Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences

Chandra R. Bhat
Dept of Civil Engineering, ECJ 6.8, Univ. of Texas at Austin, Austin, TX, 78712
Phone: (512) 471-4535, Fax: (512) 475-8744, Email: bhat@mail.utexas.edu

Abstract

The use of simulation techniques has been increasing in recent years in the transportation and related fields to accommodate flexible and behaviorally realistic structures for analysis of decision processes. This paper proposes a randomized and scrambled version of the Halton sequence for use in simulation estimation of discrete choice models. The scrambling of the Halton sequence is motivated by the rapid deterioration of the standard Halton sequence's coverage of the integration domain in high dimensions of integration. The randomization of the sequence is motivated from a need to statistically compute the simulation variance of model parameters. The resulting hybrid sequence combines the good coverage property of quasi-Monte Carlo sequences with the ease of estimating simulation error using traditional Monte Carlo methods. The paper develops an evaluation framework for assessing the performance of the traditional pseudo-random sequence, the standard Halton sequence, and the scrambled Halton sequence. The results of computational experiments indicate that the scrambled Halton sequence performs better than the standard Halton sequence and the traditional pseudo-random sequence for simulation estimation of models with high dimensionality of integration.

Keywords: Maximum simulated likelihood estimation, pseudo-random sequences, quasi-random sequences, hybrid sequences, multinomial probit model, mixed logit model, mixed probit model.

1. Introduction

The use of simulation methods to estimate econometric models with analytically intractable criterion functions has grown rapidly in the past decade. To some extent, this has been due to the availability of faster computers. But researchers have also strived to improve the efficiency of simulation techniques so that, in combination with faster computing power, simulation methods are effective and feasible to implement even for large scale model systems.

Simulation methods are now routinely applied in the estimation of limited-dependent and discrete choice models in the economics, marketing, and transportation fields. Such methods obviate the need to impose *a priori* behavioral restrictions on the mechanism underlying the decision process being examined. Many analytically tractable models, though elegant and simple in structure, maintain restrictions that are difficult to justify. For example, in a discrete choice context, the multinomial logit model has a simple form, but is saddled with the independence from irrelevant alternatives (IIA) property. Relaxing such rigid behavioral restrictions tends to lead to analytically intractable models. Some of these analytically intractable models, such as the heteroscedastic extreme value (HEV) model formulated by Bhat in 1995, involve a single-dimensional integral that can be evaluated accurately using well-established quadrature techniques. However, quadrature formulas are unable to compute integrals with sufficient precision and speed for estimation of models with higher than 1-2 dimensions of integration (see Hajivassiliou and Ruud, 1994). In fact, because of the curse of dimensionality, the quadrature method is literally useless in high dimensions. Two broad simulation methods are available in high dimensions: (a) Monte Carlo methods and (b) Quasi-Monte Carlo methods. Each of these is discussed in the next two sections. Section 1.3 discusses a hybrid of the Monte-Carlo and Quasi-Monte Carlo methods.

1.1 The Monte-Carlo method

The Monte-Carlo simulation method (or “the method of statistical trials”) to evaluating multidimensional integrals entails computing the integrand at a sequence of “random” points and computing the average of the integrand values. The basic principle is to replace a continuous average by a discrete average over randomly chosen points. Of course, in actual implementation, truly random sequences are not available; instead, deterministic pseudo-random sequences which appear random when subjected to simple statistical tests are used (see Niederreiter, 1995 for a discussion of pseudo-random sequence generation). This pseudo-Monte Carlo (or PMC) method has a slow asymptotic convergence rate with the expected integration error of the order of $N^{-0.5}$ in probability (N being the number of pseudo-random points drawn from the s-dimensional integration space). Thus, to obtain an added decimal digit of accuracy, the number of draws needs to be increased hundred fold. However, the PMC method's convergence rate is remarkable in that it is applicable for a wide class of integrands (the only requirement is that the integrand have a finite variance; see Spanier and Maize, 1994). Further, the integration error can be easily estimated using the sample values and invoking the central limit theorem, or by replicating the evaluation of the integral several times using independent sets of PMC draws and computing the variance in the different estimates of the integrand.

1.2 The quasi-Monte Carlo method

The quasi-Monte Carlo method is similar to the Monte Carlo method in that it evaluates a multidimensional integral by replacing it with an average of values of the integrand computed at discrete points. However, rather than using pseudo-random sequences for the discrete points, the quasi-Monte Carlo approach uses “cleverly” crafted non-random and more uniformly distributed

sequences (labeled as quasi-Monte Carlo or QMC sequences) within the domain of integration. The underlying idea of the method is that it is really inconsequential whether the discrete points are truly random; of primary importance is the even distribution (or maximal spread) of the points in the integration space. The convergence rate for quasi-random sequences is, in general, faster than for pseudo-random sequences. In particular, the theoretical upper bound of the integration error for reasonably well-behaved smooth functions is of the order of N^{-1} in the QMC method, where N is the number of quasi-random integration points. However, a limitation of the QMC method is that there is no practical way of statistically estimating the error in integration, because of the deterministic nature of the QMC sequences. Theoretical results are available to compute the upper bound of the error using a well-known theorem in number theory referred to as the Koksma-Hlawka inequality (Zaremba, 1968). This inequality provides the upper bound of the integration error as a function of the discrepancy of the QMC sequence and the variation of the integrand in the sense of Hardy and Krause (Niederreiter, 1992, Chapter 4; a sequence with high discrepancy is not as well distributed over the integration space as a sequence with low discrepancy). However, computing this theoretical error bound is not practical and, in fact, is much more complicated than evaluating the integral itself (Owen, 1997; Tuffin, 1996). Besides, the upper bound of the integration error from the theoretical result can be very conservative (Owen, 1998).

1.3 The hybrid method and the research objectives of the current paper

The discussion in the previous two sections indicates that quasi-MC sequences provide better accuracy than PMC sequences, while PMC sequences provide the ability to estimate the integration error easily. To take advantage of the strengths of each of these two methods, it is

desirable to develop hybrid or randomized QMC sequences (see Owen, 1995 for a history of such hybrid sequences). The essential idea is to introduce some randomness into a QMC sequence, while preserving the equidistribution property of the underlying QMC sequence. Then, by using several independent randomized QMC sequences, one can use standard statistical methods to estimate integration error.

The development of randomized QMC (RQMC) sequences has been the subject of interest in number theory for several decades now (see Cranley and Patterson, 1976 for one of the earliest references to such sequences). However, it is only in the past few years that number theorists have made significant advances in RQMC techniques (Owen, 1997, 1998; Morohosi and Fushimi, 1998; Hickernell and Yue, 2000; Yue and Hickernell, 2001). Much of this literature is focused on theoretical analysis to prove that specific types of randomization techniques retain the good properties of the underlying QMC sequence that form the basis for randomization. Some studies (for example, Tuffin, 1996 and Okten, 1996) have also conducted numerical experiments to examine the variance reduction due to the use of RQMC sequences compared to PMC sequences.

All the number-theoretical studies discussed above have focused on the use of RQMC sequences for accurate evaluation of a single multidimensional integral. In contrast, the focus of the current research is on the use of RQMC sequences for simulation estimation of an econometric discrete choice model. In such a simulation context, the objective is to accurately estimate underlying model parameters through the evaluation of multiple multidimensional integrals, each of which involves a parameterization of the model parameters and the data.

The specific objectives of this study are threefold. The first objective is to propose the use of a randomized and scrambled version of the Halton sequence for simulation estimation of

econometric models. The Halton sequence is a QMC sequence that belongs to the family of r-*adic* expansion of integers. The second objective is to develop an evaluation framework for comparing the performance of alternative simulation methods in a discrete choice context. The third objective is to undertake computational experiments to compare the performance of the pseudo-random sequence, the randomized standard Halton sequence, and the randomized scrambled Halton sequence.

The rest of the paper is structured as follows. The next section discusses the basic structure of the Halton sequence. Section 3 presents a randomization scheme for any QMC sequence. Section 4 describes the use of the randomized Halton sequence for choice model estimation. Section 5 describes the evaluation framework, including the design for the numerical experiments in the paper. Section 6 presents the computational results. The final section closes the paper.

2. The Halton Sequence

This section presents two versions of the Halton sequence: the standard Halton sequence (Section 2.1) and a scrambled Halton sequence (Section 2.2).

2.1 The standard Halton sequence

The standard Halton sequence is designed to span the domain of the *S*-dimensional unit cube uniformly and efficiently (the interval of each dimension of the unit cube is between 0 and 1). In one dimension, the standard Halton sequence is generated by choosing a prime number r ($r \geq 2$) and expanding the sequence of integers 1, 2, …g, ...G in terms of the base r:

$$g = \sum_{l=0}^{L} b_l(g) r^l, \text{ where } 0 \leq b_l(g) \leq r-1 \text{ and } r^L \leq g \leq r^{L+1}, \quad (1)$$

and l is an index of the power to which the base is raised. The first condition in the above equation indicates that the digits $b_l(g)$ in the base expansion cannot exceed the base value less one. The second condition determines the largest value of l (i.e., L) in the expansion of the integer g in the base r. Thus, g (g = 1, 2,…,G) can be represented in digitized form by the r-*adic* integer string $b_L(g)b_{L-1}(g)...b_1(g)b_0(g)$. The Halton sequence in the prime base r is obtained by taking the radical inverse of g (g = 1, 2,…,G) to the base r by reflecting through the radical point:

$$\varphi_r(g) = 0 \cdot b_0(g)b_1(g)...b_L(g) \text{ in base } r = \sum_{l=0}^{L} b_1(g)r^{-l-1}. \quad (2)$$

As an example of the sequence, consider the prime number 3 and the integer 5. The integer 5 can be written in base 3 as $5 = 1\times 3^1 + 2\times 3^0$. Thus, the digitized form of the integer 5 in base 3 is 12. The radical inverse function corresponding to the integer 5 in base 3 is then obtained by reflecting the number 12 about the "decimal point" as 0.21 and expanding 0.21 in base 3 as $2\times 3^{-1} + 1\times 3^{-2} = 7/9$. The number 7/9 then forms the fifth number in the Halton sequence. The values for other integers can be obtained similarly. The first 8 numbers in the sequence corresponding to base 3 are 1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9. As should be clear, new numbers added to the sequence tend to fill in the gaps left by the previous numbers, making for well-distributed coverage in the interval (0,1). The reader will also note that the sequence corresponding to the prime 3 comprises cycles of length 3 of monotonically increasing numbers (after adding a phantom value of zero at the beginning of the sequence). More generally, the sequence corresponding to the prime r comprises cycles of length r of monotonically increasing numbers.

The standard Halton sequence in S dimensions is obtained by pairing S one-dimensional sequences based on S pairwise relatively prime integers, $r_1, r_2, ..., r_S$ (usually the first S primes).

The Halton sequence is based on prime numbers, since the sequence based on a non-prime number will partition the unit space in the same way as each of the primes that contribute to the non-prime number. Thus, the *g*th multidimensional point of sequence is as follows:

$$\varphi(g) = (\varphi_{r_1}(g), \varphi_{r_2}(g), ..., \varphi_{r_S}(g)) . \quad (3)$$

The standard Halton sequence of length *N* is finally obtained as $\psi^{(N)} = [\varphi(1)', \varphi(2)', ..., \varphi(n)']'$. Of course, once the standard Halton sequence is generated for the unit cube, the corresponding sequence for any other integration domain can be obtained using the integral transform result. For example, the multivariate normal standard Halton points over the dimensional domain of the real line can be obtained as $X^{(N)} = \Phi^{-1}(\psi^{(N)})$, where Φ is the standard cumulative normal distribution function (see Fang and Wang, 1994; Chapter 4).

2.2 The Scrambled Halton sequence

A problem with the standard Halton sequence discussed above is that there is a strong correlation between higher coordinates of the sequence. This is because of the cycles of length *r* for the prime *r*. Thus, when two large prime-based sequences, associated with two high dimensions, are paired, the corresponding unit square face of the *S*-dimensional cube is sampled by points that lie on parallel lines. For example, the fourteenth dimension (corresponding to the prime number 43) and the fifteenth dimension (corresponding to the prime number 47) consist of 43 and 47 increasing numbers, respectively. This generates a correlation between the fourteenth and fifteenth coordinates of the sequence. This is illustrated diagrammatically in the first plot of Figure 1. The consequence is a rapid deterioration in the uniformity of the Halton sequence in high dimensions (the deterioration becomes clearly noticeable beyond five dimension).

Number theorists have proposed an approach to improve the uniformity of the Halton sequence in high dimensions. The basic method is to break the correlations between the coordinates of the standard Halton sequence by scrambling the cycles of length *r* for the prime *r*. This is accomplished by permutations of the coefficients b_l in the radical inverse function of Equation (2). The resulting scrambled Halton sequence for the prime *r* is written as:

$$\varphi_r(g) = \sum_{l=0}^{L} \sigma_r(b_l(g)) r^{-l-1}, \tag{4}$$

where σ_r is the operator of permutations on the digits of the expansion $b_1(g)$ (the standard Halton sequence is the special case of the scrambled Halton sequence with no scrambling of the digits $b_1(g)$). Different researchers (see Braaten and Weller, 1979; Hellekalek, 1984; Kocis and Whiten, 1997) have suggested different algorithms for arriving at the permutations of the coefficients b_l in Equation (4). The algorithm adopted in the current study corresponds to that of Braaten and Weller, who developed permutations that minimize the discrepancy of the resulting scrambled sequence for each prime *r*. These permutations are presented in the Appendix A for the first ten prime numbers. Braatan and Weller have also proved that their scrambled sequence retains the theoretically appealing N^{-1} order of integration error of the standard Halton sequence.

An example would be helpful in illustrating the scrambling procedure of Braaten and Weller. These researchers suggest the following permutation of (0,1,2) for the prime 3: (0,2,1). As indicated earlier, the 5th number in base 3 of the Halton sequence in digitized form is 0.21. When the permutation above is applied, the 5th number in the corresponding scrambled Halton sequence in digitized form is 0.21, which when expanded in base 3 translates to

$1\times 3^{-1}+2\times 3^{-2}=5/9$. The first 8 numbers in the scrambled sequence corresponding to base 3 are 2/3, 1/3, 2/9, 8/9, 5/9, 1/9, 7/9, 4/9.

The Braaten and Weller method involves different permutations for different prime numbers. As a result of this scrambling, the resulting sequence does not display strong correlation across dimensions as does the standard Halton sequence. This is illustrated in the second plot of Figure 1, which plots 150 scrambled Halton points in the fourteenth and fifteenth dimensions. A comparison of the two plots in Figure 1 clearly indicates the more uniform coverage of the scrambled Halton sequence relative to the standard Halton sequence.

3. Randomization of QMC Sequences

All QMC sequences (including the standard Halton and scrambled Halton sequences discussed above) are fundamentally deterministic. This deterministic nature of the sequences does not permit the practical estimation of the integration error. Theoretical results exist for estimating the integration error, but these are difficult to compute and can be very conservative.

The essential concept of randomizing QMC sequences is to introduce randomness into a deterministic QMC sequence that preserves the uniformly distributed and equidistribution properties of the underlying QMC sequence (see Shaw, 1988; Tuffin, 1996). One simple way to introduce randomness is based on the following idea. Let $\psi^{(N)}$ be a QMC sequence of length *N* over the *S*-dimensional cube $\{0,1\}^S$ and consider any *S*-dimensional uniformity distributed vector in the *S*-dimensional cube $(u\in\{0,1\}^S)$. $\psi^{(N)}$ is a matrix of dimension $N\times S$, and *u* is a vector of dimension $1\times S$. Construct a new sequence $\chi^{(N)}=\left\{\psi^{(N)}+u\otimes 1^{(N)}\right\}$, where {.} denotes the fractional part of the matrix within parenthesis, $\otimes$ represents the kronecker or tensor product, and $1^{(N)}$ is a unit column vector of size *N* (the kronecker product multiplies each element of *u*

with the vector $1^{(N)}$). The net result is a sequence $\chi^{(N)}$ whose elements χ_{ns} are obtained as $\psi_{ns} + u_s$ if $\psi_{ns} + u_s \leq 1$, and $\psi_{ns} + u_s - 1$ if $\psi_{ns} + u_s > 1$. It can be shown that the sequence $\chi^{(N)}$ so formed is also a QMC sequence of length *N* over the *S*-dimensional cube $\{0,1\}^S$. Tuffin provides a formal proof for this result, which is rather straightforward but tedious. Intuitively, the vector *u* simply shifts the points within each coordinate of the original QMC sequence $\psi^{(N)}$ by a certain value. Since all points within each coordinate are shifted by the same amount, the new sequence will preserve the equidistribution property of the original sequence. This is illustrated in Figure 2 in two dimensions. The first diagram in Figure 2 plots 100 points of the standard Halton sequence in the first two dimensions. The second diagram plots 100 points of the standard Halton sequence shifted by 0.5 in the first dimension and 0 in the second dimension. The result of shifting is as follows. For any point below 0.5 in the first dimension in the first diagram (such as the point marked 1), the point gets moved by 0.5 toward the right in the second diagram. For any point above 0.5 in the first dimension in the first diagram (such as the point marked 2), the point gets moved to the right, hits the right edge, bounces off this edge to the left edge, and is carried forward so that the total distance of the shift is 0.5 (another way to visualize this shift is to transform the unit square into a cylinder with the left and right edges "sewn" together; then the shifting entails moving points along the surface of the cylinder and perpendicular to the cylinder axis). Clearly, the two-dimensional plot in the second diagram of Figure 2 is also well-distributed because the relative positions of the points do not change from that in Figure 1; there is simply a shift of the overall pattern of points. The last diagram in Figure 2 plots the case where there is a shift in both dimensions; 0.5 in the first and 0.25 in the second. For the same reasons discussed in the context of the shift in one dimension, the sequence obtained by shifting in both dimensions is also well-distributed.

It should be clear from above that any vector $u \in \{0,1\}^S$ can be used to generate a new QMC sequence from an underlying QMC sequence. An obvious way of introducing randomness is then to randomly draw *u* from a multidimensional uniform distribution.

An important point to note here is that randomizing the standard Halton sequence as discussed earlier does not break the correlations in high dimensions because the randomization simply shifts all points in the same dimension by the same amount. Thus, randomized versions of the standard Halton sequence will suffer from the same problems of non-uniform coverage in high dimensions as the standard Halton sequence. To resolve the problem of non-uniform coverage in high dimensions, the scrambled Halton sequence needs to be used.

4. Use of the Randomized QMC Sequence in Model Estimation

The focus of theoretical and empirical work in the mathematical field, in the context of QMC sequences and their randomized versions, has been on evaluating a single multidimensional integral accurately. However, as discussed earlier, the focus in model estimation is on evaluating underlying model parameters that appear in the integrands of several multidimensional integrals. The intent in the latter case is to estimate the model parameters accurately, and not expressly on evaluating each integral itself accurately.

McFadden (1989) suggested simulation techniques using the pseudo-Monte Carlo (PMC) method for model parameter estimation in discrete response models. The method is based on drawing *S* random multidimensional points from the domain of integration for each observation in the estimation sample, and evaluating the contribution of each observation to the criterion function by averaging the value of the criterion function over the *N* random draws. The *N* random points are drawn independently for each observation. Because of the independent draws

across observations, and because each observation contributes to the criterion function, simulation errors in the evaluation of individual contributions average out. Thus, a much smaller number of draws per observation is sufficient to estimate the model parameters accurately than would be necessary to evaluate each individual observation's contribution to the criterion function accurately. The simulation variance in parameter estimates can be estimated in a straightforward manner in the PMC approach by generating several independent sets of $N*Q$ random multidimensional points (Q being the number of observations in the sample), estimating the model parameters for each set of points, and computing the standard deviation across the sets of model parameter estimates.

Bhat (2001) proposed a simulation approach for discrete response models that uses quasi-Monte Carlo (QMC) sequences. In his approach, Bhat generates a multidimensional QMC sequence of length $N*Q$, then uses the first N points to compute the contribution of the first observation to the criterion function, the second N points to compute the contribution of the second observation, and so on. This technique is also based on averaging out of simulation errors across observations. But rather than being random sets of points across observations, each set of N points fills in the gaps left by the sets of N points used for previous observations. Consequently, the averaging effect across observations is stronger when using QMC sequences than when using the PMC sequence. Train (1999) illustrates this effect very effectively and clearly by showing that the correlation in simulated probabilities between successive observations is negative and quite high when using the Halton sequence, but is close to zero when using the PMC sequence. In addition to the stronger averaging out effect across observations, the QMC sequence also provides more uniform coverage over the domain of the integration space for each observation compared to the PMC sequence. This enables more

accurate computations of the probabilities for each observation with fewer points (*i.e.*, smaller *N*) when QMC sequences are used. Of course, the presumption here is that the QMC sequence is well distributed in the integration space. As suggested earlier, this may not be so for the standard Halton sequence in high dimensions. Because the scrambled sequence provides better coverage, we expect it to perform better than the standard Halton sequence in high dimensions.

The simulation variance in parameters cannot be computed with the QMC sequences because of the deterministic nature of QMC sequences. However, randomized versions of the QMC can be used for this purpose. Essentially, several independent sets of *N***Q* multidimensional QMC points are generated using the randomization technique discussed in Section 3. The model parameters are estimated for each set of points and the variance across the sets of model parameter estimates is computed.

5. Evaluation Framework

A challenge in evaluating alternative simulation methods for estimation is to propose a setting in which the underlying model parameters associated with a given sample can be obtained exactly. Then, one can evaluate the ability of alternative simulation methods to recover the actual model parameters. In earlier simulation-related studies, the "true" model parameters have been obtained using a very large number of draws. For example, Bhat (2001) uses 20,000 random draws to estimate the parameters of a mixed logit model and declares the resulting estimates as the "true" parameter values. Similarly, Hajivassiliou et al. (1996) use 20,000 draws to estimate "true" parameters of an autoregressive random-effects model. However, the resulting parameter values may not be the true values because of simulation error even with 20,000 draws. This is high.

In the current paper, we propose a discrete choice model setting in which the true underlying model parameters associated with a given data sample can be obtained quickly and accurately. Specifically, we consider a random-coefficients formulation in which the utility that an individual q associates with alternative i is written as:

$$U_{qi} = \beta_q' x_{qi} + \varepsilon_{qi} \tag{5}$$

where x_{qi} is a column vector of exogenous attributes, β_q is a column vector corresponding coefficients that varies across individuals with density $f(\beta)$, and ε_{qi} is assumed to be an independently and identically distributed (across alternatives) normal error term with a variance of one. In addition, we assume that each element of β is normally distributed and that the elements of β are independent of one another and the error term. Next, define a vector θ that includes the means of standard deviations of the normal distributions characterizing the elements of β. With this specification, the choice probability of alternative i for individual q can be computer in one of two ways: (a) using a mixed probit framework or (b) using a multinomial probit (MNP) framework. These two methods are discussed in the two subsequent sections. Section 5.3 discusses the experimental design in the current paper.

5.1 The mixed probit framework

The unconditional choice probability of alternative i for individual q corresponding to Equation (5) is given by the following mixed probit formula (see Train, 1995):

$$P_{qi}(\theta) = \int_{\beta=-\infty}^{+\infty} L_{qi}(\beta) f(\beta \mid \theta) d(\beta), \tag{6}$$

where

$$L_{qi}(\beta) = \int_{\varepsilon_{qi}=-\infty}^{\infty} \left(\prod_{j \neq i} \left[\Phi\{\beta' x_{qi} - \beta' x_{qj} + \varepsilon_{qi}\} \right] \right) \phi(\varepsilon_{qi}) d(\varepsilon_{qi}) . \quad (7)$$

The integral involved in $L_{qi}(\beta)$ is only one-dimensional because of the IID normal distribution assumption for the error terms. The integral in Equation (6), on the other hand, is *S*-dimensional and is due to the random coefficients specification for the model parameters (*s* is the number of elements in β).

The focus of the experimental setting in the current paper is to evaluate the ability of alternative sequences (the PMC sequence, the randomized standard Halton sequence, and the randomized scrambled Halton sequence) to accurately recover the true mean and standard deviation of the random coefficients (*i.e.*, the true value for θ) using the maximum likelihood method. The single-dimensional integral in $L_{qi}(\beta)$ is evaluated accurately using Gauss-Hermite quadrature, so that differences in the ability to recover the random coefficients can be attributed entirely to the sequence used in the evaluation of the multidimensional integral in Equation (6). Thus, for each draw of β from an underlying θ vector, $L_{qi}(\beta)$ is evaluated using the quadrature method. The draws of β for an underlying value of the vector θ are based on the sequences discussed earlier. The sequences are generated using specialized code written by the author in the GAUSS matrix programming language (the code may be downloaded from the author's web site at www.ce.utexas.edu/prof/bhat/home.html).

The reader will note that assuming that the error terms are identical and independently distributed with a Gumbel distribution in Equation (5) leads to a mixed logit framework. The mathematical expressions for choice probabilities are simpler in the mixed logit than the mixed probit framework. However, we use the mixed probit framework here because the choice probabilities in the resulting model can be written using a MNP framework. When the number

of alternatives is small, the choice probabilities and the model parameters can be estimated accurately in the MNP framework (as we discuss in the next section), and can serve as the yard stick for evaluating the performance of alternative sequences for simulation estimation in the mixed probit framework.

5.2 The MNP framework

The model of Equation (5) can be estimated in another way by combining the random part of the coefficient vector with the error term, and explicitly recognizing the resulting correlations across alternatives. This formulation results in the traditional multinomial probit model (see Hausman and Wise, 1978). To explicate this formulation, define a $I \times S$ matrix $X_q = (x'_{q1}, x'_{q2}, ..., x'_{qI})'$ and define $I \times 1$ column vectors U_q and ε_q obtained by stacking the utility values and error terms, respectively, of the *I* alternatives. In vector form, Equation (5) may be written as:

$$U_q = X_q(\alpha + \delta_q) + \varepsilon_q = X_q\alpha + (X_q\delta_q + \varepsilon_q), \quad (8)$$

where α is the vector of mean values of the elements of β and δ_q is the vector of unobserved random deviations from the mean for individual *q*. Let $V_q = X_q\alpha$ and let δ_q be a realization from an underlying *S*-variate normal distribution with a mean vector of 0 and a variance matrix given by Σ_δ (in the current formulation, Σ_δ is a diagonal matrix with the *s*th diagonal element (*s* = 1, 2, …, *S*) being the variance in the *s*th element of the random parameter vector β). From Equation 8, the variance-covariance matrix of the utility vector U_q is $\Sigma_q = X_q\Sigma_\delta X'_q + ID_q$, where ID_q represents an identity matrix of size *I*. The expression for the probability of

individual q choosing alternative i can be reduced to an $(I-1)$ integral by defining the following $(I-1)\times I$ matrix operator (see Daganzo, 1979, pp. 43-44):

$$\Delta_i = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \\ i\text{–}1 \\ i \\ i\text{+}1 \\ \\ I\text{–}2 \\ I\text{–}1 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & \dots & i\text{–}1 & i & i\text{+}1 & \dots & I\text{–}1 & I \end{array} \\ \left[\begin{array}{cccccccccc} 1 & 0 & 0 & ... & 0 & -1 & 0 & ... & 0 & 0 \\ 0 & 1 & 0 & ... & 0 & -1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & -1 & 0 & ... & 0 & 0 \\ & & & \dots & & & & \dots & & \\ 0 & 0 & 0 & ... & 1 & -1 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & -1 & 1 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & -1 & 0 & ... & 0 & 0 \\ & & & \dots & & & & \dots & & \\ 0 & 0 & 0 & ... & 0 & -1 & 0 & ... & 1 & 0 \\ 0 & 0 & 0 & ... & 0 & -1 & 0 & ... & 0 & 1 \end{array}\right] \end{array} \tag{9}$$

Then, it can be shown that the required probability for alternative i is as follows:

$$P_{qi} = F_{I-1}(-W_{qi}, \Omega_{qi}), \text{ where } W_{qi} = \Delta_i V_q \text{ and } \Omega_{qi} = \Delta_i \Sigma_q \Delta'_i. \tag{10}$$

$F_{I-1}(-W_{qi}, \Omega_{qi})$ refers to the cumulative distribution function of the $(I-1)$-variate normal random variable with mean vector given by $-W_{qi}$ and covariance matrix given by Ω_{qi}. The expression in Equation (10) is, in general, quite difficult to compute. However, when the number of alternatives I is less than or equal to 4, one can compute the probabilities using well-developed and very accurate routines for evaluating the bivariate and trivariate cumulative normal distribution functions.

5.3 Experimental design

The discussion in the two previous sections on estimating a random-coefficients model with a probit kernel forms the basis for the experimental setting in this paper. Specifically, we

generate a sample of 2000 observations with 10 independent variables for three alternatives. The values for each of the 10 independent variables for the first two alternatives are drawn from a standard univariate normal distribution, while the corresponding values for each independent variable for the third alternative are drawn from a univariate normal distribution with mean 0.5 and standard deviation of 1. The coefficient to be applied to each independent variable for each observation is also drawn from a standard univariate normal distribution (i.e., $\beta_{qs} \sim N(0,1)$; q = 1,2,…, Q and s = 1,2,…, S). Finally, values of the error term ε_{qi} in Equation (5) are generated from a standard univariate normal distribution and the utility of each alternative is computed based on Equation (5). The alternative with the highest utility for each observation is then identified as the chosen alternative.

The data sample generated above is used to estimate the mean and standard deviations of the random coefficients using the mixed probit framework as well as the traditional MNP framework. The mixed probit framework requires the computation of a 10-dimensional integral for each observation and provides a good setting to evaluate the performance of the PMC sequence, the randomized standard Halton sequence, and the randomized scrambled Halton sequence. The traditional MNP framework, on the other hand, requires only the computation of a bivariate cumulative normal distribution, and provides the true parameter estimates from the sample.

An issue that needs to be resolved in the mixed probit framework is the number of quadrature points to use in the internal one-dimensional integration in $L_{qi}(\beta)$ (see Equation 6). This one-dimensional integration is a consequence of the probit kernel used for the error terms ε_{qi}. It is important that the one-dimensional integration be evaluated accurately so that the ability to recover the parameters of the random-coefficients can be solely attributed to the type of

sequence used to evaluate the outer 10-dimensional integral. In the current paper, we used 10-point Gauss-Hermite quadrature to evaluate the inner one-dimensional integral. This decision was based on a test of the accuracy of different numbers of quadrature points. Specifically we generated a data sample exactly as discussed earlier in this section, except that we retained fixed coefficients for β in Equation (5). The estimates for β can then be obtained accurately using a simple IID multinomial probit (MNP) model. The estimates can also be obtained using Equation (6), except that the outer integral is not present. The inner integral can be computed using different quadrature points to determine the number needed for accurate parameter estimation. In our analysis, we found that the parameter estimates from 10-point quadrature were almost exactly the same as the actual values from a MNP model (the mean percentage difference across all model parameters was 0.002 between the MNP and 10-point quadrature estimations; 8 of the 10 parameters had identical values, while the other two differed by 0.01%).

6. Computational Results and Comparative Performance

All the numerical estimation methods were implemented using the GAUSS matrix programming language. The log-likelihood function and the gradients of the function with respect to relevant parameters were coded.

The estimation of the true parameter values from the MNP estimation served as the benchmark to compare the performances of the pseudo-random Monte Carlo (PMC) method, the randomized standard Halton method, and the randomized scrambled Halton method. For the PMC estimations, we considered three different numbers of draws: 250, 500, and 1000. We did not go beyond 1000 draws because computation time starts to increase quite substantially beyond this many number of draws in high dimensions. For each number of draws, we estimated the

model 10 times using different random seeds to estimate simulation variance. For the Halton and scrambled Halton methods, we considered 50, 100, and 150 draws. Again, for each number of draws, we estimated the model 10 times by generating 10 randomized sequences as discussed in Section 4.

The performance evaluation of the alternative simulation techniques was based on four criteria: (a) ability to recover the true model parameters, (b) ability to estimate the overall log-likelihood function accurately, (c) ability to reproduce individual (*i.e.*, observation-specific) likelihood function values, and (d) ability to replicate the individual logarithm of the likelihood function. For each of these criteria, the evaluation of the proximity of estimated and true values was based on two performance measures: (a) root mean square error and (b) mean absolute percentage error. Further, for each criterion-performance measure combination, we computed three properties: (a) bias, or the difference between the mean of the relevant values across the 10 runs and the true values, (b) simulation variance, or the variance in the relevant parameters across the 10 runs, and (c) total error, or the difference between the estimated and true values across all runs.

Tables 1 through 4 present the computational results. In each table, the error measures decrease in magnitude with higher number of draws within each of the PMC, standard Halton, and scrambled Halton sequences. However, the decrease in error measures is much more rapid for the Halton sequences compared to the PMC sequence. For example, in Table 1, the RMSE for total error decreases from 0.114 to 0.076 as the number of PMC draws is increased from 250 to 1000. However, the RMSE decreases much more sharply from 0.191 to 0.064 as the number of scrambled Halton draws is increased from 50 to 150. A similar result can be noticed in the mean absolute percentage error measure. These results are a consequence of additional Halton

draws being strategically located to improve the coverage of the sampling domain, so that each additional draw of the Halton sequence contributes more to error reduction than does a random draw. The result is particularly noticeable in the bias of the model parameters. The reduction in bias is quite gradual with higher number of PMC draws; however, the bias decreases by about half or more for each additional 50 Halton draws. Between the standard and scrambled Halton sequences, the error measures in Table 1 are lower for each number of draws for the scrambled Halton sequence. This can be attributed to the more uniform coverage of the scrambled Halton sequence relative to the standard Halton sequence in high dimensions. Again, the difference in the two Halton sequences is particularly apparent in the bias measure. The standard Halton sequence maintains a reduction by about half for each additional 50 draws, while the scrambled Halton sequence shows a nonlinear reduction; an increase from 50 to 100 draws reduces bias by half, while the increase from 100 to 150 draws reduces the bias by more than a third.

Table 2 presents the results for the overall log-likelihood function value. The RMSE and MAPE performance measures show that all the different sequences do reasonably well in estimating the overall log-likelihood function value. However, the Halton sequences again exhibit higher rates of decrease in error than does the PMC sequence. A peculiar result in Table 2 is that the bias and total error for the MAPE performance measure is exactly identical for 50 scrambled Halton draws (both the bias and total error values are 0.173). This is because the log-likelihood value at convergence is overestimated in magnitude (relative to the actual value) by each of the 10 scrambled Halton runs.

Tables 3 and 4 provide the results for the individual likelihood and log-likelihood functions, respectively. The RMSE in these tables may appear smaller than the values for the overall log-likelihood function value in Table 2, but this is deceiving because the RMSE values

in Tables 3 and 4 are computed at the individual level, while those in Table 2 correspond to the aggregate log-likelihood value across all individuals. The MAPE provides a better perspective here, and indicates that the average percentage error is higher in the computation of individual likelihood values than in the overall likelihood function. This is to be expected because of the averaging out effect of simulation errors across individuals in the computation of the overall log-likelihood. Between the individual likelihood and log-likelihood values, the error in the latter is higher than in the former because of the logarithm transformation, which accentuates differences in the untransformed likelihood values.

To summarize, three important overall observations may be made from Tables 1 through 4. First, the standard Halton and scrambled Halton sequences provide more "bang for the buck" compared to the PMC sequence; much fewer Halton draws are necessary to provide the same level of accuracy as a higher number of PMC draws. Second, the scrambled Halton sequence is the most effective in high dimensions. In particular, about 100 scrambled draws appear to provide the same level of accuracy and precision as 150 standard Halton draws or 500 PMC draws, and 150 scrambled draws provide a higher level of accuracy and precision than 1000 PMC draws. Third, a comparison of the MAPE values across the tables shows that the individual and overall likelihood function values are more accurately estimated than the model parameter values. This suggests a rather flat likelihood function near the optimum; that is, closeness of likelihood function values to the true likelihood value does not immediately imply closeness in model parameters too.

7. Summary and Conclusions

The use of simulation techniques for estimating discrete choice and other econometric models has been increasing in recent years to accommodate more behavioral realism in decision

structures. Most of the simulation techniques used to date are based on pseudo-random draws or its variants such as importance sampling. Bhat (2001) proposed the use of a quasi-random Monte Carlo (QMC) method (specifically, the standard Halton method) for discrete choice models. However, there are two problems with the use of the standard Halton sequence. First, the Halton sequence's coverage of the integration domain deteriorates quite rapidly in high dimensions. Second, the standard Halton sequence is deterministic, and so does not permit the computation of statistical error bounds. This paper proposes the use of a scrambled Halton sequence to ensure good coverage of the integration domain even in high dimensions. In addition, it proposes the use of a randomization technique along with the Halton sequence to facilitate the estimation of simulation error.

The paper develops an evaluation framework for assessing the performance of the PMC sequence, the standard Halton sequence, and the scrambled Halton sequence. The results of the analysis indicate that the standard Halton sequence is not as good as the scrambled sequence in high dimensions. However, the standard Halton sequence still is to be preferred over the PMC sequence, given that it provides better accuracy with fewer draws. But the effectiveness of the standard Halton sequence is not as good as the results in Bhat (2001) and Train (1999). The results in these earlier papers suggest that 100 standard Halton draws provide about the same level of accuracy in model parameters as about 1000 PMC draws. But these results were in the context of a low level of dimensionality (1-5 dimensions). In the current paper, which considers 10 dimensions of integration, the results suggest that it take about 150 standard Halton draws to get the same comparable level of accuracy as 500 PMC draws. On the other hand, it takes only 100 scrambled Halton draws to achieve the same degree of accuracy and precision as 150

standard draws or 500 PMC draws, and 150 scrambled Halton draws provides a better degree of accuracy and precision than 1000 PMC draws.

This paper contributes to the existing small body of literature on the use of quasi-Monte Carlo sequences for econometric model estimation. It also generates several research issues that need careful investigation in the future. First, there is a need to evaluate alternative quasi-Monte Carlo sequences such as the Faure sequence and (*t,m,s*)-net sequences (see Niederreiter and Xing, 1998). Second, the method used here to randomize a QMC sequence is based on Tuffin (1996), but other methods have also been proposed (see Owen, 1997; Wang and Hickernell, 2000). A comparison of alternative randomization schemes would be helpful. Third, the application of variance reduction techniques to randomized quasi-Monte Carlo estimators is another important research area. Fourth, there is a need to extend the scope of the current analysis to include different orders of dimension (such as 10, 20, 30, *etc*.) to examine the impact of dimensionality on the performance of alternative sequences. Fifth, it would be useful to examine the effectiveness of quasi-Monte Carlo sequences within a Bayesian estimation framework for complex discrete choice models (see Brownstone, 2000 and Train, 2001 for reviews of Bayesian estimation for discrete choice models).

In closing, it is only fair to acknowledge that our knowledge of QMC sequences for econometric model estimation is still quite limited. Recent studies by Bhat (2001), Train (1999), and Hensher (1999) have investigated the performance of the standard Halton sequences for discrete choice models with low dimensionality (<5) of integration. The current paper has added to this literature by examining the performance of a scrambled and randomized Halton sequence in 10 dimensions. All these studies point quite clearly to the value of QMC sequences in model estimation. However, substantially more computational and empirical research into the use of

QMC sequences is needed in different settings (such as different patterns of correlation among exogenous variables, independent trials of performance using different simulated data sets, and different model structures) to draw more definitive conclusions. Notwithstanding this caveat, the current paper provides a first, but rather clear, indication that scrambled Halton sequences perform better than the standard Halton and PMC sequences when the dimensionality involved in model estimation is of the order of 10.

Acknowledgments

The author benefited from discussions and email exchanges with Bruno Tuffin during the early stages of this research. The author would also like to thank Lisa Weyant for her help in typesetting and formatting the text. Frank Koppelman, Kenneth Train, and an anonymous reviewer provided important suggestions for improvement on an earlier version of this paper.

References

Bhat, C.R., 1995. A heteroscedastic extreme value model of intercity travel mode choice, *Transportation Research*, 29B, 6, 471-483.

Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model, *Transportation Research*, 35B, 677-693.

Braaten, E. and G. Weller, 1979. An improved low-discrepancy sequence for multidimensional quasi-Monte Carlo integration, *Journal of Computational Physics*, 33, 249-258.

Brownstone, D., 2000. Discrete choice modeling for transportation, resource paper, presented at the International Association for Travel Behaviour Research Meeting, The Gold Coast, Australia, July.

Cranley, R. and T.N.L. Patterson, 1976. Randomization of number theoretic methods for multiple integrations, *Siam Journal of Numerical Analysis*, 13, 6, 904-914.

Daganzo, C.F., 1979. *Multinomial Probit: The Theory and its Applications to Demand Forecasting*, Academic Press, New York.

Fang, K.T. and Y. Wang, 1994. *Number-Theoretic Methods in Statistics*, Chapman & Hall, London.

Hajivassiliou, V.A. and P.A. Ruud, 1994. Classical estimation methods for LDV models using simulations, in R. Engle and D. McFadden, eds., *Handbook of Econometrics*, IV, 2383-2441, Elsevier, New York.

Hajivassiliou, V.A., McFadden, D., and P.A. Ruud, 1996. Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results, *Journal of Econometrics*, 72, 85-134.

Hausman, J. and D. Wise, 1978. A conditional probit model for qualitative choice: discrete decisions recognizing interdependence and heterogeneous preferences, *Econometrica*, 48, 2, 403-426.

Hellekalek, P., 1984. Regularities in the distribution of special sequences, *Journal of Number Theory*, 18, 41-55.

Hensher, D., 1999. The valuation of travel time savings for urban car drivers: Evaluating alternative model specifications, Technical paper, Institute of Transport Studies, University of Sydney.

Hickernell, F.J., and R-X Yue, 2000. The mean square discrepancy of scrambled (*T*, *S*)-sequences, *Siam Journal of Numerical Analysis*, 38, 4, 1089-1112.

Kocis, L. and W.J. Whiten, 1997. Computational investigations of low-discrepancy sequences, *ACM Transactions on Mathematical Software*, 23, 2, 266-294.

McFadden, D., 1989. A method of simulated moments for estimation of the multinomial probit without numerical integration, *Econometrica*, 57, 995-1026.

Morohosi, H. and M. Fushimi, 1998. A practical approach to the error estimation of quasi-Monte Carlo integration, in H. Niederreiter and J. Spanier, eds., *Proceedings of Monte Carlo and Quasi-Monte Carlo*, Springer-Verlag.

Niederreiter, H., 1992. Random number generation and quasi-Monte Carlo methods, 63, CBMSNSF Regional Conference Series in Applied Math, SIAM, Philadelphia, Pennsylvania.

Niederreiter, H., 1995. New developments in uniform pseudo-random number and vector generation, in H. Niederreiter and J-S. Shiue, eds., *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computin*g, 94-112, Springer, New York.

Niederreiter, H. and C. Xing, 1998. Nets, (*t*,*s*)-sequences, and algebraic geometry, in P. Hellekalek and G. Larcher, eds., Random and Quasi-Random Point Sets, *Lecture Notes in Statistics*, 138, 267-302, Springer-Verlag.

Okten, G., 1996. Error estimation for quasi-Monte Carlo methods, in H. Niederreiter, P. Hellekalek, G. Larcher and P. Zinterhof, eds., *Monte Carlo and Quasi-Monte Carlo Methods*, 351-368, Springer-Verlag.

Owen, A.B., 1995. Randomly permuted (t,m,s)-nets and (t,s)-sequences, in H. Niederreiter and J-S. Shiue, eds., *Monte Carlo Methods in Scientific Computin*g, 299-317, Springer, New York.

Owen, A.B., 1997. Scrambled net variance for integrals of smooth functions, *The Annals of Statistic*s, 25, 1541-562.

Owen, A.B., 1998. Latin supercube sampling for very high dimensional simulations, *ACM Transactions on Modeling and Computer Simulatio*n, 8, 71-102.

Shaw, J.E.H., 1988. A quasirandom approach to integration in Bayesian statistics, *The Annals of Statistics*, 16, 2, 895-914.

Spanier, J. and E.H. Maize, 1994. Quasi-random methods for estimating integrals using relatively small samples, *SIAM Revie*w, 36, 18-44.

Train, K., 1995. Simulation methods for probit and related models based on convenient error partitioning, technical paper, Department of Economics, University of California, Berkeley.

Train, K., 1999. Halton sequences for mixed logit, technical paper, Department of Economics, University of California, Berkeley.

Train, K., 2001. A comparison of hierarchical bayes and maximum simulated likelihood for mixed logit, technical paper, Department of Economics, University of California, Berkeley.

Tuffin, B., 1996. On the use of low discrepancy sequences in Monte Carlo methods, *Monte Carlo Methods and Applications*, 2, 295-320.

Wang, X. and F.J. Hickernell, 2000. Randomized Halton sequences, *Mathematical and Computer Modelling*, 32, 887-899.

Yue, R-X., and F.J. Hickernell, 2001. Integration and approximation based on scrambled sampling in arbitrary dimensions, *Journal of Complexity*, 17, 881-897.

Zaremba, S.K., 1968. The mathematical basis of Monte Carlo and quasi-Monte Carlo methods, *SIAM Review*, 10, 303-314.

Appendix

Permutations for Scrambled Halton Sequences

Prime *r*	Permutation of (0 1 2 ... *r*-1)
2	(0 1)
3	(0 2 1)
5	(0 3 1 4 2)
7	(0 4 2 6 1 5 3)
11	(0 5 8 2 10 3 6 1 9 7 4)
13	(0 6 10 2 8 4 12 1 9 5 11 3 7)
17	(0 8 13 3 11 5 16 1 10 7 14 4 12 2 15 6 9)
19	(0 9 14 3 17 6 11 1 15 7 12 4 18 8 2 16 10 5 13)
23	(0 11 17 4 20 7 13 2 22 9 15 5 18 1 14 10 21 6 16 3 19 8 12)
29	(0 15 7 24 11 20 2 27 9 18 4 22 13 26 5 16 10 23 1 19 28 6 14 17 3 25 12 8)

Source: Braaten and Waller (1979)

List of Figures

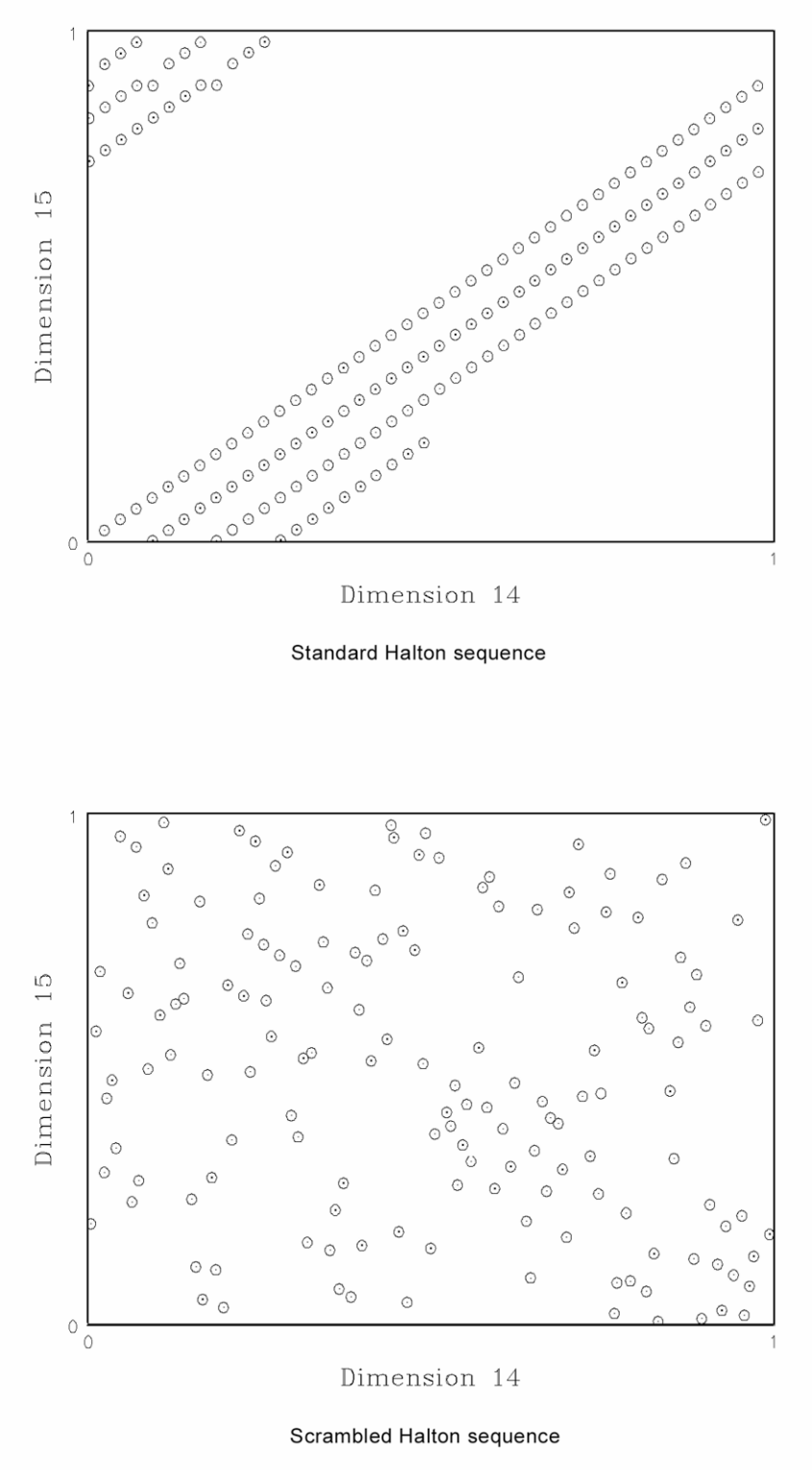

Figure 1. 150 Draws of Standard and Scrambled Halton Sequences

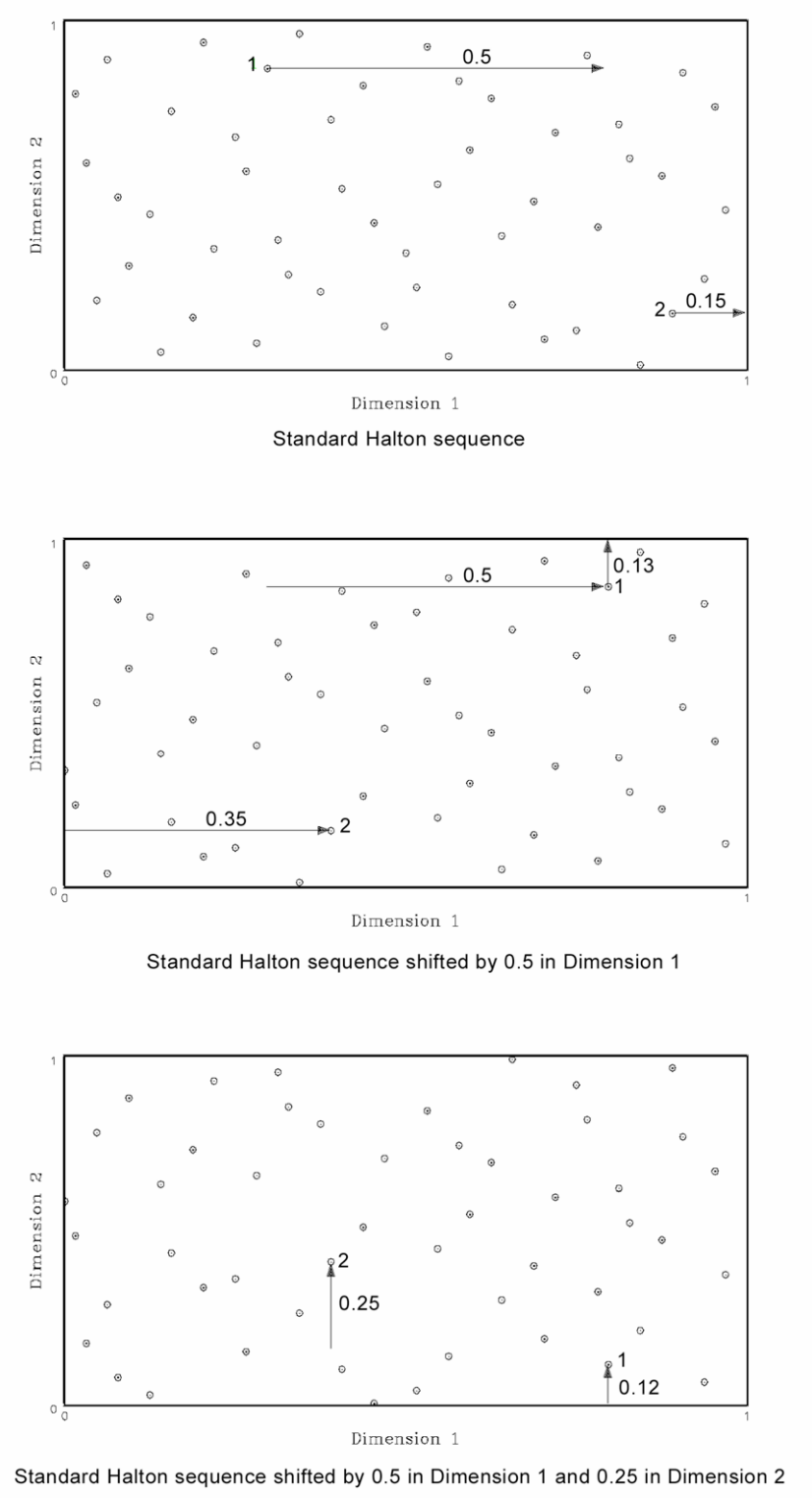

Figure 2. Shifting the Standard Halton Sequence

List of Tables

Table 1. Evaluation of Ability to Recover Model Parameters

Performance Measure	Estimator Property	Pseudo-Random Monte Carlo (PMC) Estimation			Standard Halton Estimation			Scrambled Halton Estimation		
		250	500	1000	50	100	150	50	100	150
RMSE[1]	Bias	0.054	0.041	0.035	0.204	0.102	0.068	0.166	0.084	0.026
	Simulation Variance	0.100	0.080	0.067	0.117	0.101	0.074	0.094	0.073	0.058
	Total Error	0.114	0.090	0.076	0.235	0.143	0.100	0.191	0.111	0.064
MAPE[2]	Bias	6.939	6.178	5.330	26.450	13.670	6.784	21.930	11.577	3.368
	Simulation Variance	13.030	10.160	8.259	15.939	14.202	10.402	13.097	10.577	8.034
	Total Error	14.315	11.934	9.461	28.227	19.618	12.097	24.280	12.612	8.622

[1] RMSE: Root Mean-Square Error

[2] MAPE: Mean Absolute Percentage Error

Table 2. Evaluation of Ability to Estimate Overall Log-Likelihood Function Value

Performance Measure	Estimator Property	Pseudo-Random Monte Carlo (PMC) Estimation			Standard Halton Estimation			Scrambled Halton Estimation		
		250	500	1000	50	100	150	50	100	150
RMSE[1]	Bias	1.300	1.253	1.191	3.132	1.870	1.196	2.323	1.366	0.599
	Simulation Variance	2.132	1.723	1.417	2.424	2.478	2.225	1.982	1.798	1.502
	Total Error	2.497	2.131	1.851	3.961	3.105	2.526	3.053	2.258	1.616
MAPE[2]	Bias	0.097	0.093	0.089	0.233	0.139	0.089	0.173	0.102	0.044
	Simulation Variance	0.139	0.110	0.086	0.145	0.156	0.129	0.137	0.123	0.075
	Total Error	0.161	0.127	0.109	0.269	0.178	0.161	0.173	0.136	0.079

[1] RMSE: Root Mean-Square Error

[2] MAPE: Mean Absolute Percentage Error

Table 3. Evaluation of Ability to Estimate Individual Likelihood Function

Performance Measure	Estimator Property	Pseudo-Random Monte Carlo (PMC) Estimation			Standard Halton Estimation			Scrambled Halton Estimation		
		250	500	1000	50	100	150	50	100	150
RMSE[1]	Bias	0.007	0.005	0.004	0.011	0.007	0.005	0.011	0.006	0.003
	Simulation Variance	0.020	0.014	0.010	0.027	0.018	0.014	0.025	0.015	0.008
	Total Error	0.021	0.015	0.011	0.030	0.020	0.015	0.028	0.016	0.008
MAPE[2]	Bias	1.397	0.985	0.719	2.394	1.390	1.01	2.588	1.273	0.624
	Simulation Variance	3.751	2.596	1.872	5.140	3.555	2.700	4.658	2.668	1.647
	Total Error	3.999	2.791	2.006	5.757	3.832	2.892	5.389	2.950	1.798

[1] RMSE: Root Mean-Square Error

[2] MAPE: Mean Absolute Percentage Error

Table 4. Evaluation of Ability to Estimate Individual Log-Likelihood Function

Performance Measure	Estimator Property	Pseudo-Random Monte Carlo (PMC) Estimation			Standard Halton Estimation			Scrambled Halton Estimation		
		250	500	1000	50	100	150	50	100	150
RMSE[1]	Bias	0.024	0.017	0.012	0.048	0.028	0.018	0.046	0.021	0.009
	Simulation Variance	0.064	0.043	0.032	0.089	0.065	0.057	0.085	0.048	0.026
	Total Error	0.069	0.046	0.034	0.101	0.071	0.060	0.097	0.052	0.028
MAPE[2]	Bias	2.874	2.061	1.658	5.809	3.051	2.149	5.595	2.207	1.201
	Simulation Variance	7.693	5.403	3.971	11.627	7.509	5.739	10.068	5.917	3.635
	Total Error	8.231	5.811	4.477	12.886	8.164	6.148	11.743	6.212	3.811

[1] RMSE: Root Mean-Square Error

[2] MAPE: Mean Absolute Percentage Error