91295	.58148	.14027	.53487
RMSE ($\partial p / \partial x$)	.72241	1.06065	.56793	.95549	.76255	.37453	.73135

*N=1,000; unbalanced; unrestricted; errors drawn from chi-square distribution.

Table C. 76 PSE Data - Monte Carlo MSE 61 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	74.25	74.28	74.25	74.28	74.26	74.28	74.28
PSE	70.93	71.10	69.69	71.10	69.83	70.94	71.10
Non-PSE	77.51	77.40	78.76	77.40	78.63	77.57	77.40
Hold-Out	74.05	74.03	74.06	74.03	74.07	74.04	74.04
PSE	71.39	71.52	70.28	71.52	70.41	71.37	71.52
Non-PSE	76.78	76.62	77.93	76.62	77.81	76.78	76.63
YRMSE	31.67	38.08	38.51	38.08	36.71	36.33	38.05
$YRMSE_0$	31.97	38.70	38.95	38.70	36.84	36.88	38.67
PRMSE	1.38	1.42	1.45	1.42	1.76	1.38	1.41
$PRMSE_0$	1.36	1.40	1.43	1.40	1.78	1.36	1.40
$MSE(\hat{\beta})$	.28684	5.32464	5.99729	5.32464	3.74101	3.92316	5.30748
$RMSE(\hat{\beta})$	.53558	2.30752	2.44894	2.30752	1.93417	1.98070	2.30380
$VAR(\hat{\beta})$	.28271	.84136	.76094	.84136	.02633	.70590	.83970
$Bias^2(\hat{\beta})$	.00414	4.48328	5.23635	4.48328	3.71469	3.21726	4.46778
MSE ($\partial p / \partial x$)	.03636	.04142	.03614	.04142	.65185	.03498	.04132
RMSE ($\partial p / \partial x$)	.19068	.20353	.19010	.20353	.80737	.18702	.20326

*N=1,000; balanced; unrestricted; errors drawn from normal distribution.

Table C. 77 PSE Data - Monte Carlo MSE 62 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	79.83	79.88	79.75	79.88	79.78	79.88	79.88
PSE	76.49	76.76	74.77	76.76	74.98	76.51	76.76
Non-PSE	83.03	82.87	84.55	82.87	84.39	83.12	82.87
Hold-Out	79.63	79.64	79.62	79.64	79.64	79.65	79.64
PSE	76.79	77.01	75.30	77.01	75.46	76.78	77.01
Non-PSE	82.47	82.27	83.92	82.27	83.78	82.52	82.28
YRMSE	33.00	49.98	50.39	49.98	35.79	44.99	49.92
$YRMSE_0$	33.90	51.05	51.11	51.05	36.43	45.96	50.98
PRMSE	1.46	1.38	1.51	1.38	2.69	1.44	1.38
$PRMSE_0$	1.45	1.36	1.49	1.36	2.76	1.43	1.36
$MSE(\hat{\beta})$	1.34400	16.7746	18.1626	16.7746	3.08791	11.7385	16.7068
$RMSE(\hat{\beta})$	1.15931	4.09568	4.26176	4.09568	1.75725	3.42615	4.08739
$VAR(\hat{\beta})$	.34606	1.06769	.87486	1.06769	.02153	.80185	1.06391
$Bias^2(\hat{\beta})$	.99794	15.7069	17.2877	15.7069	3.06638	10.9367	15.6428
MSE ($\partial p / \partial x$)	.09661	.10968	.08122	.10968	.58421	.06784	.10912
RMSE ($\partial p / \partial x$)	.31083	.33118	.28498	.33118	.76434	.26046	.33033

*N=1,000; balanced; unrestricted; errors drawn from t- distribution.

Table C. 78 PSE Data - Monte Carlo MSE 63 (See Table 2.4 for design)

	Probit	Logit	DA	ME	GME1	GME2	GME3
% Correctly Predicted							
In-Sample	76.17	76.23	76.35	76.23	76.36	76.25	76.23
PSE	70.89	70.72	69.22	70.72	69.39	70.52	70.71
Non-PSE	80.82	81.09	82.64	81.09	82.51	81.29	81.09
Hold-Out	75.80	75.83	75.99	75.83	75.98	75.85	75.83
PSE	71.14	70.93	69.61	70.93	69.75	70.76	70.93
Non-PSE	80.05	80.29	81.79	80.29	81.64	80.48	80.29
YRMSE	32.06	41.68	43.11	41.68	36.38	39.15	41.64
$YRMSE_0$	32.43	42.36	43.63	42.36	36.63	39.75	42.33
PRMSE	15.79	15.72	15.72	15.72	16.52	15.81	15.72
$PRMSE_0$	15.80	15.73	15.72	15.73	16.57	15.83	15.73
$MSE(\hat{\beta})$	.43100	7.99523	9.27270	7.99523	3.52750	5.83269	7.96816
$RMSE(\hat{\beta})$	.65650	2.82758	3.04511	2.82758	1.87816	2.41510	2.82279
$VAR(\hat{\beta})$	.30011	.90941	.80953	.90941	.02491	.74230	.90729
$Bias^2(\hat{\beta})$	.13088	7.08582	8.46318	7.08582	3.50259	5.09039	7.06087
MSE ($\partial p / \partial x$)	.03751	.03997	.03286	.03997	.63123	.03352	.03987
RMSE ($\partial p / \partial x$)	.19368	.19993	.18127	.19993	.79450	.18309	.19967

*N=1,000; balanced; unrestricted; errors drawn from chi-square distribution.

Table C. 79 PSE Data - Monte Carlo MSE 71 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	79.78	79.79	79.79	79.55	79.79	79.79
PSE	93.53	93.30	93.30	95.74	93.65	93.30
Non-PSE	34.64	35.42	35.42	26.36	34.29	35.41
Hold-Out	80.27	80.27	80.27	79.99	80.28	80.27
PSE	93.31	93.03	93.03	95.68	93.44	93.04
Non-PSE	36.16	37.09	37.09	26.87	35.74	37.07
YRMSE	70.78	97.51	97.51	94.30	87.03	97.36
$YRMSE_0$	31.68	44.07	44.07	42.63	39.19	44.00
PRMSE	1.28	1.43	1.43	3.99	1.30	1.42
$PRMSE_0$	0.58	0.65	0.65	1.79	0.58	0.65
$MSE(\hat{\beta})$	.06548	2.89678	2.89678	2.43023	1.65367	2.87865
$RMSE(\hat{\beta})$	.25589	1.70199	1.70199	1.55892	1.28595	1.69666
$VAR(\hat{\beta})$	.06511	.20449	.20449	.00533	.15181	.20368
$Bias^2(\hat{\beta})$	.00037	2.69229	2.69229	2.42490	1.50187	2.67497
MSE ($\partial p / \partial x$)	.00971	.01432	.01432	.42159	.00972	.01417
RMSE ($\partial p / \partial x$)	.09852	.11966	.11966	.64930	.09859	.11904

*N=5,000; unbalanced; restricted; errors drawn from normal distribution.

Table C. 80 PSE Data - Monte Carlo MSE 72 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	84.17	84.20	84.20	83.38	84.16	84.20
PSE	95.31	94.92	94.92	97.48	95.40	94.93
Non-PSE	40.18	41.88	41.88	27.70	39.78	41.86
Hold-Out	84.44	84.47	84.47	83.64	84.42	84.47
PSE	95.23	94.76	94.76	97.49	95.33	94.77
Non-PSE	41.07	43.13	43.13	27.98	40.58	43.09
YRMSE	72.95	128.35	128.35	92.63	105.17	127.99
$YRMSE_0$	32.69	58.21	58.21	41.88	47.48	58.04
PRMSE	2.19	1.57	1.57	6.46	2.39	1.58
$PRMSE_0$	0.99	0.71	0.71	2.92	1.08	0.71
$MSE(\hat{\beta})$	.27597	7.05054	7.05055	2.32254	3.67324	6.99342
$RMSE(\hat{\beta})$	.52533	2.65529	2.65529	1.52399	1.91657	2.64451
$VAR(\hat{\beta})$	.08010	.26376	.26376	.00473	.16427	.26183
$Bias^2(\hat{\beta})$	.19586	6.78678	6.78678	2.31780	3.50897	6.73160
MSE ($\partial p / \partial x$)	.02487	.04698	.04698	.41052	.01449	.04632
RMSE ($\partial p / \partial x$)	.15770	.21676	.21676	.64072	.12037	.21523

*N=5,000; unbalanced; restricted; errors drawn from *t*- distribution.

Table C. 81 PSE Data - Monte Carlo MSE 73 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	81.53	81.55	81.55	81.18	81.54	81.55
PSE	91.06	90.93	90.93	93.44	91.34	90.94
Non-PSE	51.28	51.74	51.74	42.27	50.41	51.73
Hold-Out	82.01	82.02	82.02	81.67	82.01	82.01
PSE	91.18	91.06	91.06	93.15	91.38	91.06
Non-PSE	52.09	52.49	52.49	44.13	51.39	52.47
YRMSE	76.90	134.27	134.27	93.05	108.07	133.86
$YRMSE_0$	34.64	61.25	61.25	42.06	49.10	61.05
PRMSE	69.00	68.82	68.82	71.39	69.30	68.83
$PRMSE_0$	31.74	31.66	31.66	32.86	31.89	31.67
$MSE(\hat{\beta})$	1.19691	11.7649	11.7649	1.92572	6.60128	11.6775
$RMSE(\hat{\beta})$	1.09403	3.43000	3.43000	1.38770	2.56930	3.41723
$VAR(\hat{\beta})$	.07910	.25722	.25722	.00492	.16276	.25537
$Bias^2(\hat{\beta})$	1.11780	11.5077	11.5077	1.92080	6.43852	11.4221
MSE ($\partial p / \partial x$)	.17803	.23877	.23877	.36739	.09986	.23630
RMSE ($\partial p / \partial x$)	.42193	.48864	.48864	.60613	.31600	.48611

*N=5,000; unbalanced; restricted; errors drawn from chi-square distribution.

Table C. 82 PSE Data - Monte Carlo MSE 81 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	73.53	73.54	73.54	73.52	73.53	73.54
PSE	70.27	70.43	70.43	69.35	70.29	70.43
Non-PSE	76.77	76.63	76.63	77.68	76.75	76.63
Hold-Out	72.73	72.73	72.73	72.73	72.74	72.73
PSE	68.55	68.73	68.73	67.50	68.58	68.72
Non-PSE	76.87	76.69	76.69	77.89	76.85	76.69
YRMSE	70.71	82.66	82.66	81.18	79.57	82.62
$YRMSE_0$	31.72	36.74	36.74	36.05	35.45	36.73
PRMSE	1.38	1.46	1.46	2.79	1.39	1.46
$PRMSE_0$	0.63	0.66	0.66	1.19	0.63	0.66
$MSE(\hat{\beta})$	.05347	4.46502	4.46504	3.79776	3.29111	4.45096
$RMSE(\hat{\beta})$	.23124	2.11306	2.11306	1.94878	1.81414	2.10973
$VAR(\hat{\beta})$	.05321	.15657	.15657	.00519	.13307	.15629
$Bias^2(\hat{\beta})$	.00026	4.30845	4.30847	3.79257	3.15804	4.29467
MSE ($\partial p / \partial x$)	.00669	.00759	.00759	.64692	.00752	.00756
RMSE ($\partial p / \partial x$)	.08181	.08711	.08711	.80431	.08672	.08696

*N=5,000; balanced; restricted; errors drawn from normal distribution.

Table C. 83 PSE Data - Monte Carlo MSE 82 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	79.26	79.26	79.26	79.22	79.26	79.26
PSE	75.96	76.20	76.20	74.61	75.98	76.20
Non-PSE	82.44	82.22	82.22	83.67	82.43	82.22
Hold-Out	78.44	78.44	78.44	78.37	78.43	78.44
PSE	74.18	74.48	74.48	72.62	74.20	74.48
Non-PSE	82.50	82.23	82.23	83.85	82.48	82.23
YRMSE	73.18	107.51	107.51	79.08	98.14	107.39
$YRMSE_0$	32.66	47.20	47.20	35.06	43.23	47.15
PRMSE	1.88	1.50	1.50	5.36	1.86	1.50
$PRMSE_0$	0.84	0.68	0.68	2.29	0.82	0.68
$MSE(\hat{\beta})$	1.04609	15.9629	15.9629	3.09457	11.4317	15.9039
$RMSE(\hat{\beta})$	1.02278	3.99536	3.99536	1.75914	3.38107	3.98796
$VAR(\hat{\beta})$	.06646	.20404	.20404	.00432	.15561	.20337
$Bias^2(\hat{\beta})$	.97963	15.7589	15.7589	3.09025	11.2760	15.7005
MSE ($\partial p / \partial x$)	.05477	.05759	.05759	.57529	.03196	.05725
RMSE ($\partial p / \partial x$)	.23402	.23997	.23997	.75848	.17878	.23927

*N=5,000; balanced; restricted; errors drawn from *t*- distribution.

Table C. 84 PSE Data - Monte Carlo MSE 83 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	75.45	75.47	75.47	75.56	75.48	75.47
PSE	69.71	69.54	69.54	68.38	69.39	69.54
Non-PSE	80.49	80.69	80.69	81.88	80.83	80.69
Hold-Out	74.68	74.70	74.70	74.80	74.72	74.70
PSE	67.81	67.60	67.60	66.46	67.47	67.60
Non-PSE	80.64	80.86	80.86	82.03	81.01	80.87
YRMSE	71.41	89.98	89.98	80.60	85.42	89.93
$YRMSE_0$	32.01	39.89	39.89	35.74	37.97	39.87
PRMSE	33.18	33.00	33.00	34.71	33.20	33.00
$PRMSE_0$	14.15	14.07	14.07	14.75	14.15	14.07
$MSE(\hat{\beta})$	.14526	6.80643	6.80643	3.61119	5.00014	6.78433
$RMSE(\hat{\beta})$	.38112	2.60891	2.60891	1.90031	2.23610	2.60467
$VAR(\hat{\beta})$	.05470	.16850	.16849	.00484	.13920	.16813
$Bias^2(\hat{\beta})$	.09055	6.63793	6.63794	3.60635	4.86094	6.61620
MSE ($\partial p / \partial x$)	.00754	.00788	.00788	.62967	.00900	.00788
RMSE ($\partial p / \partial x$)	.08685	.08878	.08878	.79352	.09487	.08874

*N=5,000; balanced; restricted; errors drawn from chi-square distribution.

Table C.85 PSE Data - Monte Carlo MSE 91 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	80.21	80.21	80.21	79.98	80.23	80.21
PSE	92.79	92.57	92.57	95.09	92.99	92.58
Non-PSE	39.51	40.22	40.22	31.08	38.93	40.21
Hold-Out	80.20	80.20	80.20	79.89	80.19	80.20
PSE	92.94	92.75	92.75	94.97	93.07	92.75
Non-PSE	37.33	37.97	37.97	29.10	36.81	37.95
YRMSE	44.84	64.30	64.30	61.22	55.77	64.18
$YRMSE_0$	31.91	46.32	46.32	43.94	40.01	46.23
PRMSE	1.25	1.33	1.33	2.87	1.25	1.33
$PRMSE_0$	0.90	0.97	0.97	2.05	0.90	0.97
$MSE(\hat{\beta})$	.18311	4.02765	4.02766	3.04291	2.15300	3.99665
$RMSE(\hat{\beta})$	.42792	2.00690	2.00690	1.74439	1.46731	1.99916
$VAR(\hat{\beta})$	.18145	.59842	.59842	.01203	.38923	.59426
$Bias^2(\hat{\beta})$	.00166	3.42924	3.42924	3.03089	1.76377	3.40238
MSE ($\partial p / \partial x$)	.02572	.03461	.03461	.48487	.02370	.03425
RMSE ($\partial p / \partial x$)	.16036	.18603	.18603	.69632	.15395	.18507

*N=2,000; unbalanced; restricted; errors drawn from normal distribution.

Table C. 86 PSE Data - Monte Carlo MSE 92 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	84.51	84.54	84.54	83.66	84.50	84.55
PSE	94.63	94.30	94.30	96.83	94.78	94.31
Non-PSE	45.90	47.35	47.35	33.40	45.28	47.32
Hold-Out	84.45	84.48	84.48	83.45	84.43	84.48
PSE	94.71	94.45	94.45	96.78	94.81	94.45
Non-PSE	43.90	45.13	45.13	30.75	43.44	45.11
YRMSE	46.55	84.87	84.87	60.33	67.16	84.57
$YRMSE_0$	33.16	61.29	61.29	43.27	48.27	61.06
PRMSE	1.70	1.38	1.38	4.41	1.83	1.39
$PRMSE_0$	1.23	1.03	1.03	3.12	1.32	1.03
$MSE(\hat{\beta})$	.51086	9.40553	9.40553	2.88061	4.64457	9.31233
$RMSE(\hat{\beta})$	.71474	3.06684	3.06684	1.69724	2.15513	3.05161
$VAR(\hat{\beta})$	.23597	.79775	.79775	.01029	.40614	.78647
$Bias^2(\hat{\beta})$	.27489	8.60778	8.60778	2.87032	4.23843	8.52586
MSE ($\partial p / \partial x$)	.04237	.05972	.05972	.46899	.02602	.05876
RMSE ($\partial p / \partial x$)	.20585	.24438	.24438	.68483	.16132	.24241

*N=2,000; unbalanced; restricted; errors drawn from t- distribution.

Table C. 87 PSE Data - Monte Carlo MSE 93 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	81.95	81.97	81.97	81.64	81.95	81.97
PSE	91.17	91.08	91.08	93.10	91.38	91.08
Non-PSE	53.17	53.55	53.55	45.83	52.51	53.53
Hold-Out	81.85	81.86	81.86	81.57	81.86	81.86
PSE	91.62	91.52	91.52	93.09	91.72	91.53
Non-PSE	50.34	50.69	50.69	44.41	50.05	50.68
YRMSE	49.40	89.68	89.53	60.47	68.68	88.84
$YRMSE_0$	35.23	64.89	64.78	43.32	49.48	64.29
PRMSE	45.68	45.57	45.57	47.08	45.87	45.57
$PRMSE_0$	33.23	33.15	33.15	34.34	33.38	33.15
$MSE(\hat{\beta})$	1.55953	14.9054	14.7195	2.51611	7.21100	14.2207
$RMSE(\hat{\beta})$	1.24881	3.86075	3.83660	1.58623	2.68533	3.77104
$VAR(\hat{\beta})$	.31546	1.42105	1.25666	.01130	.42808	.93818
$Bias^2(\hat{\beta})$	1.24407	13.4843	13.4656	2.50481	6.78292	13.2826
MSE ($\partial p / \partial x$)	.20910	.28481	.27651	.43078	.09213	.25711
RMSE ($\partial p / \partial x$)	.45727	.53368	.52584	.65634	.30353	.50706

*N=2,000; unbalanced; restricted; errors drawn from chi-square distribution.

Table C. 88 PSE Data - Monte Carlo MSE 101 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	73.63	73.64	73.64	73.61	73.64	73.64
PSE	70.63	70.79	70.79	69.55	70.64	70.79
Non-PSE	76.58	76.44	76.44	77.62	76.60	76.45
Hold-Out	73.07	73.06	73.06	73.08	73.07	73.06
PSE	70.03	70.17	70.17	69.07	70.03	70.17
Non-PSE	76.11	75.96	75.96	77.08	76.11	75.96
YRMSE	44.74	52.70	52.70	51.49	50.60	52.67
$YRMSE_0$	31.79	37.16	37.16	36.17	35.76	37.15
PRMSE	1.37	1.42	1.42	2.07	1.38	1.42
$PRMSE_0$	0.96	0.99	0.99	1.44	0.96	0.99
$MSE(\hat{\beta})$	.13904	4.76318	4.76320	3.82084	3.52336	4.74825
$RMSE(\hat{\beta})$	.37288	2.18247	2.18248	1.95470	1.87706	2.17905
$VAR(\hat{\beta})$	.13851	.40856	.40856	.01350	.34653	.40781
$Bias^2(\hat{\beta})$	.00053	4.35462	4.35464	3.80734	3.17683	4.34044
MSE ($\partial p / \partial x$)	.01580	.01700	.01700	.57491	.01638	.01697
RMSE ($\partial p / \partial x$)	.12570	.13038	.13038	.75823	.12800	.13028

*N=2,000; balanced; restricted; errors drawn from normal distribution.

Table C. 89 PSE Data - Monte Carlo MSE 102 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	79.22	79.25	79.25	79.17	79.25	79.25
PSE	76.17	76.46	76.46	74.60	76.20	76.45
Non-PSE	82.13	81.90	81.90	83.52	82.15	81.91
Hold-Out	78.91	78.92	78.92	78.85	78.92	78.93
PSE	75.52	75.74	75.74	74.18	75.51	75.73
Non-PSE	82.21	82.04	82.04	83.39	82.25	82.04
YRMSE	46.21	68.48	68.48	50.06	62.26	68.40
$YRMSE_0$	33.11	47.96	47.96	35.54	43.79	47.90
PRMSE	1.57	1.40	1.40	3.55	1.55	1.40
$PRMSE_0$	1.11	0.98	0.98	2.49	1.09	0.98
$MSE(\hat{\beta})$	1.19718	16.6062	16.6062	3.08510	11.8753	16.5441
$RMSE(\hat{\beta})$	1.09416	4.07507	4.07507	1.75644	3.44606	4.06744
$VAR(\hat{\beta})$	.16205	.49526	.49526	.01070	.37826	.49364
$Bias^2(\hat{\beta})$	1.03513	16.1109	16.1109	3.07440	11.4971	16.0505
MSE ($\partial p / \partial x$)	.04240	.03396	.03396	.50539	.02301	.03382
RMSE ($\partial p / \partial x$)	.20592	.18428	.18428	.71091	.15170	.18391

*N=2,000; balanced; restricted; errors drawn from *t*- distribution.

Table C. 90 PSE Data - Monte Carlo MSE 103 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	75.49	75.54	75.54	75.65	75.56	75.54
PSE	69.87	69.66	69.66	68.13	69.47	69.65
Non-PSE	80.40	80.67	80.67	82.21	80.87	80.67
Hold-Out	75.00	75.02	75.02	75.10	75.03	75.02
PSE	69.65	69.44	69.44	68.03	69.25	69.44
Non-PSE	79.68	79.90	79.90	81.29	80.09	79.90
YRMSE	45.16	57.38	57.38	51.07	54.32	57.35
$YRMSE_0$	32.10	40.35	40.35	35.93	38.29	40.32
PRMSE	20.92	20.82	20.82	21.92	20.94	20.82
$PRMSE_0$	13.48	13.41	13.41	14.12	13.49	13.41
$MSE(\hat{\beta})$	.25072	7.36695	7.36696	3.59918	5.42500	7.34307
$RMSE(\hat{\beta})$	.50072	2.71421	2.71421	1.89715	2.32916	2.70981
$VAR(\hat{\beta})$	.14168	.42787	.42787	.01243	.35402	.42695
$Bias^2(\hat{\beta})$	.10905	6.93908	6.93909	3.58675	5.07097	6.91612
MSE ($\partial p / \partial x$)	.01525	.01782	.01782	.55581	.01856	.01781
RMSE ($\partial p / \partial x$)	.12348	.13348	.13348	.74552	.13622	.13347

*N=2,000; balanced; restricted; errors drawn from chi-square distribution

Table C. 91 PSE Data - Monte Carlo MSE 111 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	80.51	80.54	80.54	80.20	80.55	80.54
PSE	93.24	93.05	93.05	95.84	93.52	93.05
Non-PSE	37.98	38.76	38.76	27.93	37.21	38.73
Hold-Out	80.06	80.06	80.06	79.43	80.06	80.05
PSE	93.08	92.86	92.86	95.31	93.31	92.86
Non-PSE	36.86	37.56	37.56	27.96	36.08	37.53
YRMSE	32.21	50.17	49.43	44.03	39.98	47.66
$YRMSE_0$	32.68	51.76	50.89	44.38	40.62	48.82
PRMSE	1.26	1.31	1.31	2.29	1.24	1.31
$PRMSE_0$	1.28	1.33	1.33	2.32	1.27	1.33
$MSE(\hat{\beta})$	.76602	9.51921	8.48793	3.66325	2.84053	6.26117
$RMSE(\hat{\beta})$	.87523	3.08532	2.91341	1.91396	1.68539	2.50223
$VAR(\hat{\beta})$	.73665	4.58188	3.65632	.02602	.83466	1.77166
$Bias^2(\hat{\beta})$	.02937	4.93733	4.83161	3.63723	2.00587	4.48951
MSE ($\partial p / \partial x$)	.11247	.27190	.22126	.73823	.05331	.11638
RMSE ($\partial p / \partial x$)	.33536	.52144	.47038	.85921	.23089	.34115

*N=1,000; unbalanced; restricted; errors drawn from normal distribution.

Table C. 92 PSE Data - Monte Carlo MSE 112 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	84.67	84.76	84.76	83.86	84.68	84.76
PSE	94.99	94.69	94.69	97.39	95.24	94.70
Non-PSE	43.71	45.32	45.32	30.10	42.75	45.29
Hold-Out	84.33	84.36	84.36	83.66	84.31	84.36
PSE	94.82	94.53	94.53	97.07	95.03	94.54
Non-PSE	43.07	44.39	44.39	30.92	42.14	44.36
YRMSE	34.97	70.37	68.26	43.55	48.20	63.72
$YRMSE_0$	35.67	73.02	70.54	44.13	48.73	65.19
PRMSE	1.45	1.30	1.30	3.25	1.51	1.30
$PRMSE_0$	1.50	1.34	1.34	3.35	1.57	1.34
$MSE(\hat{\beta})$	2.21986	25.1001	21.1680	3.54886	5.35343	13.7722
$RMSE(\hat{\beta})$	1.48992	5.01000	4.60087	1.88384	2.31375	3.71109
$VAR(\hat{\beta})$	1.68425	11.8908	8.64314	.02187	.88173	2.94598
$Bias^2(\hat{\beta})$	.53561	13.2093	12.5249	3.52699	4.47171	10.8262
MSE ($\partial p / \partial x$)	.26711	.68536	.52275	.62914	.06064	.22925
RMSE ($\partial p / \partial x$)	.51683	.82786	.72301	.79319	.24625	.47880

*N=1,000; unbalanced; restricted; errors drawn from t- distribution.

Table C. 93 PSE Data - Monte Carlo MSE 113 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	82.29	82.34	82.34	81.78	82.30	82.34
PSE	91.49	91.38	91.38	93.84	91.80	91.39
Non-PSE	52.49	53.02	53.02	42.66	51.53	53.00
Hold-Out	81.57	81.56	81.56	81.08	81.53	81.56
PSE	91.44	91.28	91.28	93.74	91.75	91.29
Non-PSE	50.41	50.85	50.85	41.09	49.23	50.82
YRMSE	37.28	74.95	72.88	43.56	49.59	68.40
$YRMSE_0$	38.16	78.13	75.71	44.04	50.61	70.49
PRMSE	32.39	32.32	32.32	33.35	32.54	32.33
$PRMSE_0$	32.34	32.28	32.28	33.45	32.52	32.28
$MSE(\hat{\beta})$	3.52895	31.7908	27.6640	3.11358	8.30309	19.9051
$RMSE(\hat{\beta})$	1.87855	5.63833	5.25966	1.76453	2.88151	4.46152
$VAR(\hat{\beta})$	1.58211	11.7168	8.43683	.02446	.86867	2.84164
$Bias^2(\hat{\beta})$	1.94684	20.0740	19.2272	3.08912	7.43442	17.0635
MSE ($\partial p / \partial x$)	.52341	1.12499	.91635	.58121	.14033	.53679
RMSE ($\partial p / \partial x$)	.72347	1.06065	.95726	.76237	.37460	.73266

*N=1,000; unbalanced; restricted; errors drawn from chi-square distribution.

Table C. 94 PSE Data - Monte Carlo MSE 121 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	74.25	74.27	74.27	74.25	74.28	74.28
PSE	70.94	71.10	71.10	69.83	70.94	71.10
Non-PSE	77.51	77.39	77.39	78.62	77.56	77.39
Hold-Out	74.05	74.04	74.04	74.07	74.04	74.04
PSE	71.39	71.51	71.51	70.41	71.37	71.51
Non-PSE	76.78	76.63	76.63	77.81	76.79	76.63
YRMSE	31.67	38.07	38.07	36.71	36.33	38.05
$YRMSE_0$	31.97	38.69	38.69	36.84	36.87	38.67
PRMSE	1.37	1.41	1.41	1.75	1.37	1.41
$PRMSE_0$	1.35	1.39	1.39	1.78	1.35	1.39
$MSE(\hat{\beta})$	.28153	5.32924	5.32925	3.73734	3.92542	5.31207
$RMSE(\hat{\beta})$	.53059	2.30851	2.30852	1.93322	1.98127	2.30479
$VAR(\hat{\beta})$	.27694	.82355	.82355	.02581	.69101	.82192
$Bias^2(\hat{\beta})$	.00459	4.50569	4.50570	3.71153	3.23441	4.49015
MSE ($\partial p / \partial x$)	.03559	.04057	.04057	.65149	.03413	.04046
RMSE ($\partial p / \partial x$)	.18866	.20141	.20141	.80715	.18474	.20114

*N=1,000; balanced; restricted; errors drawn from normal distribution.

Table C. 95 PSE Data - Monte Carlo MSE 122 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	79.83	79.88	79.88	79.78	79.88	79.88
PSE	76.49	76.76	76.76	74.98	76.51	76.76
Non-PSE	83.04	82.87	82.87	84.39	83.12	82.87
Hold-Out	79.63	79.65	79.65	79.64	79.66	79.65
PSE	76.79	77.01	77.01	75.46	76.78	77.01
Non-PSE	82.48	82.28	82.28	83.79	82.53	82.28
YRMSE	33.00	49.98	49.98	35.79	44.99	49.91
$YRMSE_0$	33.90	51.04	51.04	36.43	45.96	50.98
PRMSE	1.46	1.38	1.38	2.69	1.44	1.38
$PRMSE_0$	1.44	1.35	1.35	2.76	1.42	1.35
$MSE(\hat{\beta})$	1.34400	16.7836	16.7836	3.08698	11.7445	16.7157
$RMSE(\hat{\beta})$	1.15931	4.09678	4.09678	1.75698	3.42702	4.08848
$VAR(\hat{\beta})$	.34367	1.06025	1.06025	.02140	.79621	1.05650
$Bias^2(\hat{\beta})$	1.00033	15.7233	15.7233	3.06558	10.9482	15.6592
MSE ($\partial p / \partial x$)	.09653	.10958	.10958	.58412	.06772	.10902
RMSE ($\partial p / \partial x$)	.31069	.33103	.33103	.76428	.26023	.33018

*N=1,000; balanced; restricted; errors drawn from t- distribution.

Table C. 96 PSE Data - Monte Carlo MSE 123 (See Table 2.4 for design)

	Probit	Logit	ME	GME1	GME2	GME3
% Correctly Predicted						
In-Sample	76.17	76.23	76.23	76.36	76.25	76.23
PSE	70.89	70.72	70.72	69.39	70.52	70.71
Non-PSE	80.83	81.09	81.09	82.51	81.29	81.09
Hold-Out	75.81	75.84	75.84	75.98	75.86	75.84
PSE	71.14	70.93	70.93	69.76	70.77	70.93
Non-PSE	80.07	80.31	80.31	81.64	80.50	80.31
YRMSE	32.06	41.66	41.66	36.38	39.15	41.63
$YRMSE_0$	32.42	42.34	42.34	36.63	39.74	42.31
PRMSE	15.79	15.72	15.72	16.52	15.81	15.72
$PRMSE_0$	15.80	15.73	15.73	16.57	15.82	15.73
$MSE(\hat{\beta})$	.42303	8.01468	8.01468	3.52074	5.84497	7.98752
$RMSE(\hat{\beta})$	.65040	2.83102	2.83102	1.87636	2.41764	2.82622
$VAR(\hat{\beta})$	.28818	.87270	.87270	.02399	.71232	.87066
$Bias^2(\hat{\beta})$	.13484	7.14198	7.14199	3.49676	5.13265	7.11686
MSE ($\partial p / \partial x$)	.03608	.03831	.03831	.63056	.03191	.03820
RMSE ($\partial p / \partial x$)	.18995	.19572	.19572	.79408	.17863	.19546

*N=1,000; balanced; restricted: errors drawn from chi-square distribution.

APPENDIX D

TABLES FOR LINEAR REGRESSION MONTE CARLO EXPERIMENTS

Table D.1 Monte Carlo Linear Regression 11 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	49.49	247.17	38.25	40.15	48.10
RMSE ($\hat{\beta}$)	7.04	15.72	6.18	6.34	6.94
Var ($\hat{\beta}$)	49.44	244.38	30.19	17.92	11.84
$Bias^2(\hat{\beta})$	0.05	2.79	8.07	22.23	36.26
PRMSE	11.58	44.90	11.65	11.68	11.75
PRMSE ($\hat{p}+\hat{e}$)		42.92	0.00	0.00	0.00
$PRMSE_0$	7.54	26.45	7.52	7.49	7.50
R^2	0.77				

* N=150: unrestricted; normal errors.

Table D.2 Monte Carlo Linear Regression 12 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	46.42	507.60	285.70	133.66	91.45
RMSE ($\hat{\beta}$)	6.81	22.53	16.90	11.56	9.56
Var ($\hat{\beta}$)	46.35	500.20	277.14	113.30	57.18
$Bias^2(\hat{\beta})$	0.07	7.40	8.57	20.36	34.26
PRMSE	11.10	86.69	53.01	30.34	16.64
PRMSE ($\hat{p}+\hat{e}$)		85.46	50.77	27.46	11.29
$PRMSE_0$	7.42	50.74	31.36	18.45	10.51
R^2	0.77				

* N=150; unrestricted; standardized *t*- errors.

Table D.3 Monte Carlo Linear Regression 13 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	48.82	273.75	90.61	44.77	48.99
RMSE ($\hat{\beta}$)	6.99	16.55	9.52	6.69	7.00
Var ($\hat{\beta}$)	48.70	263.34	79.36	22.24	12.48
$\text{Bias}^2(\hat{\beta})$	0.11	10.41	11.25	22.53	36.51
PRMSE	11.51	45.64	14.14	11.65	11.68
PRMSE ($\hat{p}+\hat{e}$)		44.85	7.04	0.01	0.00
PRMSE_0	7.40	26.91	8.89	7.44	7.39
R^2	0.77				

* N=150; unrestricted; standardized chi-square errors.

Table D.4 Monte Carlo Linear Regression 21 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	69.38	267.80	47.51	50.10	58.21
RMSE ($\hat{\beta}$)	8.33	16.36	6.89	7.08	7.63
Var ($\hat{\beta}$)	69.29	265.99	35.57	20.46	12.82
$\text{Bias}^2(\hat{\beta})$	0.08	1.82	11.94	29.64	45.39
PRMSE	9.12	33.17	9.20	9.26	9.35
PRMSE ($\hat{p}+\hat{e}$)		31.24	0.00	0.00	0.00
PRMSE_0	7.81	24.25	7.73	7.68	7.65
R^2	0.70				

* N=100; unrestricted; normal errors.

Table D.5 Monte Carlo Linear Regression 22 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE $(\hat{\beta})$	60.96	532.49	215.62	103.09	85.30
RMSE $(\hat{\beta})$	7.81	23.08	14.68	10.15	9.24
Var $(\hat{\beta})$	60.84	528.64	196.54	70.54	38.99
$Bias^2(\hat{\beta})$	0.12	3.85	19.08	32.55	46.31
PRMSE	8.75	69.45	34.04	16.68	15.37
PRMSE $(\hat{p}+\hat{e})$		68.45	32.39	13.75	12.64
$PRMSE_0$	7.65	52.28	24.96	12.91	11.84
R^2	0.70				

* N=100; unrestricted; standardized *t*- errors.

Table D.6 Monte Carlo Linear Regression 23 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE $(\hat{\beta})$	67.69	306.45	82.25	53.99	58.61
RMSE $(\hat{\beta})$	8.23	17.51	9.07	7.35	7.66
Var $(\hat{\beta})$	67.59	298.22	69.48	24.56	13.63
$Bias^2(\hat{\beta})$	0.10	8.23	12.77	29.43	44.98
PRMSE	9.13	50.74	10.82	9.33	9.37
PRMSE $(\hat{p}+\hat{e})$		50.18	5.45	0.00	0.00
$PRMSE_0$	7.81	36.86	8.91	7.77	7.67
R^2	0.70				

* N=100; unrestricted; standardized chi-square errors.

Table D.7 Monte Carlo Linear Regression 31 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE $(\hat{\beta})$	211.38	284.54	79.54	76.97	81.41
RMSE $(\hat{\beta})$	14.54	16.87	8.92	8.77	9.02
Var $(\hat{\beta})$	211.31	279.43	47.32	22.37	12.56
$\text{Bias}^2(\hat{\beta})$	0.07	5.11	32.21	54.60	68.85
PRMSE	5.73	18.33	5.91	6.06	6.19
PRMSE $(\hat{p}+\hat{e})$		17.17	0.00	0.00	0.00
PRMSE_0	9.10	19.33	8.45	8.25	8.15
R^2	0.84				

* N=50; unrestricted; normal errors.

Table D.8 Monte Carlo Linear Regression 32 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE $(\hat{\beta})$	204.29	433.90	178.84	113.22	99.39
RMSE $(\hat{\beta})$	14.29	20.83	13.37	10.64	9.97
Var $(\hat{\beta})$	204.06	421.99	143.43	55.38	29.87
$\text{Bias}^2(\hat{\beta})$	0.23	11.91	35.41	57.84	69.52
PRMSE	5.69	37.90	19.21	8.58	8.13
PRMSE $(\hat{p}+\hat{e})$		37.33	18.26	5.39	4.61
PRMSE_0	9.01	38.50	20.32	10.88	10.37
R^2	0.84				

* N=50; unrestricted; standardized t- errors.

Table D.9 Monte Carlo Linear Regression 33 – PS1 (See Table 3.1 for design)

	OLS	GME1-PS1	GME2-PS1	GME3-PS1	GME4-PS1
MSE ($\hat{\beta}$)	202.24	358.57	99.90	80.70	83.08
RMSE ($\hat{\beta}$)	14.22	18.94	10.00	8.98	9.11
Var ($\hat{\beta}$)	202.05	348.99	64.97	24.91	13.46
$\text{Bias}^2(\hat{\beta})$	0.18	9.58	34.93	55.79	69.62
PRMSE	5.74	32.80	6.22	6.07	6.18
PRMSE ($\hat{p}+\hat{e}$)		32.22	1.86	0.00	0.00
PRMSE_0	9.08	32.91	8.83	8.31	8.16
R^2	0.84				

* N=50; unrestricted; standardized chi-square errors.

Table D.10 Monte Carlo Linear Regression 41 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	32.29	183.23	84.49	150.89	205.01
RMSE ($\hat{\beta}$)	5.68	13.54	9.19	12.28	14.32
Var ($\hat{\beta}$)	30.22	157.61	11.37	5.12	3.14
$\text{Bias}^2(\hat{\beta})$	2.07	25.62	73.12	145.77	201.87
PRMSE	11.64	72.33	12.20	12.94	13.85
PRMSE ($\hat{p}+\hat{e}$)		71.28	0.00	0.00	0.00
PRMSE_0	7.47	42.40	7.72	8.06	8.48

* N=150; restricted; normal errors.

Table D.11 Monte Carlo Linear Regression 42 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	31.30	314.07	200.90	194.73	222.66
RMSE ($\hat{\beta}$)	5.59	17.72	14.17	13.95	14.92
Var ($\hat{\beta}$)	29.12	257.62	119.53	50.37	24.38
Bias^2($\hat{\beta}$)	2.18	56.45	81.37	144.36	198.27
PRMSE	11.16	82.45	36.56	23.86	16.71
PRMSE ($\hat{p}+\hat{e}$)		81.22	33.38	19.44	8.18
PRMSE$_0$	7.37	48.51	21.81	14.56	10.37

* N=150; restricted: standardized *t*- errors.

Table D.12 Monte Carlo Linear Regression 43 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	32.54	189.50	113.85	153.58	206.14
RMSE ($\hat{\beta}$)	5.70	13.77	10.67	12.39	14.36
Var ($\hat{\beta}$)	29.64	153.92	36.20	7.33	3.44
Bias^2($\hat{\beta}$)	2.90	35.57	77.65	146.25	202.69
PRMSE	11.56	63.21	16.43	12.91	13.77
PRMSE ($\hat{p}+\hat{e}$)		59.00	10.25	0.02	0.00
PRMSE$_0$	7.34	37.15	10.11	8.03	8.38

* N=150; restricted: standardized chi-square errors.

Table D.13 Monte Carlo Linear Regression 51 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	45.40	181.60	112.57	184.47	239.57
RMSE ($\hat{\beta}$)	6.74	13.48	10.61	13.58	15.48
Var ($\hat{\beta}$)	41.80	154.48	12.48	5.76	3.52
Bias^2 ($\hat{\beta}$)	3.60	27.12	100.09	178.72	236.05
PRMSE	9.20	49.06	9.90	10.74	11.71
PRMSE ($\hat{p}+\hat{e}$)		48.02	0.00	0.00	0.00
PRMSE$_0$	7.68	34.76	8.05	8.64	9.38

* N=100; restricted; normal errors.

Table D.14 Monte Carlo Linear Regression 52 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	41.37	280.51	191.72	208.22	251.00
RMSE ($\hat{\beta}$)	6.43	16.75	13.85	14.43	15.84
Var ($\hat{\beta}$)	37.73	224.71	81.61	28.58	14.67
Bias^2 ($\hat{\beta}$)	3.64	55.81	110.11	179.64	236.32
PRMSE	8.82	57.17	26.02	13.79	13.52
PRMSE ($\hat{p}+\hat{e}$)		55.85	23.34	8.08	6.42
PRMSE$_0$	7.55	40.23	18.97	10.98	10.86

* N=100; restricted; standardized *t*- errors.

Table D.15 Monte Carlo Linear Regression 53 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	43.99	186.53	126.84	185.90	239.78
RMSE ($\hat{\beta}$)	6.63	13.66	11.26	13.63	15.48
Var ($\hat{\beta}$)	39.99	156.53	27.72	7.47	3.87
Bias^2($\hat{\beta}$)	4.00	30.00	99.12	178.43	235.91
PRMSE	9.21	46.39	11.55	10.80	11.71
PRMSE ($\hat{p}+\hat{e}$)		43.22	4.75	0.00	0.00
PRMSE$_0$	7.68	32.55	9.17	8.74	9.43

* N=100; restricted; standardized chi-square errors.

Table D.16 Monte Carlo Linear Regression 61 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	112.27	211.19	265.81	348.04	398.91
RMSE ($\hat{\beta}$)	10.60	14.53	16.30	18.66	19.97
Var ($\hat{\beta}$)	97.44	144.41	14.62	6.82	4.06
Bias^2($\hat{\beta}$)	14.83	66.78	251.19	341.22	394.85
PRMSE	5.90	29.12	6.94	7.85	8.82
PRMSE ($\hat{p}+\hat{e}$)		28.38	0.00	0.00	0.00
PRMSE$_0$	8.63	29.80	10.35	11.92	13.44

* N=50; restricted; normal errors.

Table D.17 Monte Carlo Linear Regression 62 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	104.99	292.33	307.85	360.64	405.34
RMSE ($\hat{\beta}$)	10.25	17.10	17.55	18.99	20.13
Var ($\hat{\beta}$)	91.37	194.79	60.56	22.67	11.73
$Bias^2(\hat{\beta})$	13.62	97.54	247.29	337.97	393.62
PRMSE	5.86	29.79	13.67	9.18	9.89
PRMSE ($\hat{p}+\hat{e}$)		28.91	11.52	4.29	4.40
$PRMSE_0$	8.53	31.06	16.36	13.31	14.47

* N=50; restricted; standardized *t*- errors.

Table D.18 Monte Carlo Linear Regression 63 – PSR1 (See Table 3.1 for design)

	IRLS	GME1-PSR1	GME2-PSR1	GME3-PSR1	GME4-PSR1
MSE ($\hat{\beta}$)	105.92	237.25	275.91	349.57	399.71
RMSE ($\hat{\beta}$)	10.29	15.40	16.61	18.70	19.99
Var ($\hat{\beta}$)	90.87	160.48	25.26	7.96	4.38
$Bias^2(\hat{\beta})$	15.05	76.77	250.65	341.61	395.33
PRMSE	5.92	28.19	8.39	7.87	8.82
PRMSE ($\hat{p}+\hat{e}$)		26.83	4.39	0.00	0.00
$PRMSE_0$	8.60	29.45	11.50	11.98	13.47

* N=50; restricted; standardized chi-square errors.

Table D.19 Monte Carlo Linear Regression 11 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	49.49	64.44	51.47	48.76	48.05
RMSE ($\hat{\beta}$)	7.04	8.03	7.17	6.98	6.93
Var ($\hat{\beta}$)	49.44	64.32	51.41	48.68	47.93
$\text{Bias}^2(\hat{\beta})$	0.05	0.12	0.06	0.08	0.11
PRMSE	11.58	12.07	11.61	11.58	11.58
PRMSE ($\hat{p}+\hat{e}$)		3.63	0.10	0.00	0.00
PRMSE_0	7.54	7.80	7.55	7.54	7.53
R^2	0.77				

* N=150; unrestricted; normal errors.

Table D.20 Monte Carlo Linear Regression 12 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	46.42	75.99	352.57	259.73	246.67
RMSE ($\hat{\beta}$)	6.81	8.72	18.78	16.12	15.71
Var ($\hat{\beta}$)	46.35	75.92	351.74	259.44	246.44
$\text{Bias}^2(\hat{\beta})$	0.07	0.08	0.82	0.29	0.23
PRMSE	11.10	12.11	17.15	13.65	13.66
PRMSE ($\hat{p}+\hat{e}$)		6.71	10.92	5.05	4.80
PRMSE_0	7.42	8.02	10.86	9.02	8.86
R^2	0.77				

* N=150; unrestricted; standardized t- errors.

Table D.21 Monte Carlo Linear Regression 13 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	48.82	56.90	106.61	57.61	50.01
RMSE ($\hat{\beta}$)	6.99	7.54	10.33	7.59	7.07
Var ($\hat{\beta}$)	48.70	56.71	105.66	57.36	49.75
$Bias^2$ ($\hat{\beta}$)	0.11	0.20	0.95	0.26	0.26
PRMSE	11.51	12.34	11.99	11.56	11.51
PRMSE ($\hat{p}+\hat{e}$)		5.97	0.67	0.04	0.00
$PRMSE_0$	7.40	7.86	7.80	7.48	7.42
R^2	0.77				

* N=150; unrestricted: standardized chi-square errors.

Table D.22 Monte Carlo Linear Regression 21 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	69.38	82.95	69.96	67.95	66.74
RMSE ($\hat{\beta}$)	8.33	9.11	8.36	8.24	8.17
Var ($\hat{\beta}$)	69.29	82.86	69.88	67.86	66.62
$Bias^2$ ($\hat{\beta}$)	0.08	0.08	0.08	0.09	0.12
PRMSE	9.12	9.38	9.14	9.12	9.12
PRMSE ($\hat{p}+\hat{e}$)		2.01	0.03	0.00	0.00
$PRMSE_0$	7.81	8.00	7.82	7.80	7.80
R^2	0.70				

* N=100; unrestricted; normal errors.

Table D.23 Monte Carlo Linear Regression 22 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	60.96	106.91	233.84	194.49	146.41
RMSE ($\hat{\beta}$)	7.81	10.34	15.29	13.95	12.10
Var ($\hat{\beta}$)	60.84	106.33	233.41	194.19	146.07
$\text{Bias}^2(\hat{\beta})$	0.12	0.58	0.43	0.30	0.34
PRMSE	8.75	9.76	10.19	9.87	9.46
PRMSE ($\hat{p}+\hat{e}$)		5.24	2.81	1.90	1.24
PRMSE_0	7.65	8.44	9.19	8.83	8.43
R^2	0.70				

* N=100; unrestricted; standardized *t*- errors.

Table D.24 Monte Carlo Linear Regression 23 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	67.69	99.31	105.41	76.30	68.09
RMSE ($\hat{\beta}$)	8.23	9.97	10.27	8.73	8.25
Var ($\hat{\beta}$)	67.59	98.94	105.26	76.20	67.99
$\text{Bias}^2(\hat{\beta})$	0.10	0.36	0.15	0.10	0.10
PRMSE	9.13	9.54	9.45	9.19	9.14
PRMSE ($\hat{p}+\hat{e}$)		3.23	0.59	0.00	0.00
PRMSE_0	7.81	8.19	8.17	7.90	7.83
R^2	0.70				

* N=100; unrestricted; standardized chi-square errors.

Table D.25 Monte Carlo Linear Regression 31 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	211.38	237.57	205.26	195.61	184.61
RMSE ($\hat{\beta}$)	14.54	15.41	14.33	13.99	13.59
Var ($\hat{\beta}$)	211.31	237.39	205.21	195.47	184.22
$Bias^2(\hat{\beta})$	0.07	0.17	0.06	0.14	0.39
PRMSE	5.73	5.90	5.74	5.74	5.74
PRMSE ($\hat{p}+\hat{e}$)		0.77	0.00	0.00	0.00
$PRMSE_0$	9.10	9.32	9.09	9.06	9.03
R^2	0.84				

* N=50; unrestricted; normal errors.

Table D.26 Monte Carlo Linear Regression 32 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	204.29	362.83	541.50	437.21	307.47
RMSE ($\hat{\beta}$)	14.29	19.05	23.27	20.91	17.53
Var ($\hat{\beta}$)	204.06	362.11	540.09	435.77	305.57
$Bias^2(\hat{\beta})$	0.23	0.72	1.41	1.44	1.90
PRMSE	5.69	6.36	7.88	7.31	6.02
PRMSE ($\hat{p}+\hat{e}$)		2.91	4.75	3.80	1.60
$PRMSE_0$	9.01	10.23	11.68	10.96	9.70
R^2	0.84				

* N=50; unrestricted; standardized *t*- errors.

Table D.27 Monte Carlo Linear Regression 33 – PS2 (See Table 3.1 for design)

	OLS	GME1-PS2	GME2-PS2	GME3-PS2	GME4-PS2
MSE ($\hat{\beta}$)	202.24	294.68	232.98	198.57	182.25
RMSE ($\hat{\beta}$)	14.22	17.17	15.26	14.09	13.50
Var ($\hat{\beta}$)	202.05	293.96	232.46	198.09	181.48
Bias^2($\hat{\beta}$)	0.18	0.72	0.52	0.48	0.77
PRMSE	5.74	6.08	5.83	5.76	5.75
PRMSE ($\hat{p}+\hat{e}$)		1.47	0.16	0.00	0.00
PRMSE$_0$	9.08	9.71	9.32	9.12	9.04
R^2	0.84				

* N=50; unrestricted; standardized chi-square errors.

Table D.28 Monte Carlo Linear Regression 41 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	32.29	520.38	55.13	124.30	268.01
RMSE ($\hat{\beta}$)	5.68	22.81	7.42	11.15	16.37
Var ($\hat{\beta}$)	30.22	153.00	29.81	24.47	21.47
Bias^2($\hat{\beta}$)	2.07	367.39	25.32	99.83	246.55
PRMSE	11.64	157.57	11.81	12.19	13.01
PRMSE ($\hat{p}+\hat{e}$)		157.18	0.74	0.00	0.01
PRMSE$_0$	7.47	95.99	7.59	7.82	8.32

* N=150; restricted; normal errors.

Table D.29 Monte Carlo Linear Regression 42 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	31.30	1697.01	357.11	307.48	378.60
RMSE ($\hat{\beta}$)	5.59	41.19	18.90	17.54	19.46
Var ($\hat{\beta}$)	29.12	686.75	274.69	186.80	120.83
Bias^2($\hat{\beta}$)	2.18	1010.26	82.42	120.68	257.77
PRMSE	11.16	293.28	103.15	75.83	53.30
PRMSE ($\hat{p}+\hat{e}$)		293.12	102.65	75.27	51.70
PRMSE$_0$	7.37	178.30	61.68	44.96	31.63

* N=150; restricted; standardized *t*- errors.

Table D.30 Monte Carlo Linear Regression 43 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	32.54	901.76	103.96	132.67	272.72
RMSE ($\hat{\beta}$)	5.70	30.03	10.20	11.52	16.51
Var ($\hat{\beta}$)	29.64	262.08	67.25	30.03	22.44
Bias^2($\hat{\beta}$)	2.90	639.68	36.72	102.64	250.28
PRMSE	11.56	210.04	30.03	12.17	12.94
PRMSE ($\hat{p}+\hat{e}$)		208.10	26.60	0.14	0.01
PRMSE$_0$	7.34	127.97	18.24	7.78	8.23

* N=150; restricted; standardized chi-square errors.

Table D.31 Monte Carlo Linear Regression 51 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	45.40	618.19	88.70	218.19	470.65
RMSE ($\hat{\beta}$)	6.74	24.86	9.42	14.77	21.69
Var ($\hat{\beta}$)	41.80	133.83	39.92	33.89	29.72
Bias^2($\hat{\beta}$)	3.60	484.37	48.78	184.30	440.93
PRMSE	9.20	149.19	9.43	10.03	11.17
PRMSE ($\hat{p}+\hat{e}$)		148.90	0.45	0.00	0.01
PRMSE$_0$	7.68	94.40	7.83	8.16	8.81

* N=100; restricted; normal errors.

Table D.32 Monte Carlo Linear Regression 52 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	41.37	1037.18	285.58	301.95	588.68
RMSE ($\hat{\beta}$)	6.43	32.21	16.90	17.38	24.26
Var ($\hat{\beta}$)	37.73	457.97	200.97	103.95	133.08
Bias^2($\hat{\beta}$)	3.64	579.20	84.61	198.00	455.60
PRMSE	8.82	193.83	68.10	35.62	74.95
PRMSE ($\hat{p}+\hat{e}$)		193.69	67.34	33.77	73.64
PRMSE$_0$	7.75	123.06	46.30	24.97	53.55

* N=100; restricted; standardized *t*- errors.

Table D.33 Monte Carlo Linear Regression 53 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	43.99	758.88	117.61	225.16	476.04
RMSE ($\hat{\beta}$)	6.63	27.55	10.84	15.01	21.82
Var ($\hat{\beta}$)	39.99	185.91	63.10	38.96	30.99
Bias^2($\hat{\beta}$)	4.00	572.97	54.51	186.20	445.05
PRMSE	9.21	166.31	20.97	10.10	11.18
PRMSE ($\hat{p}+\hat{e}$)		164.79	18.10	0.03	0.01
PRMSE$_0$	7.68	105.45	14.69	8.26	8.86

* N=100; restricted; standardized chi-square errors.

Table D.34 Monte Carlo Linear Regression 61 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	112.27	297.51	436.69	1224.26	2560.19
RMSE ($\hat{\beta}$)	10.60	17.25	20.90	34.99	50.60
Var ($\hat{\beta}$)	97.44	179.95	91.17	77.05	67.07
Bias^2($\hat{\beta}$)	14.83	117.56	345.52	1147.22	2493.12
PRMSE	5.90	54.00	6.59	8.15	10.30
PRMSE ($\hat{p}+\hat{e}$)		53.64	0.00	0.00	0.00
PRMSE$_0$	8.63	57.29	10.39	13.78	18.28

* N=50; restricted; normal errors.

Table D.35 Monte Carlo Linear Regression 62 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	104.99	574.90	782.05	1527.48	2651.71
RMSE ($\hat{\beta}$)	10.25	23.98	27.97	39.08	51.49
Var ($\hat{\beta}$)	91.37	363.50	394.11	340.50	162.64
$\text{Bias}^2(\hat{\beta})$	13.62	211.40	387.94	1186.98	2489.08
PRMSE	5.86	88.38	52.80	32.66	16.78
PRMSE ($\hat{p}+\hat{e}$)		88.19	52.46	31.50	13.40
PRMSE_0	8.53	93.06	53.54	33.66	22.91

* N=50; restricted; standardized *t*- errors.

Table D.36 Monte Carlo Linear Regression 63 – PSR2 (See Table 3.1 for design)

	IRLS	GME1-PSR2	GME2-PSR2	GME3-PSR2	GME4-PSR2
MSE ($\hat{\beta}$)	105.92	433.93	456.37	1219.11	2548.88
RMSE ($\hat{\beta}$)	10.29	20.83	21.36	34.92	50.49
Var ($\hat{\beta}$)	90.87	241.82	112.31	81.62	68.42
$\text{Bias}^2(\hat{\beta})$	15.05	192.11	344.06	1137.49	2480.46
PRMSE	5.92	73.21	11.81	8.17	10.32
PRMSE ($\hat{p}+\hat{e}$)		72.50	9.55	0.00	0.00
PRMSE_0	8.60	77.53	14.58	13.86	18.33

* N=50; restricted; standardized chi-square errors.

Table D.37 Monte Carlo Linear Regression 11 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	49.49	241.13	30.99	18.30	12.03
RMSE ($\hat{\beta}$)	7.04	15.53	5.57	4.28	3.47
Var ($\hat{\beta}$)	49.44	240.32	30.95	18.27	12.01
$Bias^2(\hat{\beta})$	0.05	0.81	0.04	0.03	0.02
PRMSE	11.58	42.96	11.63	11.61	11.64
PRMSE ($\hat{p}+\hat{e}$)		41.05	0.00	0.00	0.00
$PRMSE_0$	7.54	25.28	7.51	7.45	7.42
R^2	0.77				

* N=150; unrestricted; normal errors.

Table D.38 Monte Carlo Linear Regression 12 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	46.42	510.84	282.68	117.81	57.89
RMSE ($\hat{\beta}$)	6.81	22.60	16.81	10.85	7.61
Var ($\hat{\beta}$)	46.35	509.93	281.92	117.53	57.73
$Bias^2(\hat{\beta})$	0.07	0.90	0.76	0.28	0.16
PRMSE	11.10	64.76	51.54	24.48	15.41
PRMSE ($\hat{p}+\hat{e}$)		62.88	49.19	21.05	9.39
$PRMSE_0$	7.42	37.93	30.44	15.14	9.82
R^2	0.77				

* N=150; unrestricted; standardized *t*- errors.

Table D.39 Monte Carlo Linear Regression 13 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE $(\hat{\beta})$	48.82	270.24	82.95	22.66	12.63
RMSE $(\hat{\beta})$	6.99	16.44	9.11	4.76	3.55
Var $(\hat{\beta})$	48.70	265.62	82.19	22.62	12.61
$\text{Bias}^2(\hat{\beta})$	0.11	4.61	0.76	0.04	0.01
PRMSE	11.51	33.86	16.53	11.59	11.57
PRMSE $(\hat{p}+\hat{e})$		30.43	11.05	0.01	0.00
PRMSE_0	7.40	20.01	10.18	7.40	7.31
R^2	0.77				

* N=150; unrestricted; standardized chi-square errors.

Table D.40 Monte Carlo Linear Regression 21 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE $(\hat{\beta})$	69.38	256.48	36.53	20.83	12.99
RMSE $(\hat{\beta})$	8.33	16.01	6.04	4.56	3.60
Var $(\hat{\beta})$	69.29	256.06	36.47	20.78	12.96
$\text{Bias}^2(\hat{\beta})$	0.08	0.42	0.06	0.05	0.03
PRMSE	9.12	23.85	9.16	9.17	9.21
PRMSE $(\hat{p}+\hat{e})$		21.32	0.00	0.00	0.00
PRMSE_0	7.81	17.77	7.70	7.61	7.55
R^2	0.70				

* N=100; unrestricted; normal errors.

Table D.41 Monte Carlo Linear Regression 22 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	60.96	432.50	199.08	71.60	38.79
RMSE ($\hat{\beta}$)	7.81	20.80	14.11	8.46	6.23
Var ($\hat{\beta}$)	60.84	429.32	198.34	71.49	38.73
$\text{Bias}^2(\hat{\beta})$	0.12	3.17	0.74	0.12	0.05
PRMSE	8.75	42.36	32.32	20.61	14.91
PRMSE ($\hat{p}+\hat{e}$)		40.65	30.36	18.23	11.92
PRMSE_0	7.65	30.98	23.87	15.59	11.59
R^2	0.70				

* N=100; unrestricted; standardized *t*- errors.

Table D.42 Monte Carlo Linear Regression 23 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	67.69	294.54	70.79	25.03	13.82
RMSE ($\hat{\beta}$)	8.23	17.16	8.41	5.00	3.72
Var ($\hat{\beta}$)	67.59	292.25	70.60	24.98	13.78
$\text{Bias}^2(\hat{\beta})$	0.10	2.29	0.18	0.05	0.03
PRMSE	9.13	38.09	10.92	9.24	9.23
PRMSE ($\hat{p}+\hat{e}$)		36.31	5.61	0.00	0.00
PRMSE_0	7.81	27.76	8.95	7.70	7.56
R^2	0.70				

* N=100; unrestricted; standardized chi-square errors.

Table D.43 Monte Carlo Linear Regression 31 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	211.38	275.55	47.99	22.58	12.66
RMSE ($\hat{\beta}$)	14.54	16.60	6.93	4.75	3.56
Var ($\hat{\beta}$)	211.31	275.39	47.96	22.56	12.66
$\text{Bias}^2(\hat{\beta})$	0.07	0.16	0.03	0.01	0.01
PRMSE	5.73	19.35	5.85	5.95	6.03
PRMSE ($\hat{p}+\hat{e}$)		18.30	0.00	0.00	0.00
PRMSE_0	9.10	20.11	8.37	8.07	7.88
R^2	0.84				

* N=50; unrestricted; normal errors.

Table D.44 Monte Carlo Linear Regression 32 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	204.29	424.22	145.21	56.31	29.50
RMSE ($\hat{\beta}$)	14.29	20.60	12.05	7.50	5.43
Var ($\hat{\beta}$)	204.06	423.43	144.98	56.23	29.46
$\text{Bias}^2(\hat{\beta})$	0.23	0.79	0.23	0.08	0.04
PRMSE	5.69	36.80	15.06	8.44	6.65
PRMSE ($\hat{p}+\hat{e}$)		36.17	13.87	5.49	2.08
PRMSE_0	9.01	37.33	16.54	10.57	8.91
R^2	0.84				

* N=50; unrestricted; standardized *t*- errors.

Table D.45 Monte Carlo Linear Regression 33 – PS3 (See Table 3.1 for design)

	OLS	GME1-PS3	GME2-PS3	GME3-PS3	GME4-PS3
MSE ($\hat{\beta}$)	202.24	355.75	65.86	25.12	13.56
RMSE ($\hat{\beta}$)	14.22	18.86	8.12	5.01	3.68
Var ($\hat{\beta}$)	202.05	354.39	65.71	25.09	13.55
$\text{Bias}^2(\hat{\beta})$	0.18	1.36	0.16	0.03	0.01
PRMSE	5.74	35.00	6.19	5.96	6.03
PRMSE ($\hat{p}+\hat{e}$)		34.33	1.88	0.00	0.00
PRMSE_0	9.08	34.99	8.82	8.14	7.89
R^2	0.84				

* N=50; unrestricted; standardized chi-square errors.

APPENDIX E

TABLES FOR MONTE CARLO EXPERIMENTS EXAMINING PARAMETER INEQUALITY RESTRICTIONS

Table E.1 Monte Carlo IR 11 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	49.40	36.39	29.85	40.24
RMSE ($\hat{\beta}$)	7.04	7.03	6.03	5.46	6.34
Var ($\hat{\beta}$)	49.44	49.35	30.63	24.67	27.43
Bias^2($\hat{\beta}$)	0.05	0.05	5.76	5.18	12.81
PRMSE	11.58	11.58	11.64	11.66	11.68
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.52	7.50	7.53
R^2	0.77				

* N=150; $\beta_2 \geq -10$; $\delta_t = -5.5$; normal errors.

Table E.2 Monte Carlo IR 12 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	45.25	36.77	48.60	52.74
RMSE ($\hat{\beta}$)	7.04	6.73	6.06	6.97	7.26
Var ($\hat{\beta}$)	49.44	44.33	29.93	25.22	26.07
Bias^2($\hat{\beta}$)	0.05	0.92	6.84	23.39	26.67
PRMSE	11.58	11.59	11.64	11.74	11.75
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.55	7.56
R^2	0.77				

* N=150; $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table E.3 Monte Carlo IR 13 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	45.04	37.25	51.39	55.48
RMSE ($\hat{\beta}$)	7.04	6.71	6.10	7.17	7.45
Var ($\hat{\beta}$)	49.44	43.50	30.01	25.57	26.15
Bias^2($\hat{\beta}$)	0.05	1.53	7.23	25.82	29.33
PRMSE	11.58	11.60	11.64	11.75	11.76
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.56	7.57
R^2	0.77				

* N=150: $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table E.4 Monte Carlo IR 14 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	47.33	37.73	63.13	69.06
RMSE ($\hat{\beta}$)	7.04	6.88	6.14	7.95	8.31
Var ($\hat{\beta}$)	49.44	41.49	30.03	25.41	25.88
Bias^2($\hat{\beta}$)	0.05	5.84	7.70	37.72	43.19
PRMSE	11.58	11.63	11.65	11.81	11.83
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.52	7.60	7.61
R^2	0.77				

* N=150; $\beta_2 \geq -3$; $\delta_t = 1.5$; normal errors.

Table E.5 Monte Carlo IR 15 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	56.29	38.25	77.42	79.35
RMSE ($\hat{\beta}$)	7.04	7.50	6.18	8.80	8.91
Var ($\hat{\beta}$)	49.44	40.52	30.19	25.43	26.01
Bias^2($\hat{\beta}$)	0.05	15.77	8.07	51.98	53.34
PRMSE	11.58	11.69	11.65	11.88	11.88
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.55	7.52	7.64	7.64
R^2	0.77				

* N=150; $\beta_2 \geq -1.5$; $\delta_t = 3$; normal errors.

Table E.6 Monte Carlo IR 16 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	73.15	39.43	108.35	106.72
RMSE ($\hat{\beta}$)	7.04	8.55	6.28	10.41	10.33
Var ($\hat{\beta}$)	49.44	40.17	30.26	25.89	26.49
Bias^2($\hat{\beta}$)	0.05	32.98	9.18	82.47	80.23
PRMSE	11.58	11.77	11.65	12.03	12.01
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.60	7.52	7.73	7.72
R^2	0.77				

* N=150; $\beta_2 \geq 0$; $\delta_t = 4.5$; normal errors.

Table E.7 Monte Carlo IR 17 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	130.35	41.30	167.81	152.86
RMSE ($\hat{\beta}$)	7.04	11.42	6.43	12.95	12.36
Var ($\hat{\beta}$)	49.44	40.03	30.84	26.73	26.69
Bias^2($\hat{\beta}$)	0.05	90.33	10.46	141.09	126.17
PRMSE	11.58	12.04	11.66	12.31	12.24
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.75	7.53	7.90	7.85
R^2	0.77				

* N=150; $\beta_2 \geq 3$; $\delta_t = 7.5$; normal errors.

Table E.8 Monte Carlo IR 11 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	49.40	32.67	27.02	45.76
RMSE ($\hat{\beta}$)	7.04	7.03	5.72	5.20	6.76
Var ($\hat{\beta}$)	49.44	49.35	18.47	14.49	15.73
Bias^2($\hat{\beta}$)	0.05	0.05	14.21	12.53	30.04
PRMSE	11.58	11.58	11.64	11.67	11.72
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.47	7.46	7.51
R^2	0.77				

* N=150; $\beta_2 \geq -10$; $\delta_t = -5.5$; normal errors.

Table E.9 Monte Carlo IR 12 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	45.25	35.37	53.77	65.98
RMSE ($\hat{\beta}$)	7.04	6.73	5.95	7.33	8.12
Var ($\hat{\beta}$)	49.44	44.33	17.69	14.77	15.24
Bias^2($\hat{\beta}$)	0.05	0.92	17.67	39.00	50.74
PRMSE	11.58	11.59	11.66	11.78	11.82
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.48	7.53	7.57
R^2	0.77				

* N=150; $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table E.10 Monte Carlo IR 13 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	45.04	36.88	60.24	71.79
RMSE ($\hat{\beta}$)	7.04	6.71	6.07	7.76	8.47
Var ($\hat{\beta}$)	49.44	43.50	17.76	15.00	15.39
Bias^2($\hat{\beta}$)	0.05	1.53	19.12	45.24	56.40
PRMSE	11.58	11.60	11.67	11.80	11.85
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.48	7.55	7.59
R^2	0.77				

* N=150; $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table E.11 Monte Carlo IR 14 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	47.33	38.51	74.93	90.05
RMSE ($\hat{\beta}$)	7.04	6.88	6.21	8.66	9.49
Var ($\hat{\beta}$)	49.44	41.49	17.76	14.92	15.19
Bias^2($\hat{\beta}$)	0.05	5.84	20.75	60.01	74.86
PRMSE	11.58	11.63	11.67	11.87	11.94
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.49	7.59	7.64
R^2	0.77				

* N=150; $\beta_2 \geq -3 : \delta_t = 1.5$; normal errors.

Table E.12 Monte Carlo IR 15 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	56.29	40.15	92.25	109.72
RMSE ($\hat{\beta}$)	7.04	7.50	6.34	9.60	10.47
Var ($\hat{\beta}$)	49.44	40.52	17.92	14.90	15.42
Bias^2($\hat{\beta}$)	0.05	15.77	22.23	77.35	94.30
PRMSE	11.58	11.69	11.68	11.95	12.02
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.55	7.49	7.65	7.70
R^2	0.77				

* N=150; $\beta_2 \geq -1.5 ; \delta_t = 3$; normal errors.

Table E.13 Monte Carlo IR 16 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	73.15	44.25	134.21	160.65
RMSE ($\hat{\beta}$)	7.04	8.55	6.65	11.58	12.67
Var ($\hat{\beta}$)	49.44	40.17	17.95	14.93	15.53
Bias^2($\hat{\beta}$)	0.05	32.98	26.30	119.28	145.12
PRMSE	11.58	11.77	11.70	12.15	12.26
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.60	7.51	7.77	7.84
R^2	0.77				

* N=150; $\beta_2 \geq 0$; $\delta_t = 4.5$; normal errors.

Table E.14 Monte Carlo IR 17 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	130.35	51.14	216.98	223.78
RMSE ($\hat{\beta}$)	7.04	11.42	7.15	14.73	14.96
Var ($\hat{\beta}$)	49.44	40.03	18.66	15.05	15.53
Bias^2($\hat{\beta}$)	0.05	90.33	32.49	201.92	208.24
PRMSE	11.58	12.04	11.72	12.52	12.54
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.75	7.53	8.00	8.02
R^2	0.77				

* N=150; $\beta_2 \geq 3$; $\delta_t = 7.5$; normal errors.

Table E.15 Monte Carlo IR 21 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	68.38	43.68	34.46	49.94
RMSE ($\hat{\beta}$)	8.33	8.27	6.61	5.87	7.07
Var ($\hat{\beta}$)	69.29	68.28	36.52	28.29	31.23
Bias^2($\hat{\beta}$)	0.08	0.10	7.16	6.17	18.70
PRMSE	9.12	9.12	9.18	9.21	9.23
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.80	7.72	7.69	7.74
R^2	0.70				

* N=100; $\beta_2 \geq -10$; $\delta_t = -5.5$; normal errors.

Table E.16 Monte Carlo IR 22 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	61.68	44.44	57.66	66.49
RMSE ($\hat{\beta}$)	8.33	7.85	6.67	7.59	8.15
Var ($\hat{\beta}$)	69.29	59.83	35.13	28.91	29.96
Bias^2($\hat{\beta}$)	0.08	1.85	9.31	28.75	36.53
PRMSE	9.12	9.14	9.19	9.29	9.30
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.79	7.72	7.76	7.78
R^2	0.70				

* N=100; $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table E.17 Monte Carlo IR 23 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	61.47	45.42	62.70	70.90
RMSE ($\hat{\beta}$)	8.33	7.84	6.74	7.92	8.42
Var ($\hat{\beta}$)	69.29	58.78	35.27	29.40	30.20
Bias^2($\hat{\beta}$)	0.08	2.68	10.15	33.30	40.70
PRMSE	9.12	9.15	9.19	9.30	9.32
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.79	7.72	7.77	7.80
R^2	0.70				

* N=100; $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table E.18 Monte Carlo IR 24 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	63.39	46.40	75.91	86.85
RMSE ($\hat{\beta}$)	8.33	7.96	6.81	8.71	9.32
Var ($\hat{\beta}$)	69.29	56.20	35.29	29.24	29.83
Bias^2($\hat{\beta}$)	0.08	7.20	11.11	46.67	57.02
PRMSE	9.12	9.17	9.19	9.35	9.38
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.80	7.73	7.81	7.84
R^2	0.70				

* N=100; $\beta_2 \geq -3$; $\delta_t = 1.5$; normal errors.

Table E.19 Monte Carlo IR 25 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	71.19	47.51	91.64	101.69
RMSE ($\hat{\beta}$)	8.33	8.44	6.89	9.57	10.08
Var ($\hat{\beta}$)	69.29	54.62	35.57	29.24	30.22
Bias^2($\hat{\beta}$)	0.08	16.56	11.94	62.40	71.47
PRMSE	9.12	9.22	9.20	9.41	9.43
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.82	7.73	7.85	7.88
R^2	0.70				

* N=100; $\beta_2 \geq -1.5$; $\delta_t = 3$; normal errors.

Table E.20 Monte Carlo IR 26 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	86.62	49.98	128.56	140.67
RMSE ($\hat{\beta}$)	8.33	9.31	7.07	11.34	11.86
Var ($\hat{\beta}$)	69.29	53.89	35.66	29.52	30.69
Bias^2($\hat{\beta}$)	0.08	32.73	14.32	99.04	109.97
PRMSE	9.12	9.28	9.20	9.54	9.57
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.86	7.74	7.96	7.99
R^2	0.70				

* N=100; $\beta_2 \geq 0$; $\delta_t = 4.5$; normal errors.

Table E.21 Monte Carlo IR 27 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	140.78	54.39	199.05	192.95
RMSE ($\hat{\beta}$)	8.33	11.86	7.38	14.11	13.89
Var ($\hat{\beta}$)	69.29	53.55	36.83	30.14	30.79
Bias^2($\hat{\beta}$)	0.08	87.22	17.57	168.91	162.16
PRMSE	9.12	9.49	9.21	9.79	9.76
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	8.00	7.75	8.15	8.13
R^2	0.70				

* N=100: $\beta_2 \geq 3$; $\delta_t = 7.5$; normal errors.

Table E.22 Monte Carlo IR 21 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	68.38	37.73	30.60	54.32
RMSE ($\hat{\beta}$)	8.33	8.27	6.14	5.53	7.37
Var ($\hat{\beta}$)	69.29	68.28	21.35	16.55	17.69
Bias^2($\hat{\beta}$)	0.08	0.10	16.39	14.04	36.64
PRMSE	9.12	9.12	9.22	9.25	9.30
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.80	7.64	7.62	7.69
R^2	0.70				

* N=100: $\beta_2 \geq -10$; $\delta_t = -5.5$; normal errors.

Table E.23 Monte Carlo IR 22 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	61.68	41.98	59.20	77.21
RMSE ($\hat{\beta}$)	8.33	7.85	6.48	7.69	8.79
Var ($\hat{\beta}$)	69.29	59.83	20.14	16.87	17.31
Bias^2($\hat{\beta}$)	0.08	1.85	21.84	42.33	59.90
PRMSE	9.12	9.14	9.24	9.34	9.39
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.79	7.66	7.70	7.76
R^2	0.70				

* N=100; $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table E.24 Monte Carlo IR 23 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	61.47	44.50	68.36	85.80
RMSE ($\hat{\beta}$)	8.33	7.84	6.67	8.27	9.26
Var ($\hat{\beta}$)	69.29	58.78	20.24	17.08	17.50
Bias^2($\hat{\beta}$)	0.08	2.68	24.26	51.28	68.30
PRMSE	9.12	9.15	9.24	9.37	9.42
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.79	7.66	7.73	7.78
R^2	0.70				

* N=100; $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table E.25 Monte Carlo IR 24 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	63.39	47.18	83.62	105.15
RMSE ($\hat{\beta}$)	8.33	7.96	6.87	9.14	10.25
Var ($\hat{\beta}$)	69.29	56.20	20.23	17.03	17.32
Bias^2($\hat{\beta}$)	0.08	7.20	26.96	66.59	87.83
PRMSE	9.12	9.17	9.25	9.42	9.49
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.80	7.67	7.77	7.84
R^2	0.70				

* N=100; $\beta_2 \geq -3$; $\delta_t = 1.5$; normal errors.

Table E.26 Monte Carlo IR 25 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	71.19	50.10	101.63	133.42
RMSE ($\hat{\beta}$)	8.33	8.44	7.08	10.08	11.55
Var ($\hat{\beta}$)	69.29	54.62	20.46	17.02	17.63
Bias^2($\hat{\beta}$)	0.08	16.56	29.64	84.61	115.79
PRMSE	9.12	9.22	9.26	9.49	9.59
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.82	7.68	7.83	7.92
R^2	0.70				

* N=100; $\beta_2 \geq -1.5$; $\delta_t = 3$; normal errors.

Table E.27 Monte Carlo IR 26 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	86.62	57.06	145.99	198.76
RMSE ($\hat{\beta}$)	8.33	9.31	7.55	12.08	14.10
Var ($\hat{\beta}$)	69.29	53.89	20.47	17.05	17.71
Bias^2($\hat{\beta}$)	0.08	32.73	36.58	128.94	181.05
PRMSE	9.12	9.28	9.29	9.65	9.82
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.86	7.70	7.95	8.10
R^2	0.70				

* N=100; $\beta_2 \geq 0$; $\delta_t = 4.5$; normal errors.

Table E.28 Monte Carlo IR 27 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	140.78	71.14	241.16	280.34
RMSE ($\hat{\beta}$)	8.33	11.86	8.43	15.53	16.74
Var ($\hat{\beta}$)	69.29	53.55	21.66	17.18	17.86
Bias^2($\hat{\beta}$)	0.08	87.22	49.47	223.99	262.48
PRMSE	9.12	9.49	9.33	9.98	10.11
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	8.00	7.74	8.22	8.33
R^2	0.70				

* N=100; $\beta_2 \geq 3$; $\delta_t = 7.5$; normal errors.

Table E.29 Monte Carlo IR 31 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	205.15	68.75	56.32	82.21
RMSE ($\hat{\beta}$)	14.54	14.32	8.29	7.50	9.07
Var ($\hat{\beta}$)	211.31	205.05	48.95	39.05	41.62
Bias^2($\hat{\beta}$)	0.07	0.10	19.80	17.27	40.59
PRMSE	5.73	5.74	5.88	5.92	5.95
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.08	8.42	8.37	8.46
R^2	0.84				

* N=50; $\beta_2 \geq -10$; $\delta_t = -5.5$; normal errors.

Table E.30 Monte Carlo IR 32 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	192.14	71.78	88.19	106.80
RMSE ($\hat{\beta}$)	14.54	13.86	8.47	9.39	10.33
Var ($\hat{\beta}$)	211.31	187.55	46.68	39.79	40.78
Bias^2($\hat{\beta}$)	0.07	4.59	25.10	48.40	66.02
PRMSE	5.73	5.77	5.89	5.99	6.02
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.04	8.43	8.48	8.54
R^2	0.84				

* N=50; $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table E.31 Monte Carlo IR 33 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	191.92	74.20	97.93	115.63
RMSE ($\hat{\beta}$)	14.54	13.85	8.61	9.90	10.75
Var ($\hat{\beta}$)	211.31	185.83	46.88	40.26	41.19
$\text{Bias}^2(\hat{\beta})$	0.07	6.09	27.32	57.66	74.45
PRMSE	5.73	5.78	5.90	6.01	6.04
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	9.10	9.04	8.44	8.51	8.57
R^2	0.84				

* N=50; $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table E.32 Monte Carlo IR 34 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	194.46	76.68	114.82	136.76
RMSE ($\hat{\beta}$)	14.54	13.94	8.76	10.72	11.69
Var ($\hat{\beta}$)	211.31	181.37	46.87	40.17	40.81
$\text{Bias}^2(\hat{\beta})$	0.07	13.09	29.81	74.65	95.95
PRMSE	5.73	5.81	5.90	6.05	6.09
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	9.10	9.05	8.44	8.56	8.64
R^2	0.84				

* N=50; $\beta_2 \geq -3$; $\delta_t = 1.5$; normal errors.

Table E.33 Monte Carlo IR 35 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE $(\hat{\beta})$	211.38	203.29	79.54	134.79	165.12
RMSE $(\hat{\beta})$	14.54	14.26	8.92	11.61	12.85
Var $(\hat{\beta})$	211.31	177.97	47.32	40.18	41.48
Bias$^2(\hat{\beta})$	0.07	25.33	32.21	94.61	123.65
PRMSE	5.73	5.85	5.91	6.10	6.15
PRMSE $(\hat{p}+\hat{e})$			0.00	0.00	0.00
PRMSE$_0$	9.10	9.08	8.45	8.63	8.73
R^2	0.84				

* N=50; $\beta_2 \geq -1.5$; $\delta_t = 3$; normal errors.

Table E.34 Monte Carlo IR 36 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE $(\hat{\beta})$	211.38	220.02	85.91	183.62	231.47
RMSE $(\hat{\beta})$	14.54	14.83	9.27	13.55	15.21
Var $(\hat{\beta})$	211.31	175.62	47.39	40.32	41.80
Bias$^2(\hat{\beta})$	0.07	44.39	38.52	143.30	189.67
PRMSE	5.73	5.90	5.92	6.22	6.31
PRMSE $(\hat{p}+\hat{e})$			0.00	0.00	0.00
PRMSE$_0$	9.10	9.13	8.47	8.79	8.94
R^2	0.84				

* N=50; $\beta_2 \geq 0$; $\delta_t = 4.5$; normal errors.

Table E.35 Monte Carlo IR 37 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	280.38	99.25	285.24	313.90
RMSE ($\hat{\beta}$)	14.54	16.74	9.96	16.89	17.72
Var ($\hat{\beta}$)	211.31	173.47	49.69	40.74	42.17
Bias^2($\hat{\beta}$)	0.07	106.91	49.57	244.50	271.73
PRMSE	5.73	6.05	5.95	6.45	6.49
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.32	8.52	9.11	9.20
R^2	0.84				

* N=50: $\beta_2 \geq 3$; $\delta_t = 7.5$; normal errors.

Table E.36 Monte Carlo IR 31 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	205.15	53.75	44.96	77.98
RMSE ($\hat{\beta}$)	14.54	14.32	7.33	6.71	8.83
Var ($\hat{\beta}$)	211.31	205.05	23.31	19.59	20.17
Bias^2($\hat{\beta}$)	0.07	0.10	30.44	25.38	57.81
PRMSE	5.73	5.74	5.99	6.03	6.08
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.08	8.17	8.12	8.25
R^2	0.84				

* N=50: $\beta_2 \geq -10$: $\delta_t = -5.5$; normal errors.

Table E.37 Monte Carlo IR 32 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	192.14	62.13	79.58	105.99
RMSE ($\hat{\beta}$)	14.54	13.86	7.88	8.92	10.30
Var ($\hat{\beta}$)	211.31	187.55	22.11	19.85	20.03
Bias^2($\hat{\beta}$)	0.07	4.59	40.02	59.73	85.96
PRMSE	5.73	5.77	6.02	6.10	6.16
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.04	8.20	8.25	8.35
R^2	0.84				

* N=50: $\beta_2 \geq -5$; $\delta_t = -0.5$; normal errors.

Table E.38 Monte Carlo IR 33 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	191.92	66.69	92.68	119.20
RMSE ($\hat{\beta}$)	14.54	13.85	8.17	9.63	10.92
Var ($\hat{\beta}$)	211.31	185.83	22.19	19.93	20.16
Bias^2($\hat{\beta}$)	0.07	6.09	44.51	72.76	99.04
PRMSE	5.73	5.78	6.03	6.13	6.19
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.04	8.21	8.30	8.40
R^2	0.84				

* N=50; $\beta_2 \geq -4.5$; $\delta_t = 0$; normal errors.

Table E.39 Monte Carlo IR 34 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	194.46	71.47	110.08	141.36
RMSE ($\hat{\beta}$)	14.54	13.94	8.45	10.49	11.89
Var ($\hat{\beta}$)	211.31	181.37	22.15	19.93	20.07
Bias^2($\hat{\beta}$)	0.07	13.09	49.32	90.15	121.29
PRMSE	5.73	5.81	6.04	6.18	6.25
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.05	8.23	8.37	8.48
R^2	0.84				

* N=50; $\beta_2 \geq -3$; $\delta_t = 1.5$; normal errors.

Table E.40 Monte Carlo IR 35 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	203.29	76.97	130.60	184.18
RMSE ($\hat{\beta}$)	14.54	14.26	8.77	11.43	13.57
Var ($\hat{\beta}$)	211.31	177.97	22.37	19.96	20.29
Bias^2($\hat{\beta}$)	0.07	25.33	54.60	110.65	163.89
PRMSE	5.73	5.85	6.06	6.23	6.36
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.08	8.25	8.44	8.63
R^2	0.84				

* N=50; $\beta_2 \geq -1.5$; $\delta_t = 3$; normal errors.

Table E.41 Monte Carlo IR 36 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	220.02	89.42	181.18	271.64
RMSE ($\hat{\beta}$)	14.54	14.83	9.46	13.46	16.48
Var ($\hat{\beta}$)	211.31	175.62	22.33	20.00	20.29
Bias^2($\hat{\beta}$)	0.07	44.39	67.10	161.18	251.35
PRMSE	5.73	5.90	6.09	6.36	6.58
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.13	8.30	8.62	8.93
R^2	0.84				

* N=50: $\beta_2 \geq 0$; $\delta_t = 4.5$; normal errors.

Table E.42 Monte Carlo IR 37 – Restrictions on β_2 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	280.38	119.51	298.12	389.94
RMSE ($\hat{\beta}$)	14.54	16.74	10.93	17.27	19.75
Var ($\hat{\beta}$)	211.31	173.47	23.66	20.03	20.46
Bias^2($\hat{\beta}$)	0.07	106.91	95.86	278.09	369.49
PRMSE	5.73	6.05	6.16	6.66	6.87
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.32	8.41	9.02	9.33
R^2	0.84				

* N=50: $\beta_2 \geq 3$; $\delta_t = 7.5$; normal errors.

Table E.43 Monte Carlo IR 11 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	49.49	38.02	36.96	38.58
RMSE ($\hat{\beta}$)	7.04	7.04	6.17	6.08	6.21
Var ($\hat{\beta}$)	49.44	49.44	30.21	29.43	29.97
$\text{Bias}^2(\hat{\beta})$	0.05	0.05	7.82	7.54	8.60
PRMSE	11.58	11.58	11.65	11.66	11.65
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	7.54	7.54	7.52	7.50	7.52
R^2	0.77				

* N=150; $\beta_3 \le 10; \delta_t = -6.5$; normal errors.

Table E.44 Monte Carlo IR 12 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	49.27	38.08	40.83	40.80
RMSE ($\hat{\beta}$)	7.04	7.02	6.17	6.39	6.39
Var ($\hat{\beta}$)	49.44	49.22	30.18	29.52	29.62
$\text{Bias}^2(\hat{\beta})$	0.05	0.05	7.90	11.31	11.18
PRMSE	11.58	11.58	11.65	11.71	11.71
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	7.54	7.53	7.52	7.55	7.55
R^2	0.77				

* N=150; $\beta_3 \le 5; \delta_t = -1.5$; normal errors.

Table E.45 Monte Carlo IR 13 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	48.71	38.11	43.43	43.19
RMSE ($\hat{\beta}$)	7.04	6.98	6.17	6.59	6.57
Var ($\hat{\beta}$)	49.44	48.39	30.17	29.50	29.48
Bias^2($\hat{\beta}$)	0.05	0.32	7.94	13.93	13.71
PRMSE	11.58	11.60	11.65	11.77	11.77
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.60	7.59
R^2	0.77				

* N=150; $\beta_3 \leq 3.5$; $\delta_t = 0$; normal errors.

Table E.46 Monte Carlo IR 14 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	48.96	38.21	45.73	44.71
RMSE ($\hat{\beta}$)	7.04	7.00	6.18	6.76	6.69
Var ($\hat{\beta}$)	49.44	48.13	30.18	29.64	29.61
Bias^2($\hat{\beta}$)	0.05	0.82	8.03	16.09	15.10
PRMSE	11.58	11.62	11.65	11.82	11.80
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.52	7.64	7.62
R^2	0.77				

* N=150; $\beta_3 \leq 3$; $\delta_t = 0.5$; normal errors.

Table E.47 Monte Carlo IR 15 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	56.96	38.25	55.44	55.46
RMSE ($\hat{\beta}$)	7.04	7.55	6.18	7.45	7.45
Var ($\hat{\beta}$)	49.44	47.78	30.19	30.12	30.10
Bias^2($\hat{\beta}$)	0.05	9.18	8.07	25.33	25.36
PRMSE	11.58	11.86	11.65	12.07	12.07
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.69	7.52	7.81	7.81
R^2	0.77				

* N=150; $\beta_3 \leq 1$; $\delta_t = 2.5$; normal errors.

Table E.48 Monte Carlo IR 16 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	65.62	38.34	65.11	62.36
RMSE ($\hat{\beta}$)	7.04	8.10	6.19	8.07	7.90
Var ($\hat{\beta}$)	49.44	47.80	30.19	31.08	30.89
Bias^2($\hat{\beta}$)	0.05	17.82	8.15	34.04	31.48
PRMSE	11.58	12.09	11.65	12.28	12.22
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.86	7.52	7.97	7.92
R^2	0.77				

* N=150; $\beta_3 \leq 0$; $\delta_t = 3.5$; normal errors.

Table E.49 Monte Carlo IR 17 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	49.49	77.24	38.44	77.77	77.64
RMSE ($\hat{\beta}$)	7.04	8.79	6.20	8.82	8.81
Var ($\hat{\beta}$)	49.44	47.80	30.20	33.72	33.71
Bias^2($\hat{\beta}$)	0.05	29.44	8.24	44.05	43.93
PRMSE	11.58	12.38	11.65	12.54	12.54
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	8.08	7.52	8.15	8.15
R^2	0.77				

* N=150; $\beta_3 \le -1$; $\delta_t = 4.5$; normal errors.

Table E.50 Monte Carlo IR 11 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	49.49	39.32	37.88	41.53
RMSE ($\hat{\beta}$)	7.04	7.04	6.27	6.15	6.44
Var ($\hat{\beta}$)	49.44	49.44	17.97	17.04	17.55
Bias^2($\hat{\beta}$)	0.05	0.05	21.35	20.84	23.98
PRMSE	11.58	11.58	11.68	11.71	11.70
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.54	7.49	7.47	7.51
R^2	0.77				

* N=150; $\beta_3 \le 10$; $\delta_t = -6.5$; normal errors.

Table E.51 Monte Carlo IR 12 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	49.27	39.55	46.66	47.29
RMSE ($\hat{\beta}$)	7.04	7.02	6.29	6.83	6.88
Var ($\hat{\beta}$)	49.44	49.22	17.92	17.08	17.22
Bias^2($\hat{\beta}$)	0.05	0.05	21.63	29.58	30.06
PRMSE	11.58	11.58	11.68	11.82	11.83
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.53	7.49	7.59	7.60
R^2	0.77				

* N=150: $\beta_3 \leq 5$; $\delta_t = -1.5$; normal errors.

Table E.52 Monte Carlo IR 13 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	48.71	39.67	51.25	50.62
RMSE ($\hat{\beta}$)	7.04	6.98	6.30	7.16	7.12
Var ($\hat{\beta}$)	49.44	48.39	17.90	17.10	17.11
Bias^2($\hat{\beta}$)	0.05	0.32	21.77	34.16	33.52
PRMSE	11.58	11.60	11.68	11.93	11.91
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.49	7.67	7.66
R^2	0.77				

* N=150: $\beta_3 \leq 3.5$; $\delta_t = 0$; normal errors.

Table E.53 Monte Carlo IR 14 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	48.96	40.00	56.48	55.07
RMSE ($\hat{\beta}$)	7.04	7.00	6.32	7.52	7.42
Var ($\hat{\beta}$)	49.44	48.13	17.90	17.13	17.21
Bias^2 ($\hat{\beta}$)	0.05	0.82	22.10	39.35	37.86
PRMSE	11.58	11.62	11.68	12.05	12.01
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.52	7.49	7.77	7.75
R^2	0.77				

* N=150; $\beta_3 \le 3$; $\delta_i = 0.5$; normal errors.

Table E.54 Monte Carlo IR 15 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	56.96	40.15	62.35	66.29
RMSE ($\hat{\beta}$)	7.04	7.55	6.34	7.90	8.14
Var ($\hat{\beta}$)	49.44	47.78	17.92	17.02	17.09
Bias^2 ($\hat{\beta}$)	0.05	9.18	22.23	45.33	49.20
PRMSE	11.58	11.86	11.68	12.20	12.29
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.54	7.69	7.49	7.88	7.96
R^2	0.77				

* N=150; $\beta_3 \le 1$; $\delta_i = 2.5$; normal errors.

Table E.55 Monte Carlo IR 16 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	65.62	40.48	78.16	74.36
RMSE ($\hat{\beta}$)	7.04	8.10	6.36	8.84	8.62
Var ($\hat{\beta}$)	49.44	47.80	17.93	17.14	17.17
Bias^2($\hat{\beta}$)	0.05	17.82	22.55	61.02	57.20
PRMSE	11.58	12.09	11.68	12.58	12.48
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
$PRMSE_0$	7.54	7.86	7.50	8.18	8.11
R^2	0.77				

* N=150; $\beta_3 \le 0$; $\delta_t = 3.5$; normal errors.

Table E.56 Monte Carlo IR 17 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	49.49	77.24	40.83	87.22	85.06
RMSE ($\hat{\beta}$)	7.04	8.79	6.39	9.34	9.22
Var ($\hat{\beta}$)	49.44	47.80	17.94	17.22	17.17
Bias^2($\hat{\beta}$)	0.05	29.44	22.89	70.00	67.89
PRMSE	11.58	12.38	11.68	12.80	12.75
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
$PRMSE_0$	7.54	8.08	7.50	8.35	8.31
R^2	0.77				

* N=150; $\beta_3 \le -1$; $\delta_t = 4.5$; normal errors.

Table E.57 Monte Carlo IR 21 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	69.38	46.82	44.42	48.44
RMSE ($\hat{\beta}$)	8.33	8.33	6.84	6.67	6.96
Var ($\hat{\beta}$)	69.29	69.29	35.62	33.97	35.04
Bias^2($\hat{\beta}$)	0.08	0.08	11.20	10.46	13.40
PRMSE	9.12	9.12	9.19	9.22	9.21
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.81	7.73	7.71	7.73
R^2	0.70				

* N=100; $\beta_3 \le 10$; $\delta_t = -6.5$; normal errors.

Table E.58 Monte Carlo IR 22 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	68.32	47.00	52.97	53.30
RMSE ($\hat{\beta}$)	8.33	8.27	6.86	7.28	7.30
Var ($\hat{\beta}$)	69.29	68.26	35.55	34.21	34.43
Bias^2($\hat{\beta}$)	0.08	0.06	11.45	18.77	18.87
PRMSE	9.12	9.12	9.19	9.29	9.29
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.80	7.73	7.76	7.76
R^2	0.70				

* N=100; $\beta_3 \le 5$; $\delta_t = -1.5$; normal errors.

Table E.59 Monte Carlo IR 23 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	66.94	47.10	57.30	56.80
RMSE ($\hat{\beta}$)	8.33	8.18	6.86	7.57	7.54
Var ($\hat{\beta}$)	69.29	66.64	35.52	34.23	34.21
Bias^2($\hat{\beta}$)	0.08	0.50	11.58	23.07	22.59
PRMSE	9.12	9.15	9.19	9.36	9.35
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.78	7.73	7.80	7.79
R^2	0.70				

* N=100; $\beta_3 \le 3.5$; $\delta_t = 0$; normal errors.

Table E.60 Monte Carlo IR 24 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	67.05	47.38	61.66	60.14
RMSE ($\hat{\beta}$)	8.33	8.19	6.88	7.85	7.75
Var ($\hat{\beta}$)	69.29	65.90	35.54	34.44	34.48
Bias^2($\hat{\beta}$)	0.08	1.15	11.84	27.22	25.65
PRMSE	9.12	9.17	9.20	9.42	9.40
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.79	7.73	7.84	7.83
R^2	0.70				

* N=100; $\beta_3 \le 3$; $\delta_t = 0.5$; normal errors.

Table E.61 Monte Carlo IR 25 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	75.73	47.51	72.34	74.62
RMSE ($\hat{\beta}$)	8.33	8.70	6.89	8.51	8.64
Var ($\hat{\beta}$)	69.29	65.15	35.57	36.88	37.03
$\text{Bias}^2(\hat{\beta})$	0.08	10.58	11.94	35.45	37.59
PRMSE	9.12	9.37	9.20	10.20	10.20
PRMSE ($\hat{p}+\hat{e}$)			0.00	3.31	3.24
PRMSE_0	7.81	7.89	7.73	8.31	8.32
R^2	0.70				

* N=100; $\beta_3 \leq 1$; $\delta_t = 2.5$; normal errors.

Table E.62 Monte Carlo IR 26 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	85.56	47.77	86.97	82.79
RMSE ($\hat{\beta}$)	8.33	9.25	6.91	9.33	9.10
Var ($\hat{\beta}$)	69.29	65.11	35.59	38.11	37.97
$\text{Bias}^2(\hat{\beta})$	0.08	20.45	12.18	48.86	44.81
PRMSE	9.12	9.55	9.20	11.67	11.70
PRMSE ($\hat{p}+\hat{e}$)			0.00	6.12	6.27
PRMSE_0	7.81	7.99	7.73	9.28	9.29
R^2	0.70				

* N=100; $\beta_3 \leq 0$; $\delta_t = 3.5$; normal errors.

Table E.63 Monte Carlo IR 27 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	69.38	98.75	48.05	96.98	96.37
RMSE ($\hat{\beta}$)	8.33	9.94	6.93	9.85	9.82
Var ($\hat{\beta}$)	69.29	65.10	35.62	38.59	38.88
Bias^2($\hat{\beta}$)	0.08	33.65	12.44	58.40	57.49
PRMSE	9.12	9.78	9.20	11.89	13.10
PRMSE ($\hat{p}+\hat{e}$)			0.00	6.31	8.30
PRMSE$_0$	7.81	8.14	7.73	9.45	10.26
R^2	0.70				

* N=100; $\beta_3 \le -1; \delta_t = 4.5$; normal errors.

Table E.64 Monte Carlo IR 21 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	69.38	48.19	45.66	52.89
RMSE ($\hat{\beta}$)	8.33	8.33	6.94	6.76	7.27
Var ($\hat{\beta}$)	69.29	69.29	20.58	19.04	19.76
Bias^2($\hat{\beta}$)	0.08	0.08	27.60	26.62	33.14
PRMSE	9.12	9.12	9.26	9.31	9.30
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.81	7.68	7.67	7.69
R^2	0.70				

* N=100; $\beta_3 \le 10; \delta_t = -6.5$; normal errors.

Table E.65 Monte Carlo IR 22 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE $(\hat{\beta})$	69.38	68.32	48.71	60.21	62.45
RMSE $(\hat{\beta})$	8.33	8.27	6.98	7.76	7.90
Var $(\hat{\beta})$	69.29	68.26	20.46	19.11	19.35
Bias$^2(\hat{\beta})$	0.08	0.06	28.26	41.10	43.10
PRMSE	9.12	9.12	9.26	9.43	9.45
PRMSE $(\hat{p}+\hat{e})$			0.00	0.00	0.00
PRMSE$_0$	7.81	7.80	7.67	7.75	7.78
R^2	0.70				

* N=100; $\beta_3 \leq 5$; $\delta_t = -1.5$; normal errors.

Table E.66 Monte Carlo IR 23 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE $(\hat{\beta})$	69.38	66.94	48.99	67.50	66.82
RMSE $(\hat{\beta})$	8.33	8.18	7.00	8.22	8.17
Var $(\hat{\beta})$	69.29	66.44	20.41	19.20	19.24
Bias$^2(\hat{\beta})$	0.08	0.50	28.58	48.31	47.58
PRMSE	9.12	9.15	9.26	9.54	9.53
PRMSE $(\hat{p}+\hat{e})$			0.00	0.00	0.00
PRMSE$_0$	7.81	7.78	7.67	7.83	7.82
R^2	0.70				

* N=100; $\beta_3 \leq 3.5$; $\delta_t = 0$; normal errors.

Table E.67 Monte Carlo IR 24 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	67.05	49.76	75.58	75.16
RMSE ($\hat{\beta}$)	8.33	8.19	7.05	8.69	8.67
Var ($\hat{\beta}$)	69.29	65.90	20.42	19.23	19.40
Bias^2 ($\hat{\beta}$)	0.08	1.15	29.33	56.36	55.77
PRMSE	9.12	9.17	9.26	9.68	9.67
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.79	7.68	7.91	7.91
R^2	0.70				

* N=100; $\beta_3 \leq 3$; $\delta_t = 0.5$; normal errors.

Table E.68 Monte Carlo IR 25 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	75.73	50.10	81.15	87.86
RMSE ($\hat{\beta}$)	8.33	8.70	7.08	9.01	9.37
Var ($\hat{\beta}$)	69.29	65.15	20.46	19.17	19.23
Bias^2 ($\hat{\beta}$)	0.08	10.58	29.64	61.98	68.63
PRMSE	9.12	9.37	9.26	9.78	9.90
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.89	7.68	7.97	8.05
R^2	0.70				

* N=100; $\beta_3 \leq 1$; $\delta_t = 2.5$; normal errors.

Table E.69 Monte Carlo IR 26 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	85.56	50.87	102.91	101.02
RMSE ($\hat{\beta}$)	8.33	9.25	7.13	10.14	10.05
Var ($\hat{\beta}$)	69.29	65.11	20.48	19.26	19.39
Bias^2($\hat{\beta}$)	0.08	20.45	30.39	83.64	81.64
PRMSE	9.12	9.55	9.27	10.16	10.12
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	7.99	7.68	8.22	8.20
R^2	0.70				

* N=100; $\beta_3 \leq 0$; $\delta_t = 3.5$; normal errors.

Table E.70 Monte Carlo IR 27 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	69.38	98.75	51.72	116.14	110.77
RMSE ($\hat{\beta}$)	8.33	9.94	7.19	10.78	10.52
Var ($\hat{\beta}$)	69.29	65.10	20.51	19.39	19.31
Bias^2($\hat{\beta}$)	0.08	33.65	31.21	96.75	91.46
PRMSE	9.12	9.78	9.27	10.39	10.30
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	7.81	8.14	7.69	8.37	8.31
R^2	0.70				

* N=100; $\beta_3 \leq -1$; $\delta_t = 4.5$; normal errors.

Table E.71 Monte Carlo IR 31 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	210.33	80.28	75.14	79.27
RMSE ($\hat{\beta}$)	14.54	14.50	8.96	8.67	8.90
Var ($\hat{\beta}$)	211.31	210.27	48.41	41.90	43.65
Bias^2($\hat{\beta}$)	0.07	0.06	31.87	33.24	35.61
PRMSE	5.73	5.74	5.90	5.96	5.95
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.09	8.45	8.38	8.45
R^2	0.84				

* N=50; $\beta_3 \leq 10; \delta_t = -6.5$; normal errors.

Table E.72 Monte Carlo IR 32 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	204.22	78.74	83.37	89.51
RMSE ($\hat{\beta}$)	14.54	14.29	8.87	9.13	9.46
Var ($\hat{\beta}$)	211.31	203.66	47.31	42.12	42.80
Bias^2($\hat{\beta}$)	0.07	0.56	31.44	41.25	46.71
PRMSE	5.73	5.76	5.90	6.02	6.04
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.04	8.44	8.50	8.58
R^2	0.84				

* N=50; $\beta_3 \leq 5; \delta_t = -1.5$; normal errors.

Table E.73 Monte Carlo IR 33 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	202.70	78.46	93.30	96.93
RMSE ($\hat{\beta}$)	14.54	14.24	8.86	9.66	9.85
Var ($\hat{\beta}$)	211.31	200.37	46.98	42.47	43.12
Bias^2($\hat{\beta}$)	0.07	2.33	31.48	50.82	50.82
PRMSE	5.73	5.78	5.91	6.07	6.07
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.03	8.44	8.62	8.62
R^2	0.84				

* N=50; $\beta_3 \le 3.5$; $\delta_t = 0$; normal errors.

Table E.74 Monte Carlo IR 34 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	202.92	79.00	102.85	108.84
RMSE ($\hat{\beta}$)	14.54	14.25	8.89	10.14	10.43
Var ($\hat{\beta}$)	211.31	199.46	47.08	42.53	43.21
Bias^2($\hat{\beta}$)	0.07	3.46	31.92	60.32	65.63
PRMSE	5.73	5.79	5.91	6.13	6.15
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.03	8.45	8.74	8.81
R^2	0.84				

* N=50; $\beta_3 \le 3$; $\delta_t = 0.5$; normal errors.

Table E.75 Monte Carlo IR 35 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	209.51	79.54	110.71	121.20
RMSE ($\hat{\beta}$)	14.54	14.47	8.92	10.52	11.01
Var ($\hat{\beta}$)	211.31	196.89	47.32	42.83	43.07
Bias^2($\hat{\beta}$)	0.07	12.62	32.21	67.88	78.13
PRMSE	5.73	5.86	5.91	6.17	6.23
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.07	8.45	8.83	8.95
R^2	0.84				

* N=50; $\beta_3 \le 1$; $\delta_t = 2.5$; normal errors.

Table E.76 Monte Carlo IR 36 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	217.03	80.61	136.24	145.83
RMSE ($\hat{\beta}$)	14.54	14.73	8.98	11.67	12.08
Var ($\hat{\beta}$)	211.31	196.06	47.49	42.98	43.65
Bias^2($\hat{\beta}$)	0.07	20.97	33.12	93.27	102.18
PRMSE	5.73	5.91	5.92	6.33	6.38
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.14	8.47	9.12	9.22
R^2	0.84				

* N=50; $\beta_3 \le 0$; $\delta_t = 3.5$; normal errors.

Table E.77 Monte Carlo IR 37 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME2-PM	GME2-R1	GME2-R2
MSE ($\hat{\beta}$)	211.38	227.83	82.07	158.51	133.29
RMSE ($\hat{\beta}$)	14.54	15.09	9.06	12.59	11.55
Var ($\hat{\beta}$)	211.31	195.58	47.71	43.39	43.89
Bias^2 ($\hat{\beta}$)	0.07	32.26	34.36	115.12	89.40
PRMSE	5.73	6.98	5.92	6.46	6.29
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
$PRMSE_0$	9.10	9.22	8.49	9.36	9.09
R^2	0.84				

* N=50; $\beta_3 \leq -1$; $\delta_t = 4.5$; normal errors.

Table E.78 Monte Carlo IR 31 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	210.33	75.70	72.64	80.04
RMSE ($\hat{\beta}$)	14.54	14.50	8.70	8.52	8.95
Var ($\hat{\beta}$)	211.31	210.27	23.38	19.52	20.09
Bias^2 ($\hat{\beta}$)	0.07	0.06	52.33	53.12	59.95
PRMSE	5.73	5.74	6.04	6.10	6.11
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
$PRMSE_0$	9.10	9.09	8.22	8.17	8.29
R^2	0.84				

* N=50; $\beta_3 \leq 10$; $\delta_t = -6.5$; normal errors.

Table E.79 Monte Carlo IR 32 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	204.22	73.98	84.25	94.93
RMSE ($\hat{\beta}$)	14.54	14.29	8.60	9.18	9.74
Var ($\hat{\beta}$)	211.31	203.66	22.36	19.57	19.80
Bias^2($\hat{\beta}$)	0.07	0.56	51.62	64.68	75.13
PRMSE	5.73	5.76	6.04	6.16	6.22
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	9.10	9.04	8.20	8.35	8.49
R^2	0.84				

* N=50; $\beta_3 \le 5$; $\delta_t = -1.5$; normal errors.

Table E.80 Monte Carlo IR 33 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	202.70	74.04	98.26	107.68
RMSE ($\hat{\beta}$)	14.54	14.24	8.60	9.91	10.38
Var ($\hat{\beta}$)	211.31	200.37	22.09	19.68	19.99
Bias^2($\hat{\beta}$)	0.07	2.33	51.95	78.59	87.69
PRMSE	5.73	5.78	6.04	6.24	6.29
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE_0	9.10	9.03	8.21	8.53	8.65
R^2	0.84				

* N=50; $\beta_3 \le 3.5$; $\delta_t = 0$; normal errors.

Table E.81 Monte Carlo IR 34 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	202.92	75.72	110.22	125.47
RMSE ($\hat{\beta}$)	14.54	14.25	8.70	10.50	11.20
Var ($\hat{\beta}$)	211.31	199.46	22.17	19.67	19.96
Bias^2($\hat{\beta}$)	0.07	3.46	53.55	90.56	105.51
PRMSE	5.73	5.79	6.05	6.32	6.40
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.03	8.23	8.68	8.87
R^2	0.84				

* N=50; $\beta_3 \leq 3$; $\delta_t = 0.5$; normal errors.

Table E.82 Monte Carlo IR 35 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	209.51	76.97	124.00	141.04
RMSE ($\hat{\beta}$)	14.54	14.47	8.77	11.14	11.88
Var ($\hat{\beta}$)	211.31	196.89	22.37	19.85	19.97
Bias^2($\hat{\beta}$)	0.07	12.62	54.60	104.15	121.07
PRMSE	5.73	5.86	6.06	6.40	6.50
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.07	8.25	8.85	9.05
R^2	0.84				

* N=50; $\beta_3 \leq 1$; $\delta_t = 2.5$; normal errors.

Table E.83 Monte Carlo IR 36 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	217.03	80.06	154.88	182.99
RMSE ($\hat{\beta}$)	14.54	14.73	8.95	12.45	13.53
Var ($\hat{\beta}$)	211.31	196.06	22.50	19.80	20.14
Bias^2($\hat{\beta}$)	0.07	20.97	57.56	135.08	162.84
PRMSE	5.73	5.91	6.07	6.59	6.74
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.14	8.30	9.21	9.52
R^2	0.84				

* N=50; $\beta_3 \le 0$; $\delta_t = 3.5$; normal errors.

Table E.84 Monte Carlo IR 37 – Restrictions on β_3 (See Table 4.1 for design)

	OLS	IRLS	GME3-PM	GME3-R1	GME3-R2
MSE ($\hat{\beta}$)	211.38	227.83	84.28	187.47	172.67
RMSE ($\hat{\beta}$)	14.54	15.09	9.18	13.69	13.14
Var ($\hat{\beta}$)	211.31	195.58	22.64	19.88	20.47
Bias^2($\hat{\beta}$)	0.07	32.26	61.64	167.59	152.20
PRMSE	5.73	6.98	6.10	6.78	6.67
PRMSE ($\hat{p}+\hat{e}$)			0.00	0.00	0.00
PRMSE$_0$	9.10	9.22	8.35	9.57	9.41
R^2	0.84				

* N=50; $\beta_3 \le -1$; $\delta_t = 4.5$; normal errors.

VITA

Randall C. Campbell was born in Hammond, Indiana. He was raised in Indianapolis, Indiana, with his sister, Kathy, and his brothers, Scott and Darin, by their parents, Jean A. and Charles S. Campbell. He delivered papers for the Indianapolis Star/News from sixth grade through high school. Randall began school at Purdue University in August 1985. He received his degree of Bachelor of Science in Industrial Management from Purdue University in West Lafayette, Indiana, in May 1989. He worked as a statistician in the Indiana Department of Public Welfare from June 1990 until July 1992. In August 1992 he began work on a degree of Doctor of Philosophy in Economics at Louisiana State University in Baton Rouge, Louisiana. Randall received his degree of Master of Science in Economics from Louisiana State University in May 1994. He married Rebecca J. Campbell, whom he met at L.S.U. in the economics program, on July 10, 1994. Randall currently works as a demand analyst for SBC Communications, Inc. and will complete his degree of Doctor of Philosophy in Economics in May 1999. After completing his doctoral work. Randall intends to continue his research in econometrics and hopes to obtain a position in academics. In addition to his econometric research, Randall enjoys taking trips with his wife, visiting major league baseball stadiums, going to movies and barbecuing. Randall and Rebecca are raising two Vizslas, Bleau and Duquesa. They have recently been elected as the first members of the Texas Gulf Coast Vizsla Club's Vizsla rescue Hall of Fame.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Randall C. Campbell

Major Field: Economics

Title of Dissertation: An Empirical Examination of Maximum Entropy Estimation

Approved:

R. Carter Hill

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

W. Douglas McMillin

M. R. Terrell

B. Charlene Henderson

Date of Examination:

2/18/99

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989