UNIVERSITÄT HOHENHEIM

Hohenheimer Agrarökonomische Arbeitsberichte

Standard and Shuffled Halton Sequences in a Mixed Logit Model

Alexander Staus

Arbeitsbericht Nr. 17

Institut für Agrarpolitik und Landwirtschaftliche Marktlehre (420)
Universität Hohenheim, 70593 Stuttgart

Veröffentlichung des Institutes für
Agrarpolitik und Landwirtschaftliche Marktlehre der Universität Hohenheim

ISSN 1615-0473

Herausgeber: Institut für Agrarpolitik und Landwirtschaftliche Marktlehre
Universität Hohenheim (420)
70593 Stuttgart
Tel.: 0711/459-22599
Fax.: 0711/459-22601
e-mail: marktlehre@uni-hohenheim.de

Gesamtherstellung: Institut für Agrarpolitik und Landwirtschaftliche Marktlehre
Universität Hohenheim (420)
70593 Stuttgart

Standard and Shuffled Halton Sequences in a Mixed Logit Model

Alexander Staus*

Institute for Agricultural Policy and Agricultural Markets

University of Hohenheim

September 2008

* Alexander Staus is Research Associate at the University of Hohenheim (Institute for Agricultural Policy and Agricultural Markets).

Standard and Shuffled Halton Sequences in a Mixed Logit Model

Alexander Staus

Abstract

Modeling consumer choice in different areas has lead to an increase use of discrete choice models. Probit or Multinomial Logit Models are often the base of further empirical research of consumer choice. In some of these models the equations to solve have no closed-form expression. They include multi-dimensional integrals which can not be solved analytically. Simulation methods have been developed to approximate a solution for these integrals. This paper describes the Standard Halton sequence and a modification of it, the Shuffled Halton sequence. Both are simulation methods which can reduce computational effort compared to a random sequence. We compare the simulation methods in their coverage of the multi-dimensional area and in their estimation results using data of consumer choice on grocery store formats.

Keywords: simulation, mixed logit, halton sequence

JEL: C15, C25

Contents

1 Introduction **1**

2 Simulation **2**

3 Standard Halton Sequence **4**

4 Shuffled Halton Sequence **6**

5 Empirical application **11**

5.1 Data 11

5.2 Model 12

5.3 Performance 13

5.4 Importance of the coefficients 16

6 Conclusion **17**

References **18**

1 Introduction

In the last decade the use of discrete choice models has increased in modeling consumer choice in many areas like e.g. the choice of mode of transportation, the choice of energy supplier, the choice between organic or conventional food or brand-choice. Many of these discrete choice models are based on Probit or some type of Multinomial Logit Models (MNL). Some of these models have equations with no closed-form expression and include multi-dimensional integrals which can not be solved analytically. Therefore methods trying to approximate a solution for the multi-dimensional integral became more and more important. One approach is doing that by simulation. While researchers knew the theory of simulation methods, they were seldom used in practice. One reason for that was the need of high computational effort. With increasing computer speed this problem is now a minor one, but still calculations of hours or even days are not uncommon.
The probably most famous simulation method in econometrics is the Monte Carlo simulation, which was first based on pseudo-random numbers (pseudo-Monte Carlo). It can be called "pseudo", because every programmed random number generator generates the numbers not really randomly but rather by a code. As a consequence the numbers we get from this programmed routine are called "pseudo-random". Alternative simulation methods are the so-called quasi-random number sequences, which can provide a better coverage of the area of integration (quasi-Monte Carlo). These quasi-random numbers are even not programmed to appear randomly, but follow a specific predetermined method. One aim of using quasi-random numbers is to save computational time by using less draws. Train (2003) describes some methods of taking draws. First he introduces pseudo-random draws for e.g. a standard normal, a uniform or a truncated density and then he describes some variance reduction draws (quasi-random draws) like antihetics, systematic sampling and Halton sequences. The focus in this paper is on Halton sequences which were first introduced by Halton (1960). Train (2000) and Bhat (2001) show that Halton sequences provide better accuracy with fewer draws and less computational time than pseudo-random draws do. They both demonstrate that 100 Halton draws provide better accuracy than using

1000 pseudo-random draws. The use of Halton draws for higher dimensional integrals can lead to problems because of the correlation between the generated draws. Hess & Polak (2003b) showed some modification of the Halton sequence to remove the correlation between the draws, the so-called "shuffled" Halton sequence. The object of this paper is to give a short introduction into the shuffled method of a Halton sequence and to use and compare the different simulation methods (random, Halton and shuffled Halton), using a Mixed Multinomial Logit model.
The models are used with panel data on consumer choice of different grocery store formats (discounters, conventional supermarkets, small and large hypermarkets and specialized dealer shops). We estimate the choice of the grocery store format with random coefficients for the intercept and with random coefficients for the variables age, gender and net income.

The paper is organized as follows. Section 2 gives a small introduction how simulation methods work generally. Section 3 and 4 describe the Standard Halton and the Shuffled Halton sequence respectively. Section 5 explains the data, the model used for simulation, compares the different simulation methods and interprets the influence of the variables on the chosen grocery store formats. Section 6 concludes.

2 Simulation

In general a function of the following form has to be calculated:

$$P = \int S(\beta) f(\beta) d\beta \tag{1}$$

$f(\beta)$ is a density function and $S(\beta)$ is the actual function of interest. $S(\beta)$ can be e.g. a Mixed Multinomial Logit (MMNL) probability term where the random coefficients in the model follow the density $f(\beta)$.[1] In this case, the function P has no closed-form and cannot be calculated analytically, but it can be approximated by simulation.

[1]For these random coefficients e.g. the mean and the variance can be calculated.

Simulation is based on drawing from a density $f(\beta)$ and replacing a continuous average by a discrete average (Bhat, 2003). We can get this discrete average by taking randomly points. The standard routine for simulation is (compare Train, 2003)

1. Draw a value of β^r from its density function $f(\beta)$ where r specifies the rth draw with $r = 1$ as the first draw and $r = R$ as the last draw. A standard uniform draw in the 0-1 interval which is the basis of these draws can be transformed into the assumed density function $f(\beta)$.

2. Calculate the function of interest $S(\beta^r)$.

3. Repeat this process for R (= number of draws) times and average the results, accordingly we get an estimate for equation (1).

With that procedure the function P in equation (1) is approximated by

$$\hat{P} = \frac{1}{R} \sum_{r=1}^{R} S(\beta^r). \tag{2}$$

This is just the discrete average of R randomly taken points.
In case of the MMNL model for panel data with random coefficients over individuals, $S(\beta)$ is the likelihood function for one individual i:

$$S_i(\beta_i) = \prod_{t=1}^{T} L_{ij(i,t)t}(\beta_i) \tag{3}$$

This function is the Multinomial Logit (MNL) probability. Taking the integral over the density of random terms, if any, we get the MMNL with:

$$L_{ijt}(\beta_i) = \frac{\exp(\beta_i' x_{ijt})}{\sum_{m=1}^{J} \exp(\beta_i' x_{imt})} \tag{4}$$

The $[K \times 1]$ vector x_{ijt} includes the K explanatory variables from individual i for alternative j at choice situation t. β_i is the coefficient vector to be estimated, including fixed or random coefficients. The distribution of the random coefficient vector can be normal, lognormal, uniform, triangular or of any other form. In case of the normal density function the mean and variance can be estimated. $j(i,t)$ in equation (3) denotes the alternative which individual i choose in time period t, so $S_i(\beta_i)$

is the conditional probability of individual i's observed sequence of choices and $\hat{P}_i = \frac{1}{R}\sum_{r=1}^{R} S(\beta_i^r)$ is the simulated unconditional probability of person i's sequence of choices. For independently draws from density f, the simulated probability is unbiased and consistent for the true probability (Sandor & Train, 2004). The variance decreases as R increases.

The simulated log-likelihood function over all individuals is:

$$SLL = \sum_{i=1}^{N} \ln(\hat{P}_i) \tag{5}$$

This log transformation of equation (2) is non-linear, therefore the estimator based on maximizing SLL in equation (5) is biased. The bias decreases if the number of draws (R) rises faster than the square root of the number of observations, so the estimator is consistent and equivalent to the maximum likelihood estimator.

3 Standard Halton Sequence

Halton sequences are one of the most popular quasi-random types (Hess et al., 2003) and were first introduced by Halton (1960). A Halton sequence is one way to take draws from a density. To understand how the sequence is generated, we go through an example (compare Train, 2003):

1. Take a prime number, e.g. 3.

2. Divide the unit interval, which is between 0 and 1, into 3 (=number of the prime) equal parts. We get $\frac{1}{3}$ and $\frac{2}{3}$. These are the first two draws.

3. Divide each of the three parts again into three equal shares and add the first part of the share to the breakpoints from the first draws. $\frac{1}{9}$, $\frac{4}{9}$ and $\frac{7}{9}$ are the next three draws. Then add the second part of the share to the same breakpoints (we get $\frac{2}{9}$, $\frac{5}{9}$ and $\frac{8}{9}$).

4. Divide each of the nine parts into thirds and follow routine in point 3. We get the following sequence: $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \frac{1}{27}, \frac{10}{27}, \frac{19}{27}, \frac{4}{27}, \frac{13}{27}, \frac{22}{27}, \frac{7}{27}, \frac{16}{27}, \frac{25}{27}$.

In general to get a Halton sequence we can follow Braaten & Weller (1979). For prime p we can write any integer g $(g = 1, ..., G)$ in terms of the base p:

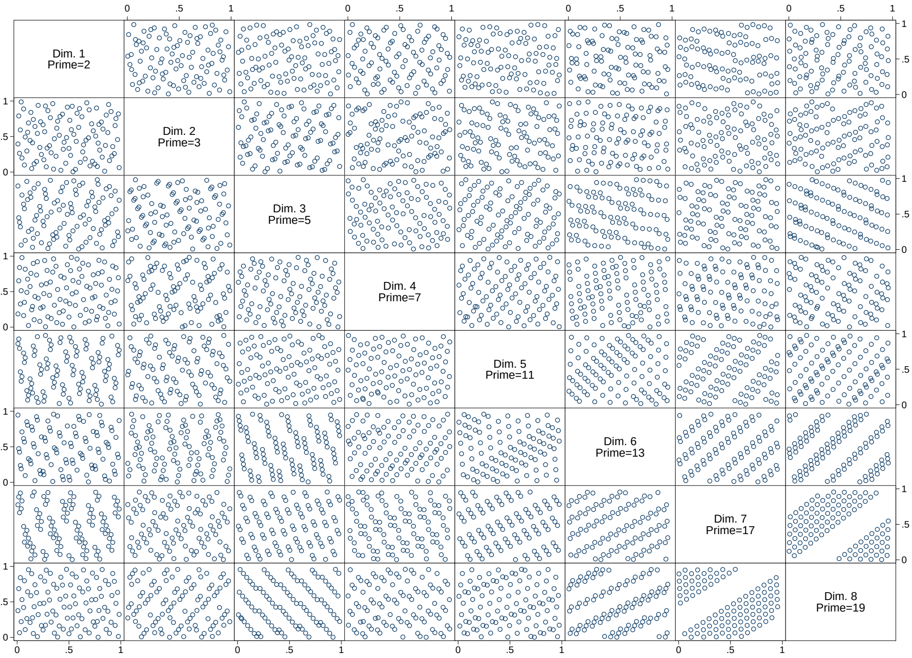
$$g = e_j p^j + \cdots + e_1 p + e_0 p^0, \quad \text{where} \quad 0 \leq e_i \leq p - 1. \tag{6}$$

So g can be represented in digitized form by the integer string $e_j \cdots e_1 e_0$. We take now the radical inverse of g to the base p by reflecting through the radical point $(= 0.e_0 e_1 \cdots e_j)$ and get the Halton sequence for prime p:

$$\varphi_p(g) = e_0 p^{-1} + e_1 p^{-2} + \cdots + e_j p^{j+1} \tag{7}$$

As an illustration we take the prime number 3 and the integer 7. We can express the integer 7 in base 3 as: $7 = 2 \times 3^1 + 1 \times 3^0$. The important parts of g are $e_1 = 2$ and $e_0 = 1$, so the digitized form is 21, the radical inverse of it is 0.12 and the seventh draw of the Halton sequence can be written as $\varphi_3(7) = 1 \times 3^{-1} + 2 \times 3^{-2} = \frac{5}{9}$.
Halton sequences are structured that way, that in one sequence they fill in the gaps of the previous sequence. This property leads to negatively correlated draws and therefore it reduces the variance in the simulated log-likelihood function. Furthermore this characteristic of the Halton sequence ensures a better coverage of the multi-dimensional area of integration compared to random draws. With the better coverage less draws need to be taken than with pseudo-random numbers and this reduces computational time. For discrete choice models Train (2000) and Bhat (2001) show that Halton sequences provide better accuracy with 100 Halton draws than with 1000 pseudo-random draws.
While for lower-dimensional integration the Halton sequence covers the 0-1 multi-dimensional space quite good, for higher-dimensional integrals the Halton sequences can be highly correlated. The consequence is an unequal coverage of the multi-dimensional area of integration and poor estimation results. Figure 1 shows a scatter-plot matrix for different two-dimensional Halton sequences for the first 8 primes.
It can be seen that the correlation increases while moving to the south-east, to higher dimensional Halton sequences. In figure 2 the correlation for dimensions 9 to 16 becomes very obvious.
Table 6 at the end of the paper shows a correlation matrix of the Standard Halton sequence of 100 draws with primes 5 to 71. The correlations of primes higher than

Figure 1: Scatter-plot matrix of two-dimensional Halton sequences with 100 points from dimension 1 (prime 2) to dimension 8 (prime 19)



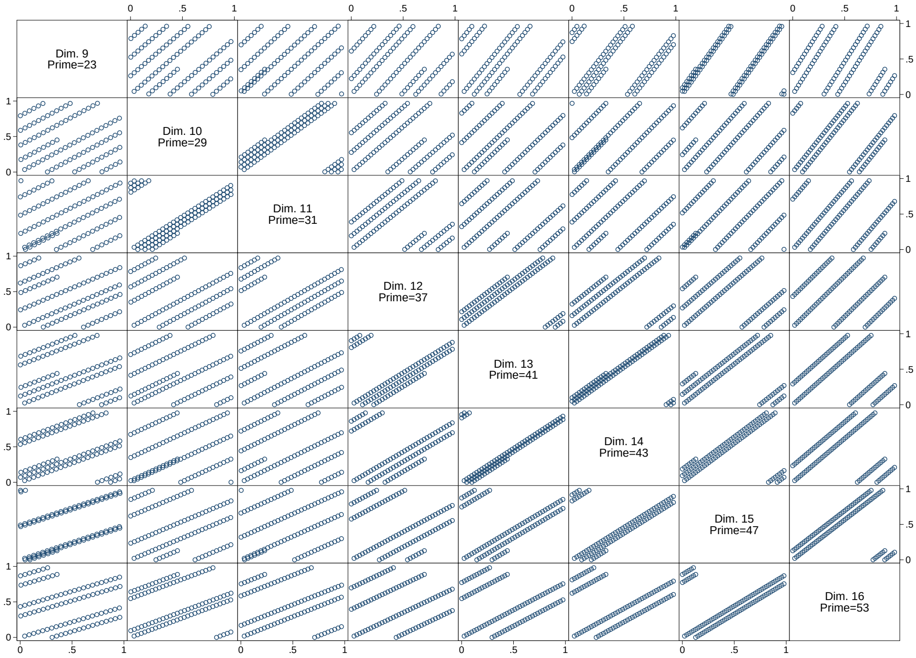
37 with its next prime is for all listed primes higher than 0.30 in absolute value, the correlation of prime 67 and prime 71 is about 0.756. This is a quite high relationship and simulation using Halton sequences with these primes should be carried out with caution.

The correlation between the prime numbers of higher dimensions is caused by the identical generating behaviour of the different sequences. For a ratio of two primes close to an integer value (especially 1) the correlation between these primes increases. The length of cycles used are then very similar (Hess & Polak, 2003a). This is actually the reason why primes have to be taken for the sequence. For nonprimes the sequence can be an exact multiply of each other.

4 Shuffled Halton Sequence

Since high correlation between the prime numbers leads to an unequal coverage of the multi-dimensional area of integration and therefore to poor estimation results,

Figure 2: Scatter-plot matrix of two-dimensional Halton sequences with 100 points from dimension 9 (prime 23) to dimension 16 (prime 53)



especially for higher primes, the Standard Halton sequence is not recommended to use for an integral with more than six or seven dimensions. There have been different variations of the Standard Halton sequence to avoid this problem like e.g. Randomized Halton draws or Scrambled Halton draws.[2] Bhat (2003) shows that Scrambled Halton sequences perform better than Standard Halton sequences. Hess & Polak (2003b) say that even if the correlation between different primes is lower in general for the scrambled sequence, it still exhibits a very high correlation for some primes and in this cases even a pseudo-random number sequence can perform better. Besides this problem the scrambled sequence is hard to calculate and only for the first sixteen primes the code to generate a Scrambled Halton sequence can be found and downloaded (from Bhat).

For that reason Hess & Polak (2003b) present another variation, the Shuffled Halton sequence. The idea is to use randomly shuffled sequences of the one-dimensional

[2]For an introduction to these Halton variations see Train (2003).

Standard Halton sequence. Using a pseudo-random generator to shuffle the one-dimensional Standard Halton sequence will not influence the good coverage of the original one-dimensional sequence, since the order of draws is not important to the coverage. With different permutations of the one-dimensional sequences we get different multi-dimensional draws. The order of the Standard Halton sequence gets randomized and with a sequence of length R there are $R!$ different possible permutations. This is even high with a low length of R and therefore the probability of using the same random permutation to two different sequences is very close to zero. With this process new multi-dimensional sequences will always differ because of the use of a pseudo-random generator. Figure 3 shows a scatter-plot matrix for different two-dimensional Halton sequences for dimensions 9 to 16. As one can see for this generated Shuffled Halton sequence the correlation is far less compared to the correlation of the Standard Halton sequence (see figure 2).

Table 7 at the end of the paper shows a correlation matrix of one Shuffled Halton sequence of 100 draws with primes 5 to 71. This is just one generated shuffled sequence out of $9.33 * 10^{157}$ possible sequences per dimension, so the correlation is not fixed on the values in Table 7.

Nearly no correlation in the shuffled sequence is higher than 0.3 in absolute value except of two (prime 17 - prime 37 and prime 13 - prime 59). Hess & Polak (2003b) computed the correlation for primes 43 and 47 over 500 runs with 100 draws. The mean absolute correlation is 0.0876 (variance of 0.0045) compared to a mean absolute correlation of 0.1075 (variance of 0.0236) for the Standard Halton sequence. This is very similar to the correlation of pseudo-random number sequences. Figure 4 shows a scatter-plot of four runs with primes 67 and 71 and Table 4 the according correlations. We can conclude that the correlation can be significantly reduced by using a shuffled version of the Halton sequence. This leads to a better coverage of the multi-dimensional area of integration and to better estimation results even with high dimensions.

Figure 3: Scatter-plot matrix of two-dimensional Shuffled Halton sequences with 100 points from dimension 9 (prime 23) to dimension 16 (prime 53)

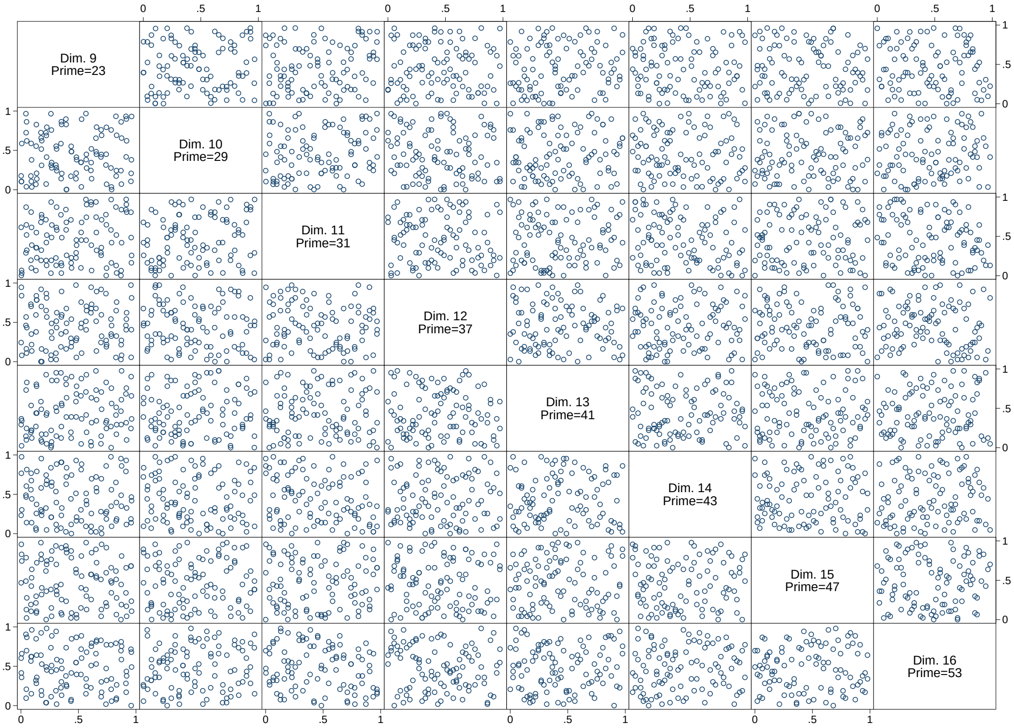
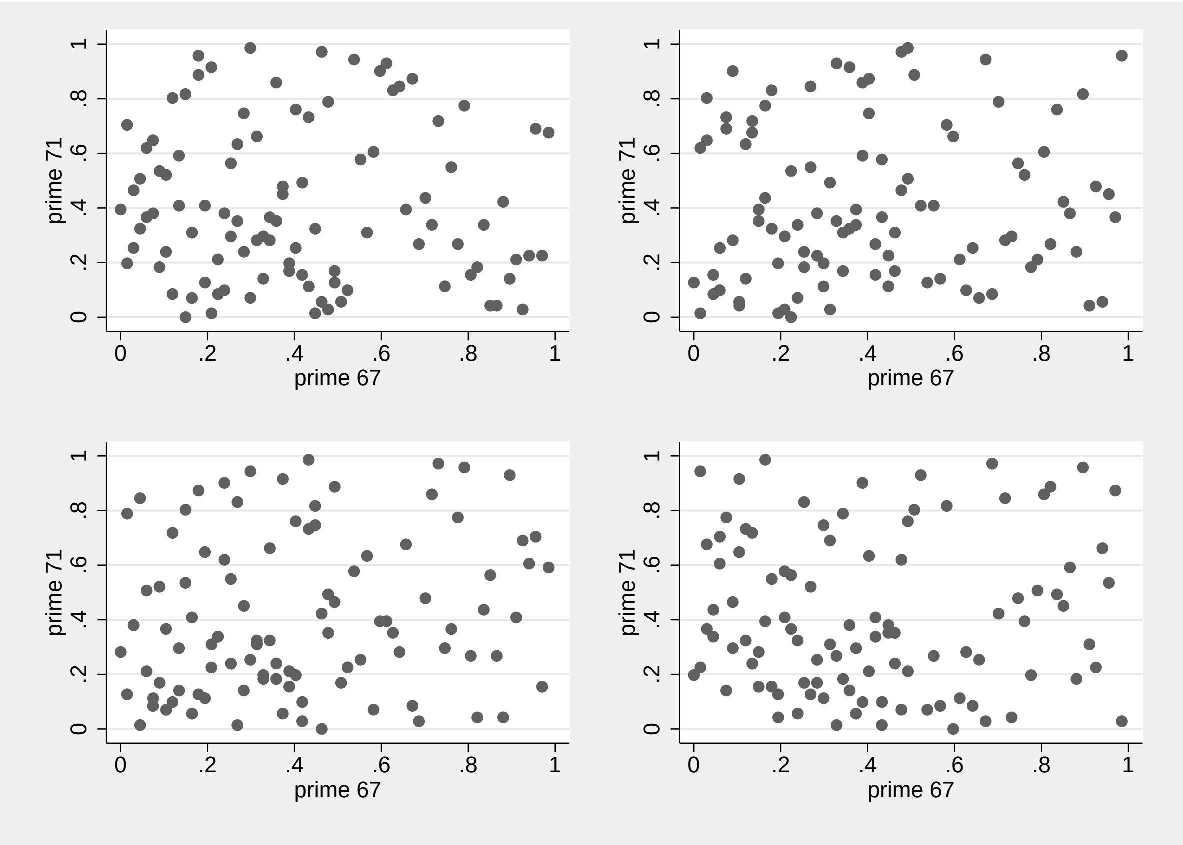


Table 1: Correlation between Shuffled Halton sequences of 100 draws of four runs with primes 67 and 71

		prime 67			
		1st run	2nd run	3rd run	4th run
prime 71	1st run	-0.0553	-0.0584	0.0804	0.0841
	2nd run	-0.0721	0.0636	-0.0236	0.1675
	3rd run	-0.0833	-0.1038	0.1531	-0.1326
	4th run	-0.0995	-0.0833	-0.0765	0.0210

Figure 4: Correlations of four runs of two-dimensional Shuffled Halton sequences with 100 points with primes 67 and 71



5 Empirical application

5.1 Data

The different simulation methods are used on a household panel data set which was provided by the GfK Group. The original data set contains a representative sample of 23,466 households who reported their purchases on various consumer goods between 1st January 2002 and 30th June 2006 in Germany. Because of the amount of data and computational speed limits a random subset of 140 households with purchases on fruits and vegetables between the 1st January 2006 and the 30th June 2006 is used. The variables included in the simulation process are the chosen grocery store format (like discounters, conventional supermarkets, small and large hypermarkets and a specialized dealer shop), age, gender and net income. The total number of observations in the data set is 4,288. In the six months of observation every household visited a grocery store 30.63 times. The summary statistics for the variables are presented in Table 2.

Table 2: Summary statistics of the used variables

Variable	Description	Mean	St.dev.
gender	gender of the purchasing person Male = 0, Female = 1	0.72	0.45
age	age of the purchasing person 11 categories $7.55 \approx 52$ years	7.55	2.98
net income	net income of the purchasing person $6.93 \approx 1115$ Euro	6.93	2.70

Table 3 shows the grocery store format shares of the households in the sample.

Table 3: grocery store format choices

Discount stores	49.63 %
Conventional supermarkets	15.88 %
Small hypermarkets	12.29 %
Large hypermarkets	14.95 %
Specialized shop	7.25 %

5.2 Model

A Mixed Multinomial Logit model (MMNL) is used to analyze the data. Contrary to a Multinomial Logit model (MNL), the MMNL allows for random taste variation across decision-makers. This implies that individuals with the same observed characteristics do not need to have the same "tastes". Our model allows correlation in unobserved factors over time which takes the nature of the panel data into account. The strong assumption of "independence from irrelevant alternatives" (IIA) in the MNL model does not apply in the MMNL model (Revelt & Train, 1998). The MMNL is highly flexible and can approximate any random utility model (McFadden & Train, 2000).

In this paper the model described in equations (2)-(5) is used with the three explanatory variables (age, gender and net income) assumed to be random in model M1. Dummy variables for every choice possibility were generated to include an alternative specific random intercept. Besides the model described in section 2 we assume that all coefficients can differ between all choice possibilities, that means that $\beta_{ij} \neq \beta_{ik} \quad \forall \quad j \neq k$, where j, k can be discount stores, conventional supermarkets, small hypermarkets, large hypermarkets or specialized shops. Furthermore we allow for correlation between the coefficients according to the particular variable or the intercept.

It can be assumed that the coefficients for all variables can differ between the individuals and don't need to be restricted to have the same sign for the whole population. Since they can be either positive or negative, a Normal distribution is used for all coefficients. To identify the model we use the discount store as the base category, so all coefficients for that choice are normalized to zero. This leads all coefficients to interpret relative to the discount store.

With one choice normalized to zero, three variables, a constant and five choice possibilities we have at all 16 different coefficients. Among with Hess et al. (2006) this is likely the highest simulated multi-dimensional integral compared with different types of Halton sequences in a MMNL.

We estimate the model using the three different simulation methods: Random sequence, Standard Halton sequence and Shuffled Halton sequence. Since the first draws of the Halton sequence are highly correlated, we follow Train (2000) and drop the first 50 draws. For the Halton and the shuffled Halton sequence we use the first 16 prime numbers. To compare the estimation results we use draws of 50, 100, 200 and 500 for every simulation method (M1). Using the routine of Hole (2007) within Stata for the Standard Halton sequence and modifications of it for the random sequence and the Shuffled Halton sequence, we can estimate the Model with the different simulation methods.

For an additional comparison we add further models. We let the random coefficients vary from one to four (models M2-M5) with simulation draws of 50, 100, 200, 500 and 1000. While using the random sequence as a basis, we compare the results of the Halton and the shuffled Halton sequence in their performance. For the Halton and the shuffled Halton sequence we use first the initial and then higher primes. Our aim is to show that the Halton sequence with higher primes performs worse than the shuffled Halton sequence with higher primes due to a worse coverage of the multi-dimensional area.

5.3 Performance

Tables 8-10 show the estimation results of our first model M1. “Sup” to “Special” are the shortcuts for the alternative specific coefficients for the grocery store formats. The other variables are the respectively grocery store specific differences of gender, age and net income in relation to the discount store. Table 8 illustrates the results using a random sequence. Compared to the other two tables (Halton and shuffled Halton) it can be seen that the log likelihood varies less and becomes quite stable after 200 draws. This is not the case for the other two simulation methods contrary to our expectations. For nearly all coefficients, except e.g. "Age Large hyp", the coefficients get not robustly estimated. The differences between the used draws inside of one simulation method are mostly higher than 0.1. The differences between the models are even higher. Therefore we can't conclude that one model is superior to another one. The differences of the coefficients between the random sequence and the Halton sequence are much higher than the differences between these two models compared to the shuffled Halton sequence.

To find an explanation we analyze the other 4 models (M2-M5) with 1 to 4 random coefficients and varied prime numbers.

Model M2 (see Table 11) includes only one random coefficient for the alternative specific variable "Supermarket". For the standard and shuffled Halton sequences we use for each two different primes, 2 and 11, for a better comparison. The random model gets fairly stable with 500 draws for the log likelihood and the coefficient, while the Halton sequence with the prime number 2 is already quite stable with 50 draws. Using the Halton sequence with the prime 11, we get good results with 100 draws. With the shuffled sequence models with prime 2 and and also with prime 11 we get stable results with 200 draws for the log likelihood and with 500 draws for the coefficient. The deviation for the coefficient for less draws is quite high. This result is very unexpected.

What happens if we use two random coefficients (model M3), one alternative specific variable for the "Supermarket" and one for "Small hypermarkets" (see Table 12). We get similar results as for model M2 with one random coefficient. Most sequences lead to fair results with 200 draws, the performance of the random model is a bit worse than in M2 and the shuffled sequences lead again to bad coefficient results up to 200 draws.

For the higher dimension models M4 and M5 with three respectively four random coefficcients the results do not change very much. The estimation results for the Halton sequence with higher primes (11, 13 and 17 for model M4 and 11, 13, 17 and 19 for model M5) is worse compared with the models using the first primes.

Table 4 lists the highest correlation between the draws for the different sequences. For the Halton sequence with high dimension primes the highest correlation is always between the last two used primes, 17 and 19, which are used in M5. So the poor performance of model M5 is not surprising. The differences in the shuffled Halton sequences are at least stable at around 500 draws and have the highest deviation of the estimated coefficients of all models. This is rather strange, since we expected to get a better performance with the shuffled sequence compared to the standard Halton sequence. A deeper view onto the sequences and onto the correlations between the different draws exhibits some interesting relationships. Even if the mean absolute correlation is less for some shuffled sequences compared to some standard Halton sequences, it's not guaranteed that the model performs better. It's more likely that e.g. if just one single correlation between two sequences is higher in the shuffled model than in the standard Halton model, even if the mean absolute correlation is smaller, the model with the single higher correlation will perform worse. That's the case for some values of the shuffled and standard Halton sequences and it even holds for some random sequences. For proving that statement further analysis has to be carry out. Table 5 shows the mean absolute correlation of the different sequences.

Table 4: Highest correlation between draws for the diffferent sequences

	Random	**Halton**	**Shuffle**	**Halton**	**Shuffle**
		primes = 2,3,5,7		**primes = 11,13,17,19**	
		(primes)	(primes)	(primes)	(primes)
50	-30.58	-5.29	-9.58	-51.96	18.98
		(5-7)	(2-3)	(17-19)	(11-13)
100	27.14	2.57	11.91	-23.54	-17,92
		(3-5)	(2-5)	(17-19)	(13-17)
200	-15.42	1.27	12.62	-10.60	-8.51
		(5-7)	(2-7)	(17-19)	(11-17)
500	-6.68	-0.49	9.40	-2.26	-6.09
		(2-3)	(2-7)	(11-17)	(11-17)
1000	4.45	0.28	-3.59	-1.86	3.45
		(3-7)	(2-3)	(17-19)	(17-19)

The standard Halton sequences do very well for small primes and also for high

Table 5: Mean absolute correlation between draws for the diffferent sequences

	Halton	**Shuffle**	**Halton**	**Shuffle**
	primes = 2,3,5,7		**primes = 11,13,17,19**	
50	1.86	5.39	16.22	13.83
100	0.92	5.09	10.43	10.20
200	0.70	6.00	2.21	5.54
500	0.23	3.14	1.06	2.50
1000	0.12	2.10	0.86	1.42

primes with more than 200 draws. In our case the shuffled sequences outperform only in two cases the standard Halton sequences (high primes with 50 and 100 draws), but in these two cases the shuffled models perform worse. In general for our models we need at least 200 or better 500 draws for the different kinds of the Halton sequences (shuffled and standard). And with this amount of draws these models do not outperform the random sequence models.

5.4 Importance of the coefficients

With the discount store as the base category all the estimated coefficients have to be interpreted with respect to the discount store. Since in all our models the direction of the estimators are the same, we take a look at Table 9 for interpretation. The intercepts have all a negative sign, so most people prefer the discount store for fruits and vegetables, followed by the large hypermarket, the supermarket, the small hypermarket and the specialized dealer shop. The gender variable uncovers that women prefer small hypermarkets most, followed by the specialized dealer shop, the supermarket and the large hypermarket. According to the results, women don't fancy discount stores very much. The age variable shows that younger people prefer the discount store and older peoples preference is the specialized shop and the supermarket. The influence of net income to the chosen store format is quite consistent with our expectations. People with higher income prefer the specialized dealer shop most, followed by a small hypermarket, but they prefer a discount store compared to a large hypermarket.

6 Conclusion

The increasing use of discrete choice models with multi-dimensional integrals which can not be solved analytically requires simulation methods to approximate a solution for these integrals. Simulation with generated pseudo-random numbers are very common, but there are other simulation procedures to get better simulation results in a shorter time. We introduced the Halton sequence and a extended version, the shuffled Halton sequence. We compared these simulation methods (random, Halton, shuffled Halton), using a Mixed Multinomial Logit Model with 1 up to 16 random coefficients. The model uses panel data about purchases on fruits and vegetables of 140 households between the 1st January 2006 and 30th June 2006 in Germany with a total of 4,288 observations. The estimation results are contrary to our expectations that the Halton sequence needs less draws to get stable results compared to the random sequence. And already with using primes of 11 or 13 the results are not satisfactory. The results are even more confusing since the shuffled sequence leads to quite large differences in the estimation of the coefficients for less than 200 draws. A detailed view on the generated sequences, especially on the correlations between the different sequences within a simulation method, shows a possible reason for that behaviour. Not the mean absolute correlation is the driving force to get better estimation results, but just one single correlation between two sequences within a simulation method. This is not unreasonable since that can lead to an unequal coverage of the multi-dimensional area of integration and therefore to poor estimation results. By using simulation methods for at least two-dimensional integrals, we propose to inspect the correlation of all generated sequences. Further research is required to verify that conclusion.

References

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

Bhat, C. R. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled halton sequences. Transportation Research Part B, 37, 837–855.

Braaten, E. & Weller, G. (1979). An improved low-discrepancy sequence for multidimensional quasi-monte carlo integration. Journal of Computational Physics, 33(2), 249–258.

Halton, J. H. (1960). On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numerische Mathematik, 2, 84–90.

Hess, S. & Polak, J. (2003a). An alternative method to the scrambled halton sequence for removing correlation between standard halton sequences in high dimensions. ERSA conference papers ersa03p406, European Regional Science Association. available at http://ideas.repec.org/p/wiw/wiwrsa/ersa03p406.html.

Hess, S. & Polak, J. (2003b). The shuffled halton sequence. CTS Working paper, Centre for Transport Studies, Imperial College London.

Hess, S., Polak, J. W., & Daly, A. (2003). On the performance of the shuffled halton sequence in the estimation of discrete choice models. European Transport Conference 2003.

Hess, S., Train, K. E., & Polak, J. W. (2006). On the use of a modified latin hypercube sampling (mlhs) method in the estimation of a mixed logit model for vehicle choice. Transportation Research Part B, 40, 147–163.

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. Stata Journal, 7(3), 388–401.

McFadden, D. & Train, K. (2000). Mixed mnl models for discrete response. Journal of Applied Econometrics, 15, 447–470.

Revelt, D. & Train, K. (1998). Mixed logit with repeated choices: Households' choices of appliance efficiency level. The Review of Economics and Statistics, 80(4), 647–657.

Sandor, Z. & Train, K. (2004). Quasi-random simulation of discrete choice models. Transportation Research Part B, 38, 313–327.

Train, K. (2000). Halton sequences for mixed logit. Economics Working Papers E00-278, University of California at Berkeley.

Train, K. E. (2003). Discrete Choice Methods with Simulation. Cambridge University Press.

Table 6: Correlation between Standard Halton sequences of 100 draws with primes 5 to 71

	p 5	p 7	p 11	p 13	p 17	p 19	p 23	p 29	p 31	p 37	p 41	p 43	p 47	p 53	p 59	p 61	p 67
p 7	-0.043																
p 11	-0.038	0.010															
p 13	0.001	-0.002	0.009														
p 17	0.010	-0.030	-0.017	-0.043													
p 19	0.012	-0.030	-0.030	-0.005	-0.131												
p 23	-0.015	0.017	0.025	0.010	-0.018	-0.068											
p 29	-0.058	-0.030	0.016	-0.070	0.006	0.003	-0.086										
p 31	-0.072	-0.011	-0.058	0.011	-0.090	0.043	0.121	0.405									
p 37	0.021	-0.060	-0.030	0.037	0.051	0.270	-0.058	-0.137	-0.030								
p 41	-0.034	0.023	-0.048	0.034	0.002	0.036	-0.044	0.012	-0.120	0.343							
p 43	-0.041	0.047	0.094	-0.049	0.074	-0.063	0.120	0.209	-0.046	0.136	0.669						
p 47	0.003	0.069	-0.036	-0.011	-0.072	0.114	0.392	0.008	0.205	-0.053	0.227	0.436					
p 53	0.042	0.048	0.073	0.136	0.189	0.043	-0.024	0.099	0.014	-0.032	-0.022	0.062	0.374				
p 59	0.016	0.005	-0.000	-0.023	0.004	0.136	-0.067	0.327	0.250	-0.027	-0.109	-0.076	0.121	0.599			
p 61	0.008	0.011	-0.016	0.006	0.004	0.048	-0.035	0.218	0.362	0.007	-0.111	-0.098	0.056	0.477	0.860		
p 67	-0.016	-0.030	0.093	0.073	0.188	-0.048	0.191	0.003	0.153	0.193	-0.037	-0.086	-0.067	0.180	0.497	0.616	
p 71	0.030	0.015	-0.022	-0.025	0.077	0.023	0.194	-0.046	0.020	0.380	0.072	-0.019	-0.089	0.046	0.314	0.417	0.756

Table 7: Correlation between Shuffled Halton sequences of 100 draws with primes 5 to 71

	p 5	p 7	p 11	p 13	p 17	p 19	p 23	p 29	p 31	p 37	p 41	p 43	p 47	p 53	p 59	p 61	p 67
p 7	0.148																
p 11	-0.095	-0.208															
p 13	-0.009	0.097	-0.064														
p 17	-0.054	-0.040	0.103	0.018													
p 19	-0.088	-0.275	0.073	-0.011	0.015												
p 23	-0.239	-0.076	0.044	0.074	0.037	-0.044											
p 29	-0.029	-0.115	-0.095	0.051	0.058	0.208	-0.040										
p 31	0.029	-0.148	0.052	-0.045	0.026	-0.011	-0.162	0.011									
p 37	-0.004	-0.073	-0.040	-0.101	-0.319	0.080	-0.217	-0.128	-0.033								
p 41	0.016	-0.096	0.094	0.065	-0.185	-0.049	0.184	-0.094	-0.187	0.073							
p 43	-0.062	-0.079	0.078	-0.007	0.050	-0.044	0.064	0.006	0.122	-0.114	0.232						
p 47	-0.084	-0.048	0.254	-0.009	-0.125	0.098	0.011	0.035	-0.093	0.022	0.068	-0.096					
p 53	-0.023	-0.065	0.080	0.054	-0.020	0.174	0.313	-0.123	-0.153	0.054	0.049	0.055	-0.073				
p 59	0.048	0.031	0.072	0.303	0.146	-0.031	-0.039	-0.149	0.123	-0.042	0.048	0.070	-0.151	-0.031			
p 61	-0.237	-0.038	0.052	0.021	0.148	-0.120	-0.084	-0.047	0.112	-0.152	-0.127	-0.126	-0.062	-0.067	0.163		
p 67	-0.034	-0.133	-0.016	0.031	-0.205	0.088	-0.012	0.227	0.098	0.068	0.061	-0.059	0.197	-0.141	-0.070	0.134	
p 71	-0.017	0.098	-0.043	0.049	-0.023	-0.116	0.153	0.022	-0.040	-0.101	-0.088	0.071	-0.050	0.060	0.064	-0.002	-0.027

Table 8: Estimation results with the “random” simulation method - model M1

Number of draws	**50**		**100**		**200**		**500**	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Sup	-4.009	0.490	-4.241	0.456	-6.914	0.642	-5.731	0.548
Small hyp	-8.296	1.055	-9.799	1.207	-9.807	1.047	-8.280	1.143
Large hyp	0.421	0.415	0.272	0.453	-1.965	0.489	-1.638	0.437
Special	-7.270	1.119	-7.777	1.124	-9.706	1.467	-10.624	1.424
Gender Sup	0.151	0.271	0.278	0.264	0.667	0.418	0.913	0.528
Gender Small hyp	2.388	0.508	2.780	0.594	2.966	0.469	1.772	0.486
Gender Large hyp	0.123	0.261	0.271	0.292	0.958	0.263	0.739	0.269
Gender Special	2.291	0.986	2.323	0.798	1.356	0.470	0.924	0.519
Age Sup	-0.218	0.042	0.196	0.041	0.454	0.053	0.361	0.048
Age Small hyp	0.335	0.064	0.399	0.065	0.351	0.056	0.225	0.058
Age Large hyp	0.097	0.034	0.108	0.041	0.124	0.053	0.189	0.051
Age Special	0.901	0.125	0.936	0.170	0.937	0.203	0.814	0.192
Net income Sup	0.003	0.063	0.014	0.046	0.060	0.049	0.047	0.058
Net income Small hyp	0.180	0.046	0.241	0.049	0.290	0.095	0.366	0.089
Net income Large hyp	-0.231	0.040	-0.251	0.049	-0.061	0.052	-0.092	0.046
Net income Special	-0.367	0.105	-0.315	0.100	-0.105	0.083	0.131	0.112
Log likelihood	**-4113.443**		**-4167.609**		**-4157.235**		**-4155.410**	

Table 9: Estimation results with the “Halton” simulation method - model M1

Number of draws	**50**		**100**		**200**		**500**	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Sup	-2.153	0.465	-3.311	0.481	-4.426	0.518	-5.079	0.531
Small hyp	-8.086	0.707	-9.049	0.896	-8.326	0.872	-8.801	0.794
Large hyp	-3.236	0.371	-3.795	0.449	-3.109	0.432	-2.716	0.538
Special	-9.140	1.127	-11.630	1.109	-10.613	1.276	-11.313	1.136
Gender Sup	-0.751	0.274	-0.161	0.245	0.382	0.261	1.072	0.309
Gender Small hyp	2.236	0.412	2.419	0.379	1.668	0.382	1.976	0.424
Gender Large hyp	0.372	0.265	0.478	0.255	0.581	0.243	0.782	0.289
Gender Special	0.680	0.549	1.049	0.376	1.371	0.489	1.108	0.427
Age Sup	0.093	0.040	0.076	0.034	0.269	0.043	0.371	0.048
Age Small hyp	0.143	0.080	0.265	0.066	0.368	0.074	0.319	0.081
Age Large hyp	0.088	0.063	0.206	0.037	0.175	0.038	0.166	0.042
Age Special	0.555	0.106	0.340	0.059	0.646	0.087	0.792	0.094
Net income Sup	0.002	0.071	0.013	0.057	0.136	0.069	0.092	0.072
Net income Small hyp	0.235	0.047	0.185	0.058	0.382	0.061	0.275	0.084
Net income Large hyp	0.035	0.045	-0.062	0.040	-0.142	0.048	-0.096	0.051
Net income Special	0.122	0.090	0.665	0.097	0.427	0.082	0.359	0.114
Log likelihood	**-4083.347**		**-4102.078**		**-4137.587**		**-4159.271**	

Table 10: Estimation results with the “Shuffled Halton” simulation method - model M1

Number of draws	**50**		**100**		**200**		**500**	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Sup	-4.523	0.499	-6.005	0.480	-7.042	0.607	-5.354	0.976
Small hyp	-5.502	0.608	-6.448	0.524	-6.765	0.739	-6.068	2.339
Large hyp	-3.827	0.530	-3.528	0.448	-2.887	0.453	-1.951	0.581
Special	-11.940	1.399	-14.421	1.782	-13.061	1.442	-7.698	1.849
Gender Sup	1.936	0.221	1.026	0.217	1.008	0.345	0.719	0.661
Gender Small hyp	0.991	0.382	0.894	0.455	0.273	0.599	0.454	0.618
Gender Large hyp	-0.114	-0.114	-0.224	0.274	-0.657	0.263	-0.509	0.559
Gender Special	4.432	0.672	4.851	0.675	2.280	0.566	0.413	1.093
Age Sup	0.277	0.041	0.350	0.043	0.375	0.048	0.303	0.080
Age Small hyp	-0.053	0.049	-0.042	0.046	0.042	0.046	0.043	0.186
Age Large hyp	0.252	0.050	0.226	0.055	0.191	0.044	0.072	0.096
Age Special	0.706	0.116	0.888	0.167	0.794	0.115	0.308	0.239
Net income Sup	-0.063	0.041	0.168	0.050	0.239	0.042	0.049	0.064
Net income Small hyp	0.492	0.056	0.574	0.075	0.546	0.066	0.389	0.068
Net income Large hyp	-0.050	0.061	-0.125	0.077	-0.076	0.052	-0.081	0.061
Net income Special	-0.008	0.079	0.044	0.117	0.218	0.075	0.226	0.135
Log likelihood	**-4085.928**		**-4109.203**		**-4138.799**		**-4168.654**	

Table 11: Estimation results of model M2 - one random coefficient

	Random	**Halton p=2**	**Halton p=11**	**Shuffle p=2**	**Shuffle p=11**
Number of draws			**50**		
Log likelihood	-6195.348	-6191.359	-6197.747	-6191.334	-6192.880
Sup	-1.963	-2.042	-1.773	-1.455	-1.695
Number of draws			**100**		
Log likelihood	-6193.808	-6191.963	-6190.724	-6189.469	-6197.682
Sup	-2.391	-2.068	-2.168	-1.768	-1.831
Number of draws			**200**		
Log likelihood	-6193.490	-6191.858	-6190.685	-6191.659	-6191.319
Sup	-2.221	-2.114	-2.178	-1.669	-1.787
Number of draws			**500**		
Log likelihood	-6191.073	-6191.662	-6190.939	-6191.965	-6192.428
Sup	-2.161	-2.111	-2.136	-2.116	-2.136
Number of draws			**1000**		
Log likelihood	-6191.967	-6191.759	-6191.262	-6191.693	-6191.561
Sup	-2.141	-2.107	-2.119	-2.158	-2.108

Table 12: Estimation results of model M3 - two random coefficients

	Random	**Halton p=2,3**	**Halton p=11,13**	**Shuffle p=2,3**	**Shuffle p=11,13**
Number of draws			**50**		
Log likelihood	-5748.563	-5751.844	-5755.472	-5762.386	-5753.277
Sup	-2.745	-2.147	-2.283	-1.534	-1.771
Small hyp	-2.945	-1.866	-2.414	-2.609	-2.725
Number of draws			**100**		
Log likelihood	-5739.551	-5739.811	-5742.100	-5739.607	-5753.977
Sup	-2.841	-2.302	-2.283	-1.670	-1.688
Small hyp	-2.519	-2.574	-2.295	-2.764	-2.268
Number of draws			**200**		
Log likelihood	-5739.407	-5742.028	-5737.943	-5743.663	-5740.123
Sup	-2.495	-2.455	-2.429	-2.171	-1.498
Small hyp	-2.538	-2.441	-2.333	-2.608	-2.559
Number of draws			**500**		
Log likelihood	-5740.982	-5742.163	-5740.999	-5741.960	-5739.112
Sup	-2.409	-2.238	-2.345	-2.302	-2.272
Small hyp	-2.283	-2.446	-2.319	-2.519	-2.524
Number of draws			**1000**		
Log likelihood	-5741.747	-5742.671	-5742.459	-5741.341	-5741.787
Sup	-2.178	-2.212	-2.226	-2.377	-2.365
Small hyp	-2.421	-2.419	-2.449	-2.381	-2.348

Table 13: Estimation results of model M4 - three random coefficients

	Random	**Halton p=2,3,5**	**Halton p=11,13,17**	**Shuffle p=2,3,5**	**Shuffle p=11,13,17**
Number of draws			**50**		
Log likelihood	-5266.491	-5281.751	-5286.768	-5309.940	-5271.793
Sup	-0.988	-2.637	-3.440	-1.547	-1.472
Small hyp	-4.058	-1.551	-1.599	-2.098	-2.553
Large hyp	-0.814	-0.783	-1.219	-1.307	-1.025
Number of draws			**100**		
Log likelihood	-5253.562	-5259.680	-5255.974	-5304.329	-5256.807
Sup	-2.127	-2.744	-3.007	-2.630	-1.668
Small hyp	-2.831	-2.202	-2.012	-1.267	-2.508
Large hyp	-1.555	-1.365	-1.122	-1.523	-1.261
Number of draws			**200**		
Log likelihood	-5247.030	-5246.615	-5255.175	-5270.926	-5246.925
Sup	-2.629	-2.303	-2.941	-2.163	-1.708
Small hyp	-2.617	-2.428	-1.809	-1.640	-2.380
Large hyp	-1.497	-1.266	-1.208	-1.314	-0.878
Number of draws			**500**		
Log likelihood	-5247.267	-5245.633	-5247.222	-5245.646	-5241.692
Sup	-2.440	-2.370	-2.210	-2.585	-2.439
Small hyp	-2.178	-2.525	-2.578	-2.475	-2.765
Large hyp	-1.404	-1.312	-1.291	-1.466	-1.356
Number of draws			**1000**		
Log likelihood	-5245.511	-5248.607	-5248.350	-5254.662	-5244.442
Sup	-2.250	-2.176	-2.396	-2.149	-2.234
Small hyp	-2.582	-2.479	-2.356	-2.674	-2.684
Large hyp	-1.412	-1.383	-1.140	-1.361	-1.344

Table 14: Estimation results of model M5 - four random coefficients

	Random	**Halton p=2,3,5,7**	**Halton p=11,13,17,19**	**Shuffle p=2,3,5,7**	**Shuffle p=11,13,17,19**
Number of draws			**50**		
Log likelihood	-4329.232	-4297.051	-4316.397	-4335.803	-4330.395
Sup	-1.943	-2.977	-2.701	-1.574	-2.243
Small hyp	-2.490	-2.584	-2.159	-2.756	-1.763
Large hyp	-1.662	-1.828	-1.901	-2.011	-1.420
Special	-2.776	-2.989	-3.270	-3.586	-1.859
Number of draws			**100**		
Log likelihood	-4284.558	-4291.312	-4242.447	-4307.689	-4264.594
Sup	-2.213	-1.974	-2.508	-1.994	-2.365
Small hyp	-3.210	-2.242	-2.746	-2.492	-2.400
Large hyp	-2.168	-1.222	-1.696	-2.105	-1.359
Special	-2.776	-1.819	-3.319	-2.554	-2.312
Number of draws			**200**		
Log likelihood	-4266.697	-4264.125	-4248.160	-4269.429	-4256.240
Sup	-2.998	-2.176	-2.347	-1.888	-2.420
Small hyp	-1.959	-2.290	-2.734	-1.857	-2.040
Large hyp	-1.854	-2.171	-1.517	-2.366	-1.301
Special	-3.524	-3.315	-2.948	-2.746	-3.257
Number of draws			**500**		
Log likelihood	-4242.530	-4234.116	-4238.551	-4241.001	-4237.217
Sup	-2.826	-3.162	-2.218	-2.753	-3.539
Small hyp	-2.609	-2.414	-2.890	-2.862	-2.478
Large hyp	-1.992	-2.118	-2.095	-2.115	-1.803
Special	-3.419	-3.100	-3.222	-3.627	-3.045
Number of draws			**1000**		
Log likelihood	-4225.988	-4228.463	-4230.798	-4224.732	-4240.638
Sup	-2.672	-3.184	-2.948	-2.825	-2.448
Small hyp	-3.098	-3.144	-2.925	-3.226	-2.655
Large hyp	-1.746	-1.927	-1.912	-1.993	-1.964
Special	-3.639	-3.978	-3.310	-3.535	-3.286

Bisher in der Reihe *Hohenheimer Agrarökonomische Arbeitsberichte* erschienen:

Becker, T.; Benner, E (2000): Zur Problematik der Herkunftsangabe im regionalen Marketing
Arbeitsbericht Nr. 1

Chaipan, C. (2000): The Euro and its Impact on ASEAN Economies
Arbeitsbericht Nr. 2

Henze, A. (2000): Marktwirtschaft - Wirtschaftliche Freiheit, motiviertes Handeln und Wettbewerb: Quellen des Wohlstands
Arbeitsbericht Nr. 3

Benner, E. (2000): Zur effizienten Herkunftsangabe im europäischen Binnenmarkt
Arbeitsbericht Nr. 4

Vorgrimler, D. (2000): Wettbewerbstheorie und stagnierende Märkte
Arbeitsbericht Nr. 5

Beerbaum, S. (2001): Grundzüge einer internationalen Zusammenarbeit im Klimaschutz aus ökonomischer Sicht
Arbeitsbericht Nr. 6

Vorgrimler, D.; Wübben, D. (2001): Prognose der Entwicklung des Agrartechnikmarktes - Eine Expertenbefragung nach der Delphi-Methode
Arbeitsbericht Nr. 7

Tesch, I. (2003): Informationsbedarf und Informationsbeschaffung von Konsumenten bei Lebensmitteln pflanzlicher Herkunft - Eine empirische Untersuchung anhand von Fokus-Gruppen -
Arbeitsbericht Nr. 8

Benner, D. (2004): Quality Ambiguity and the Market Mechanism for Credence Goods
Arbeitsbericht Nr. 9

Benner, E., Kliebisch, C. (2004): Regio-Marketing-Strategien des Lebensmitteleinzelhandels
Arbeitsbericht Nr. 10

Benner, E., Heidecke, S.-J. (2005): Grundpreisaufschläge bei Groß- und Familienpackungen - eine empirische Untersuchung im deutschen und französischen Lebensmitteleinzelhandel -
Arbeitsbericht Nr. 11

Becker, T. (2006): Zur Bedeutung geschützter Herkunftsangaben.
Arbeitsbericht Nr. 12, 1. und 2. Auflage.

Elsäßer, A., Benner, E., Becker, T. (2006): Marketing auf Wochenmärkten.
Arbeitsbericht Nr. 13

Becker, T. (2006): Die CMA auf dem Prüfstand.
Arbeitsbericht Nr. 14

Elsäßer, A., Benner, E., Becker, T. (2006): Marketing auf Wochenmärkten.
Arbeitsbericht Nr. 13

Becker, T. (2006): Die CMA auf dem Prüfstand.
Arbeitsbericht Nr. 14

Staus, A. (2007): An Ordinal Regression Model using Dealer Satisfaction Data.
Arbeitsbericht Nr. 15

Kliebisch, C., Rügge, M. (2007): Alte und neue Wege des Gemeinschaftsmarketings für Agrarprodukte und Lebensmittel.
Arbeitsbericht Nr. 16

jeweils als pdf-file unter: https://marktlehre.uni-hohenheim.de (Kapitel Forschung – Arbeitsberichte)