The Order of Variables, Simulation Noise and Accuracy of Mixed Logit Estimates

Marco A. Palma[a], Yajuan Li[b] , Dmitry V. Vedenov[c], and David Bessler[d]

[a] *Corresponding author*. Associate Professor and Extension Economist, Department of Agricultural Economics. Texas A&M University. 2124 TAMU, College Station, Texas, 77843. (979) 845-5284. E-mail: mapalma@tamu.edu

[b] Research Assistant, Department of Agricultural Economics, TAMU.

[c] Associate Professor, Department of Agricultural Economics, TAMU.

[d] Professor, Department of Agricultural Economics, TAMU.

Selected Paper prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, Massachusetts, July 31-August 2

The Order of Variables, Simulation Noise and Accuracy of Mixed Logit Estimates

Abstract

The simulated choice probabilities in Mixed Logit are approximated numerically from a multidimensional integral with a mixing distribution from a multivariate density function of the random parameters. Theoretically the order in which the variables are estimated should not matter; however, due to the inherent simulation 'noise' the magnitude of the estimated coefficients differs depending on the arbitrarily selected order in which the random variables enter the estimation procedure. This problem is exacerbated with a low number of draws or if correlation among coefficients is allowed. If correlation among the random parameters is allowed the variable ordering effects arise from simulation noise and from the Cholesky factorization used to allow for correlation. Ignoring the potential ordering effects in simulated maximum likelihood estimation methods seriously compromises the ability for replicating the results and can inadvertently influence policy recommendations. The simulation noise is independent of the number of integrating dimensions for random draws, but it increases for Halton draws. Hence, better coverage is achieved with Halton draws for small integrating dimensions, but random draws provide better coverage for larger dimensions.

JEL codes: C25, C63

1. Introduction

Mixed Logit (ML) is one of the most widely used models to analyze discrete choice data. Substantial improvements in computing power and speed and the model flexibility to approximate any random utility model make it ideal for modeling behavioral choices (McFadden and Train 2000). ML is a generalization of the standard closed-form Logit model that allows the parameters associated with each observed variable to vary randomly across respondents (Revelt and Train 1998) and choice situations (Hess and Rose 2009). Mixed Logit relaxes the independence of irrelevant alternatives restriction of standard logit specifications. Many popular statistical analysis software packages include built-in routines for estimating mixed logit using simulated maximum likelihood (SML). See Hensher and Greene (2003) and Train (2009) for an overview. The simulation procedure involves generating *pseudorandom* (commonly referred to as random) or *quasirandom* draws from the mixing density of the random coefficients to numerically approximate the multidimensional integral of the choice probabilities (Train 2009, Greene 2012, Hensher and Greene 2003, Geweke, Keane, and Runkle 1994, McFadden and Ruud 1994). The most commonly used quasirandom sequence of draws is Halton (Train 2000). If more than one random coefficient is specified, as in practically every ML application, then the choice probabilities involve a multidimensional integral with the mixing distribution of the random coefficient taken from a multivariate density (Cappellari and Jenkins 2006).

By construction, for a given number of observations, the Monte Carlo or Halton numerical integration has some inherent 'noise' depending on the number of draws taken and the sequence in which the draws are generated. In all statistical packages, the order in which the variables are specified in the estimation procedure determines which draws are used for each parameter. Theoretically, the order of variables should not matter. However, due to simulation noise, the arbitrary decision on the order specified by the user affects the parameter estimates. In other words, a model with three variables estimated in the following order X_1, X_2, and X_3, would produce different coefficients than the model with the same three variables ordered X_3, X_2, and X_1 or any other possible combination. The problem is exacerbated if correlation among the parameters is allowed. The correlation of random parameters is introduced through Cholesky factorization, which is not invariant to the order of variables (Koop, Pesaran, and Potter 1996, Pesaran and Shin 1998). The replication of the results is compromised if not enough information is provided about the estimation procedures. Furthermore, policy recommendations based on the estimates from a single ordering may be unreliable unless estimated coefficients are properly tested for stability. The objective of this paper is to quantify the effects of the variable ordering on parameter estimates and willingness-to-pay measures (WTP) in mixed logit models. In order to achieve the general objective, the following specific objectives will be accomplished: 1) Separating and comparing the variability of estimated coefficients arising from simulation noise versus Cholesky factorization for Halton and random draws. 2) Evaluating differences in simulation noise of Halton and random sequences based on the number of integrating dimensions. 3) Suggesting approaches to deal with the ordering effects.

We evaluate the ordering effects empirically using a widely known data set documented in commonly available software routines for mixlogit (Hole 2007, Croissant 2011). The results provide important contributions to the choice modeling literature. If no correlation is allowed among the random parameters the variable ordering effects arise exclusively due to simulation noise. Different variable orderings produce different mean and standard deviation parameters, especially with a low number of draws. As expected, we find that increasing the number of draws (both Halton and Random) reduces the simulation noise. The parameter estimates and model fit converge as the number of draws is artificially high (more than 2000 draws). Comparing how the simulation noise is affected by the number of integrating dimensions is difficult, since adding additional parameters produce a different model specification. We deal with this limitation by estimating the same coefficients. For random draws the procedure involves taking different random draws for the coefficients which is invariant to the number of integrating dimensions (Drukker and Gates 2006). For Halton draws, high prime numbers (31, 37, 41, 43, 47) are used as a proxy to represent simulation noise in high integrating dimensions (>10) compared to the normally used prime numbers in low integrating dimensions (2, 3, 5, 7, 9, 11). Halton draws provide a better model fit and lower simulation noise with smaller integrating dimensions. However, random draws perform better with larger integrating dimensions. Although the simulation noise effect reduces the importance of the type of draws used when the number of draws goes beyond 2000, the results are quite important, particularly since most applications of mixed logit are seldom estimated with such a large number of draws.

If correlation among the random parameters is allowed the variable ordering effects arise from simulation noise and also from the Cholesky factorization both of which are similar in size. The variability of the average of the mean parameters was higher for the Cholesky factorization effect. The standard deviation parameters were smaller for the Cholesky effect but their variance was considerable higher. Ignoring variable ordering effects compromise the replication of results and may have impact policy recommendations. The rest of the paper is organized as follows. Section 2 describes the mixed logit and the source of the simulation noise, Section 3 presents an empirical application using a well-known data set to illustrate the extend of the problem, and Section 4 gives some recommendations on how to deal with the ordering effects and concludes.

2. **The Mixed Logit Model**

An individual *n* faces a choice amongst a set of *J* available alternatives at different T time periods. The utility person *n* derives from alternative *i* at choice situation *t* is represented by $U_{nit} = \beta_n x_{nit} + \varepsilon_{nit}$ where x_{nit} is a vector of observed variables, β_n is an unobserved vector of coefficients for each individual *n* that varies in the population with density $f(\beta_n|\theta^*)$ where θ^* are the true parameters of the multivariate distribution of β and ε_{nit} is an unobserved random term independently and identically distributed (IID) extreme value. Conditional on β_n, the probability that individual *n* chooses alternative *i* in period *t* is given by:

$$L_{nit}(\beta_n) = \frac{e^{\beta_n X_{nit}}}{\sum_j e^{\beta_n X_{njt}}} \qquad (1)$$

The unconditional probability is the integral of the conditional probability over all possible values of β_n:

$$P_n(\theta^*) = \int L_n(\beta_n) f(\beta_n|\theta) d\,\beta_n \qquad (2)$$

The β_n are individual specific taste parameters and the β vector takes the form $\beta = b + \eta_n$ where b is the population mean and η_n are individual deviations from the average population (Calfee, Winston, and Stempski 2001). The estimated parameters can take a number of distributional forms including normal, log-normal, triangular, etc. The log-likelihood function is given by $LL(\theta) = \sum_n lnP_n(\theta)$. The integral in equation (2) cannot be solved analytically and it is approximated numerically generally using SML. The procedure involves for any θ drawing a value of β_n from its density $f(\beta_n|\theta)$ and using it to calculate $L_n(\beta_n)$ in equation (2) from this draw, repeating the procedure R times and averaging the results. The simulated probability then becomes $SP_{nit} = (1/R) \sum_{r=1}^{R} L_{nit}\left(\beta_n^{r|\theta}\right)$ and the simulated log-likelihood function is $SLL(\theta) = \sum_n ln(SP_{nit})$. The simulated probability SP_{nit} is an unbiased estimator of the actual probability

P_{nit}. For a review of the properties of the estimators see Train (2009). Note that the variance of the estimator, and hence the simulation noise decreases as the number of draws (R) increases. If more than one parameter is specified to be random, as in virtually every mixlogit application, then the SLL procedure has to numerically approximate a multidimensional integral drawing from a multivariate density of $f(\beta_n|\theta)$. In selecting the type and number of draws R, randomness is not as important as the uniform coverage over the domain of integration. While pseudorandom draws (hereafter random) are not correlated, quasirandom draws such as Halton, are negatively correlated by design (Train 2009). The use of quasirandom draws, however, comes at a cost. The bound of the error for the numerical approximation of a multivariate integral is independent of the number of integrating dimensions, d (i.e. number of random variables) for random draws, but it increases exponentially for Halton draws. See Drukker and Gates (2006) for a formal treatise. Better coverage is achieved with Halton draws for small integrating dimensions, but random draws provide better coverage for larger dimensions. In order to deal with this problem, when specifying a large number of parameters, several variations of Halton sequences are normally used, including Halton-random, Halton-shuffled, and Halton-scrambled. Modified Latin Hypercube Sampling (MLHS) draws outperformed all of these types of draws in an application for vehicle choices (Hess, Train, and Polak 2006). However, its use has not expanded in the literature, perhaps because of the lack of availability in common computer software.

The simulation noise for random draws can be observed by changing the seed number, since a different seed produces different draws. An alternative way to observe the simulation noise is to predefine a matrix of draws D of $d \times (N \cdot R)$, where d is the number of integrating dimensions determined by the number of random parameters, N is the number of observations and R is the number of draws. The simulation noise can be assessed by changing the row of the draw matrix *D* corresponding to each random parameter. Halton draws are not random, but a deterministic sequence based on prime numbers, with one prime number for each integrating dimension starting with 2. Hence, the second approach is used to generate the simulation noise by changing the rows of the matrix of Halton draws *D* corresponding to each parameter. A similar procedure is used in Train (2000) for five possible combinations. In our application, when applied to random draws, this procedure resulted in equivalent parameter distributions as using a different seed number to generate the random draws in 10 of the 11 estimated parameters (Kolmogorov-Smirnov $p<0.05$).

3. The Data

We employ a widely used data set discussed in Train (2009), and included in the documentation of the mixed logit routines for the Stata and R software packages (Hole 2007, Croissant 2011). The data consists of stated-preference choices of 361 residential customers for electricity supplier. Each customer was presented with 8-12 hypothetical choice situations consisting of four alternatives with differences in *prices* (in cents per kilowatthour), contract length (contract), time of the day prices (tod), or seasonal pricing (seasonal), and whether the supplier was their local utility company (local), a well-known company (wknown) or an unfamiliar company. The price coefficient was fixed and all other five coefficients were specified to be normally distributed for a total of 11 parameters to be estimated (1 fixed, 5 means and 5 standard deviations). With 5! possible ordering combinations of the random variables, 120 models for Halton and Random sequences for different number of draws with and without correlation are estimated.

4. Results

The distribution of the seasonal pricing parameter and associated mean WTP estimates (in cents) over the 120 possible orderings of the observed variables assuming no correlation among the random parameters is shown in Figure 1. The results show a wide range of variation of the coefficient based only in changing the order of the observed variables in the estimation procedure (Panel A). Although the sign of the parameters did not change, their fluctuation was significant enough to produce mean WTP values that differ in magnitude (Panel B).

Table 1 shows the average parameter estimates of the mean and standard deviations representing the 'simulation noise' over all 120 possible orderings without correlation for different number of Halton and random draws. The numbers in parenthesis are the standard deviations of the estimated parameters over all 120 possible ordering combinations. The results are shown for 100 and 2000 draws. The models were also computed for 200, 500, and 1000 draws. The results were systematic for increases in the number of draws. The full tables are available in the appendix. The results with the Halton heading were computed using the primes 2, 3, 5, 7, and 11 in a predefined matrix of draws D of dimensions [5 x (361xR)]. The simulation noise was assessed by systematically changing the assigned prime position for all possible combinations (5!=120). This is equivalent to changing the order of the variables in the estimation procedure. For the random sequences the D matrix of draws is generated by drawing pseudorandom numbers from a random uniform distribution. Similarly, the simulation noise was generated by systematically changing the row vectors of draws assigned to each variable. The results with the Halton High Primes heading are a quasi-representation of the simulation noise for large integrating dimensions. Comparing the simulation noise effects of the number of integrating dimensions on the parameter estimates is difficult, since adding additional parameters results in a different model specification. We deal with this limitation by estimating the same coefficients adapted to a large integrating dimension scenario. For random draws the procedure involves taking different random draws for the coefficients which is invariant to the number of integrating dimensions (Drukker and Gates 2006). For Halton draws, high prime numbers (31, 37, 41, 43, 47) are used as a proxy to represent simulation noise in high integrating dimensions (>10 random parameters) compared to the normally used primes in low integrating dimensions (2, 3, 5, 7, 9, 11).

When correlation among the random parameters is not allowed the variable ordering effect comes exclusively from the simulation noise. The variable order produced variations in the mean and standard deviation parameters, especially with a low number of draws. The reported estimates in the table are the average over all the 120 possible order combinations. The average appears to be moving away from zero as R increases, with larger average estimates for positive coefficients and smaller average estimates for negative coefficients. This result was robust across all specifications. As expected, we find that increasing the number of draws (both Halton and Random) reduces the simulation noise. This can be seen as a reduction in the standard deviation of the average (in parenthesis) of going from 100 to 2000 draws. With the same number of estimated coefficients and same number of draws, the log-likelihood (LL) value itself provides a measure of model fit. Halton draws provide a better model fit and lower simulation noise with smaller integrating dimensions. The LL increases with R and it is higher with Halton draws compared to random draws. The variance of the LL decreases as the number of draws is increased. Random draws perform better than Halton draws with larger integrating dimensions (using large prime numbers). The D-error reported in the table is the determinant of the entire variance-covariance (VC) matrix normalized by the number of coefficients K and multiplying it by the number of respondents $D = |VC|^{1/K} \cdot N$. A lower D-error indicates higher efficiency of the models by providing lower variance in the VC matrix. The D-error decreases monotonically, only after taking more than 200 draws (results not shown but available in the appendix). This result seems to indicate that for this dataset there is not enough uniform coverage for random or Halton sequences with less than 200 draws. The average parameter estimates, LLs and D-errors converge as the number of draws is considerably increased (more than 2000 draws). Intuitively, if R is sufficiently high, the coverage of the draws becomes good enough that the type and sequence of draws becomes less important. Although the simulation noise effect fades when the number of draws goes beyond 2000, the results are quite important, particularly since most researchers seldom estimate mixlogit models with such a large number of draws. Recall that table 1 shows the average of the parameters over all 120 combinations. Table 2 presents the average and range of the associated mean willingness-to-pay values. Depending on the arbitrary order chosen by the investigator there is some variation in the WTP values even with a vast number of draws.

If correlation is allowed the variable ordering effects arise from simulation noise and also from the Cholesky factorization both of which are similar in size. The average parameter estimates of the mean and standard deviations representing the 'simulation noise' and 'Cholesky factorization effect' over all 120 possible orderings with correlation for different number of Halton and random draws are shown in Tables 3 and 4 respectively. The procedure used to capture the simulation noise was the same as before. In order to capture the 'Cholesky factorization effect' the order of the variables was systematically changed, but the prime number associated with each variable remained with the variable for all 120 order combinations. Therefore the variable ordering effects are separated into the simulation noise effect and the

Cholesky factorization effect. The results for simulation noise were similar to the 'no-correlation' results presented above, except that the simulation noise was amplified when correlation is allowed. The variability of the average of the mean parameters was even higher for the Cholesky factorization effect. The standard deviation parameters were smaller for the Cholesky effect but their variance was considerable higher.

The results provide a cautionary warning that simply choosing an arbitrary order of the observed variables will provide one solution in the above distribution of parameter estimates and associated WTP values. Since Halton draws are a deterministic sequence, the results should be congruent if the same procedure is used. At a minimum, in order for the results to be replicated, mixed logit estimated using Halton draws should report the order of the variables or more generally the prime number used for each coefficient, the number of draws initially burned, the maximization algorithm used and convergence criteria and threshold.

One possible solution to deal with the variable ordering effects is to estimate all possible variable orders and report the averages just as we did in our application. However, this approach may be computationally expensive and impractical in many situations, but more importantly, it may be considered "mongrel" estimates that may not adequately represent the correct causal structure of the data. So what criteria should be used to select the appropriate ordering structure? If the data comes from behavioral data (rather than constructed from experimental design) a method can be applied to orthogonalize shocks in the Cholesky factorization and placing some restrictions on the variables ordering. The approach involves placing variables that are weakly exogenous with respect to other variables earlier in the ordering.

In order to establish causality direction we use machine learning to create directed acyclic graphs (DAG). DAGs are used to sort out causal flows of the observed random variable information of the covariance matrix of ε^*_{nit} (Pearl 2000). In a DAG, directed arrows are used to represent causal flows. If variables are not connected by arrows, then it implies that there is no direct causal effect. The approach we applied requires weakly exogenous variables to appear earlier in the ordering using the LiNGAM (Linear, Non-Gaussian, Acyclic causal Models) algorithm developed by Shimizu et al. (2006) and applied by Lai and Bessler (2015) in TETRAD V (Spirtes, Glymour, and Scheines 2000).

The arrow heads in Figure 2 represent the flows of causal patters among the observed random variables. The more (less) arrowheads that point to a specific variable, the more endogenous (exogenous) the variable is and should be placed later in the ordering. The variable 'contract' is not a cause of any of the other variables and hence should be placed first. The second variable would be 'local'. Next 'wknown' and 'tod' have the same number of arrowheads, 'wknown' causes 'tod', therefore it is suggested to place 'wknown' third, followed by 'tod'. Finally, the last variable would be 'seasonal'. How does the suggested order estimation represent the data? It is fairly close to the average over all 120 possible orderings (Table 6).

Recommendations and Conclusions

The results clearly show that due to simulation noise and Cholesky factorization (if correlation among random parameters is allowed), the arbitrary decision of the order in which the variables are estimated affect the magnitude of the estimated coefficients. This problem is exacerbated with a low number of draws or if correlation among coefficients is allowed. This situation seriously hinders the replicability of the results if not enough information is provided (Chang and Lusk 2011). At a minimum, when reporting results one needs to include the seed number and the estimation order if random draws are used, and the order or more generally the prime number used for each coefficient and the number of draws initially burned. Also, the maximization algorithm used and convergence criteria and threshold should also be reported.

One approach to dealing with the ordering effect and simulation noise is to increase the number of draws to the point where the log-likelihood and parameters (or the D-error) stabilizes. Besides increasing the computing time, the actual number of draws depends on the particular data set used; furthermore, although smaller, even with a very large number of draws there is still ordering effects due to simulation noise or correlation among the parameters.

Another approach would be estimating all or a subset of the possible ordering combinations and report the averages. This procedure is computationally expensive and impractical in many situations. If the data comes from observed real (behavioral) variables, using the average of different orderings would be 'mongrel' estimates that may not adequately represent the correct causal structure of the data. In such cases, it is recommended to select one ordering based on causality. In order to establish causality direction machine learning can be used to create directed acyclic graphs (DAG) to sort out causal flows of the observed random variable information of the covariance matrix of ε_{nit}^{*} (Pearl 2009). The approach involves placing weakly exogenous variables to appear earlier in the ordering. Continuing to ignore the potential ordering effects in simulated maximum likelihood estimation methods seriously compromises the ability for replicating the results and can inadvertently influence policy recommendations.

5. References

Calfee, John, Clifford Winston, and Randolph Stempski. 2001. "Econometric Issues in Estimating Consumer Preferences from Stated Preference Data: A Case Study of the

Value of Automobile Travel Time." *Review of Economics and Statistics* 83 (4):699-707. doi: 10.1162/003465301753237777.

Cappellari, L., and S. P. Jenkins. 2006. "Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation." *Stata Journal* 6 (2):156-189.

Chang, Jae Bong, and Jayson L. Lusk. 2011. "Mixed logit models: accuracy and software choice." *Journal of Applied Econometrics* 26 (1):167-172. doi: 10.1002/jae.1201.

Croissant, Yves. 2011. mlogit: multinomial logit model.

Drukker, D. M., and R. Gates. 2006. "Generating Halton sequences using Mata." *Stata Journal* 6 (2):214-228.

Geweke, John, Michael Keane, and David Runkle. 1994. "Alternative Computational Approaches to Inference in the Multinomial Probit Model." *The Review of Economics and Statistics* 76 (4):609-632. doi: 10.2307/2109766.

Greene, William H. 2012. *Econometric analysis*. Boston: Boston : Prentice Hall.

Hensher, DavidA, and WilliamH Greene. 2003. "The Mixed Logit model: The state of practice." *Transportation* 30 (2):133-176. doi: 10.1023/A:1022558715350.

Hess, Stephane, and John M. Rose. 2009. "Allowing for intra-respondent variations in coefficients estimated on repeated choice data." *Transportation Research Part B: Methodological* 43 (6):708-719. doi: http://dx.doi.org/10.1016/j.trb.2009.01.007.

Hess, Stephane, Kenneth E. Train, and John W. Polak. 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: Methodological* 40 (2):147-163. doi: http://dx.doi.org/10.1016/j.trb.2004.10.005.

Hole, Arne Risa. 2007. "Estimating mixed logit models using maximum simulated likelihood." *Stata J* 7 (3):388-401.

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter. 1996. "Impulse response analysis in nonlinear multivariate models." *Journal of Econometrics* 74 (1):119-147. doi: http://dx.doi.org/10.1016/0304-4076(95)01753-4.

Lai, Pei-Chun, and David A. Bessler. 2015. "PRICE DISCOVERY BETWEEN CARBONATED SOFT DRINK MANUFACTURERS AND RETAILERS: A DISAGGREGATE ANALYSIS WITH PC AND LiNGAM ALGORITHMS." *Journal of Applied Economics* 18 (1):173-197. doi: http://dx.doi.org/10.1016/S1514-0326(15)30008-8.

McFadden, Daniel, and Paul A. Ruud. 1994. "Estimation by Simulation." *The Review of Economics and Statistics* 76 (4):591-608. doi: 10.2307/2109765.

McFadden, Daniel, and Kenneth Train. 2000. "Mixed MNL models for discrete response." *Journal of applied Econometrics* 15 (5):447-470.

Pearl, Judea. 2000. *Causality : models, reasoning, and inference*. Cambridge, U.K. New York: Cambridge, U.K. New York : Cambridge University Press.

Pesaran, H. Hashem, and Yongcheol Shin. 1998. "Generalized impulse response analysis in linear multivariate models." *Economics Letters* 58 (1):17-29. doi: http://dx.doi.org/10.1016/S0165-1765(97)00214-0.

Revelt, David, and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics* 80 (4):647-657. doi: 10.1162/003465398557735.

Shimizu, Shohei, Patrik O. Hoyer, Aapo Hyv, #228, rinen, and Antti Kerminen. 2006. "A Linear Non-Gaussian Acyclic Model for Causal Discovery." *J. Mach. Learn. Res.* 7:2003-2030.

Spirtes, Peter, Clark N Glymour, and Richard Scheines. 2000. *Causation, prediction, and search*. Vol. 81: MIT press.

Train, Kenneth. 2000. "Halton sequences for mixed logit." *Department of Economics, UCB.*

Train, Kenneth E. 2009. *Discrete choice methods with simulation*: Cambridge university press.

Table 1. Average parameter estimates of the mean and standard deviations representing the "simulation noise" over all 120 possible orderings without correlation for different number of Halton and random draws.

Variable	**Halton**		**Random**		**Halton High Primes** [a]	
	100 [b,c]	**2000**	**100**	**2000**	**100**	**2000**
			Mean Parameters			
Price	-0.8791	-0.9374	-0.8661	-0.9328	-0.7843	-0.9300
	(0.007)	(0.002)	(0.008)	(0.003)	(0.011)	(0.004)
Contract	-0.2114	-0.2281	-0.2031	-0.2283	-0.1796	-0.2277
	(0.018)	(0.005)	(0.018)	(0.006)	(0.015)	(0.009)
Local	2.1393	2.3099	2.0770	2.2894	1.8603	2.2886
	(0.065)	(0.026)	(0.061)	(0.035)	(0.067)	(0.041)
Wknown	1.5199	1.6503	1.4837	1.6301	1.3192	1.6274
	(0.036)	(0.014)	(0.036)	(0.019)	(0.029)	(0.023)
Tod	-8.3854	-9.1980	-8.2527	-9.1277	-7.4720	-9.0665
	(0.159)	(0.062)	(0.149)	(0.077)	(0.193)	(0.085)
Seasonal	-8.6003	-9.3526	-8.4853	-9.3166	-7.5925	-9.2636
	(0.098)	(0.046)	(0.098)	(0.050)	(0.133)	(0.067)
			Standard Deviation Parameters			
Contract	0.3768	0.4034	0.3688	0.4017	0.2597	0.3988
	(0.015)	(0.005)	(0.015)	(0.006)	(0.033)	(0.007)
Local	1.5002	1.8142	1.0719	1.7979	0.7692	1.7695
	(0.285)	(0.026)	(0.976)	(0.031)	(0.630)	(0.039)
Wknown	0.3810	1.2207	0.1494	1.2017	0.1386	1.1549
	(0.902)	(0.018)	(0.900)	(0.025)	(0.500)	(0.219)
Tod	2.7127	2.9474	2.6830	2.9652	1.9819	2.8999
	(0.150)	(0.060)	(0.129)	(0.075)	(0.418)	(0.077)
Seasonal	1.9779	2.2123	1.8934	2.1839	1.1176	2.1605
	(0.102)	(0.046)	(0.093)	(0.051)	(0.631)	(0.056)
Log-Likelihood	-3973.884	-3908.861	-4009.320	-3914.525	-4199.030	-3914.728
	(9.918)	(2.479)	(13.173)	(3.508)	(33.583)	(4.442)
D-Error	3.132	2.818	3.035	2.889	2.205	2.891
	(0.345)	(0.167)	(0.278)	(0.176)	(0.263)	(0.209)
Time	1.320	21.437	1.529	21.322	1.734	37.113
	(0.233)	(2.807)	(0.205)	(2.758)	(0.315)	(50.252)

[a] Halton High Primes are used as a proxy to represent simulation noise in high integrating dimensions (>10) which use high prime numbers (31, 37, 41, 43, 47)

[b] The results are shown for 100 and 2000 draws. The models were also computed for 200, 500, and 1000 draws. The results were systematic for increases in the number of draws. The full tables are available in the appendix.

[c] The values in parenthesis are the standard deviations of the estimated parameters over all 120 possible ordering combinations

Table 2. Average and Range of Willingness-to-Pay estimates representing the "simulation noise" over all 120 possible orderings without correlation for different number of Halton and random draws.

Variable		**Halton**		**Random**		**Halton High Primes** [a]	
		100 [b]	**2000**	**100**	**2000**	**100**	**2000**
Contract	*Average*	-0.24	-0.24	-0.23	-0.24	-0.23	-0.24
	Min	-0.30	-0.26	-0.29	-0.26	-0.28	-0.27
	Max	-0.20	-0.23	-0.18	-0.23	-0.18	-0.22
Local	*Average*	2.43	2.46	2.40	2.45	2.37	2.46
	Min	2.24	2.40	2.25	2.34	2.20	2.37
	Max	2.67	2.54	2.61	2.52	2.54	2.56
Wknown	*Average*	1.73	1.76	1.71	1.75	1.68	1.75
	Min	1.65	1.72	1.60	1.70	1.60	1.69
	Max	1.85	1.79	1.82	1.81	1.74	1.80
Tod	*Average*	-9.54	-9.81	-9.53	-9.79	-9.53	-9.75
	Min	-9.92	-9.99	-10.12	-9.95	-10.09	-9.92
	Max	-9.13	-9.66	-9.10	-9.60	-8.83	-9.51
Seasonal	*Average*	-9.78	-9.98	-9.80	-9.99	-9.68	-9.96
	Min	-10.03	-10.08	-10.06	-10.07	-10.05	-10.10
	Max	-9.57	-9.88	-9.58	-9.87	-9.34	-9.84

(a) Halton High Primes are used as a proxy to represent simulation noise in high integrating dimensions (>10) which use high prime numbers (31, 37, 41, 43, 47)

(b) The results are shown for 100 and 2000 draws. The models were also computed for 200, 500, and 1000 draws. The results were systematic for increases in the number of draws. The full tables are available in the appendix.

Table 3. Average parameter estimates of the mean and standard deviations representing the “simulation noise” over all 120 possible orderings with correlation for different number of Halton and random draws.

Variable	**Halton**		**Random**	
	100[(a,b)]	**2000**	**100**	**2000**
		Mean Parameters		
Price	-0.8997	-0.9558	-0.8876	-0.9535
	(0.016)	(0.007)	(0.016)	(0.005)
Contract	-0.2097	-0.2291	-0.2026	-0.2270
	(0.017)	(0.006)	(0.018)	(0.007)
Local	2.4443	2.6166	2.3879	2.6099
	(0.110)	(0.040)	(0.117)	(0.054)
Wknown	1.7889	1.9167	1.7437	1.9094
	(0.073)	(0.025)	(0.082)	(0.035)
Tod	-8.6468	-9.3479	-8.5120	-9.3049
	(0.211)	(0.090)	(0.197)	(0.082)
Seasonal	-8.7356	-9.4351	-8.6163	-9.4003
	(0.150)	(0.077)	(0.150)	(0.069)
		Standard Deviation Parameters		
Contract	0.3770	0.4008	0.3594	0.4000
	(0.014)	(0.006)	(0.068)	(0.006)
Local	0.4167	0.3781	0.4338	0.3797
	(0.198)	(0.064)	(0.236)	(0.082)
Wknown	0.1685	0.1479	0.1867	0.1417
	(0.137)	(0.048)	(0.171)	(0.059)
Tod	-0.0637	-0.1487	-0.0759	-0.1527
	(0.314)	(0.131)	(0.372)	(0.165)
Seasonal	-0.3144	-0.3236	-0.3168	-0.3343
	(0.208)	(0.085)	(0.220)	(0.099)
Log-Likelihood	-3849.628	-3793.790	-3878.622	-3797.291
	(12.694)	(6.744)	(15.907)	(3.656)
D-Error	5.220	4.886	5.443	5.092
	(0.713)	(0.569)	(0.939)	(0.631)
Time	10.730	81.003	7.372	81.649
	(2.551)	(9.595)	(1.805)	(11.558)

[(a)] The results are shown for 100 and 2000 draws. The models were also computed for 200, 500, and 1000 draws. The results were systematic for increases in the number of draws. The full tables are available in the appendix.

[(b)] The values in parenthesis are the standard deviations of the estimated parameters over all 120 possible ordering combinations

Table 4. Average parameter estimates of the mean and standard deviations representing the "Cholesky factorization effect" over all 120 possible orderings with correlation for different number of Halton and random draws.

Variable	**Halton**		**Random**	
	100[a,b]	**1000**	**100**	**1000**
		Mean Parameters		
Price	-0.9031	-0.9497	-0.8823	-0.9445
	(0.022)	(0.014)	(0.016)	(0.011)
Contract	-0.2086	-0.2242	-0.1973	-0.2252
	(0.017)	(0.008)	(0.017)	(0.008)
Local	2.4236	2.5882	2.3293	2.6092
	(0.149)	(0.070)	(0.111)	(0.059)
Wknown	1.7745	1.9014	1.7205	1.9235
	(0.097)	(0.046)	(0.080)	(0.040)
Tod	-8.6280	-9.2930	-8.4234	-9.1745
	(0.203)	(0.148)	(0.197)	(0.127)
Seasonal	-8.7811	-9.3921	-8.5454	-9.2984
	(0.205)	(0.149)	(0.160)	(0.105)
		Standard Deviation Parameters		
Contract	0.0852	0.0868	0.0869	0.0849
	(0.150)	(0.160)	(0.151)	(0.165)
Local	0.8655	1.0560	1.0039	1.0461
	(0.977)	(1.000)	(0.938)	(0.992)
Wknown	0.5744	0.7144	0.6280	0.7032
	(0.670)	(0.713)	(0.631)	(0.708)
Tod	0.8141	0.8766	1.0034	0.9336
	(1.237)	(1.346)	(1.275)	(1.321)
Seasonal	0.3692	0.4056	0.4551	0.4589
	(1.026)	(1.124)	(1.011)	(1.115)
Log-Likelihood	-3845.458	-3798.531	-3897.819	-3811.020
	(25.311)	(13.924)	(14.744)	(9.147)
D-Error	3.655	3.521	4.153	3.760
	(0.867)	(0.978)	(1.008)	(1.021)
Time	9.140	67.586	9.308	66.655
	(2.549)	(16.008)	(3.206)	(15.939)

[a] The results are shown for 100 and 2000 draws. The models were also computed for 200, 500, and 1000 draws. The results were systematic for increases in the number of draws. The full tables are available in the appendix.

[b] The values in parenthesis are the standard deviations of the estimated parameters over all 120 possible ordering combinations

Table 5. Average and Range of Willingness-to-Pay estimates representing the “simulation noise” over all 120 possible orderings with correlation for different number of Halton and random draws.

Variable		Halton		Random	
		100[a,b]	**2000**	**100**	**2000**
Contract	*Average*	-0.23	-0.24	-0.23	-0.24
	Min	-0.27	-0.25	-0.28	-0.25
	Max	-0.18	-0.22	-0.18	-0.22
Local	*Average*	2.72	2.74	2.69	2.74
	Min	2.35	2.65	2.31	2.61
	Max	3.02	2.84	2.92	2.91
Wknown	*Average*	1.99	2.01	1.96	2.00
	Min	1.79	1.95	1.71	1.92
	Max	2.19	2.09	2.12	2.11
Tod	*Average*	-9.61	-9.78	-9.59	-9.76
	Min	-10.21	-9.97	-10.16	-9.98
	Max	-9.13	-9.60	-9.07	-9.57
Seasonal	*Average*	-9.71	-9.87	-9.71	-9.86
	Min	-9.97	-9.96	-10.01	-9.99
	Max	-9.39	-9.77	-9.41	-9.69

[a] The results are shown for 100 and 2000 draws. The models were also computed for 200, 500, and 1000 draws. The results were systematic for increases in the number of draws. The full tables are available in the appendix.

[b] The values in parenthesis are the standard deviations of the estimated parameters over all 120 possible ordering combinations

Table 6. Willingness-to-Pay estimates of the proposed ordering assuming no-correlation for different number of Halton and random draws.

Variable	Halton		Random		Halton High Primes [a]	
	100 [b]	2000	100	2000	100	2000
Contract	-0.24	-0.24	-0.18	-0.25	-0.22	-0.25
Local	2.37	2.49	2.44	2.45	2.46	2.51
Wknown	1.73	1.76	1.71	1.75	1.69	1.78
Tod	-9.57	-9.80	-9.75	-9.71	-9.59	-9.75
Seasonal	-9.67	-9.98	-9.86	-9.97	-9.73	-9.89

(a) Halton High Primes are used as a proxy to represent simulation noise in high integrating dimensions (>10) which use high prime numbers (31, 37, 41, 43, 47)

(b) The results are shown for 100 and 2000 draws. The models were also computed for 200, 500, and 1000 draws. The results were systematic for increases in the number of draws. The full tables are available in the appendix.

Panel A Panel B

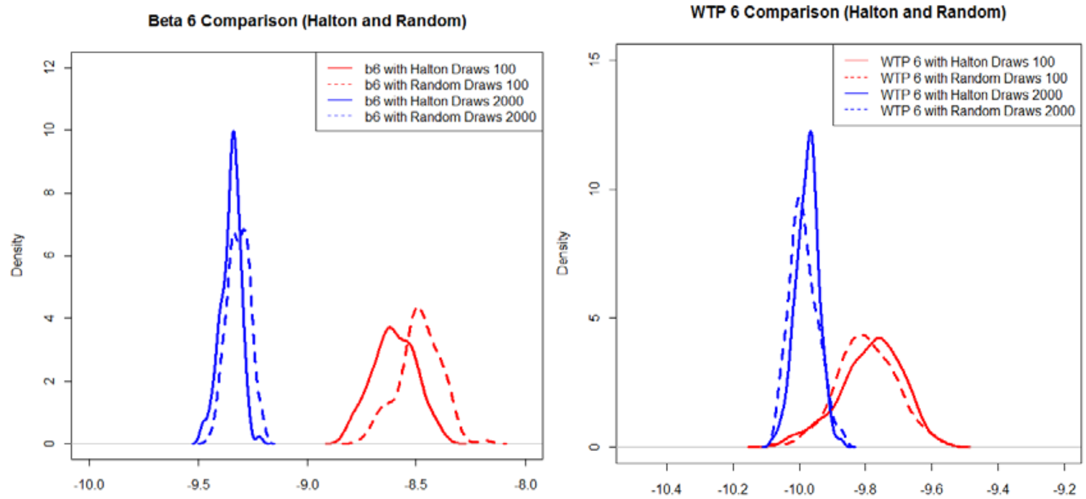

Figure 1. Distribution of Coefficient Estimates and Willingness-To-Pay for Seasonal Pricing for All 120 Possible Orderings.

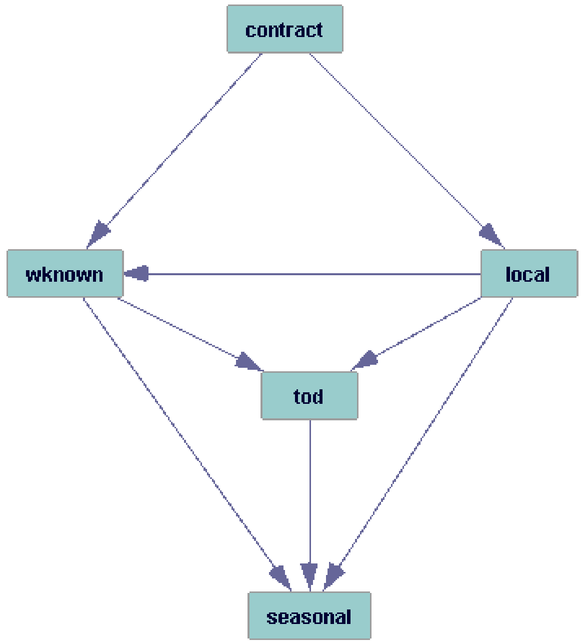

Figure 2. Directed Acyclic Graphs of Causality of the Observed Random Variables.