5-1-2009

A Comparison of Maximum Likelihood and Expected A Posteriori Estimation for Polychoric Correlation Using Monte Carlo Simulation

Jinsong Chen
The George Washington University, cjs@gwmail.gwu.edu

Jaehwa Choi
The George Washington University, jaehwa.choi@yahoo.com

A Comparison of Maximum Likelihood and Expected A Posteriori Estimation for Polychoric Correlation Using Monte Carlo Simulation

Jinsong Chen   Jaehwa Choi
The George Washington University

This study aims to compare the maximum likelihood (ML) and expected a posterior (EAP) estimation for polychoric correlation (PCC) under diverse conditions, especially when considering a sample size. As the ML is the classical solution to estimate PCC, the EAP is a new method based on Bayes' theorem. Different types of prior distributions are also adapted to investigate the sensitivity of prior distribution onto the PCC estimate for the EAP case. The Monte Carlo simulation is used for this comparison by a specialized program code in MATLAB.

Key words: Polychoric correlation, maximum likelihood, expected a posterior.

Introduction

It is fairly common that observed variables are measured using ordinal scales, which represent categorizations of underlying constructs that are continuous. This scenario is especially relevant in psychological and educational measurement. As an estimate of the relation between the two continuous constructs underlying two such ordinal variables, the polychoric correlation (PCC) has been widely employed. For instance, PCC has been used in many confirmatory factor analysis (CFA) or structural equation model (SEM) scenarios recently (e.g., Flora, 2002; Flora & Curran, 2004; Rigdon & Ferguson, 1991). The estimation of PCC has been conducted using maximum likelihood (ML) methods (e.g., Olsson, 1979), which can be accomplished using several popular statistical applications such as PRELIS (Jöreskog, 2002-2005) or SAS PROC FREQ (SAS Institute Inc., 2004).

Jinsong Chen is a doctoral candidate in the Graduate School of Education and Human Development. Email: cjs@gwmail.gwu.edu. Jaehwa Choi is an Assistant Professor of Educational Research in the Graduate School of Education and Human Development. Email: jaehwa.choi@yahoo.com.

Regarding the ML estimation of PCC, research showed that it: 1) produces an unbiased estimate of the correlation between the original bivariate normal variables (Babakus & Ferguson, 1988; Olsson, 1979); 2) outperforms Pearson's product-moment correlation (PPMC), Spearman's rho, and Kendall's tau-b for ordinal data (Babakus & Ferguson, 1988); and 3) is rather robust to modest violation of the underlying normality assumptions (Quiroga, 1992).

Even though estimating PCC with the ML method has been quite satisfactory as stated above, empirical or simulation results from previous research are based on relatively large sample sizes. For instance, the sample size was 500 for Olsson (1979), and 200 or above for Quiroga (1992). In many situations however, the sample size could be much smaller (e.g., less than 100), and the performance of the ML estimator has not been studied yet in the case of smaller sample sizes. Furthermore, due to the properties of numerical procedure of ML (i.e., iterative hill-climbing method using gradients of the target function), the ML estimation method for PCC also has several disadvantages such as, local maxima and non-converged solution.

Recently, expected a posteriori (EAP) estimation for PCC was introduced (Choi, Chen, & Kim, in press). As the EAP method is based on Bayes' theorem (Bock & Aitken, 1981), the

estimation of PCC can incorporate prior information regarding the correlation. The EAP method has been spotlighted in social and behavioral methodologies, such as Item Response Theory (IRT) models (e.g., Mislevy & Stocking, 1989). Also, this estimator has been compared with the ML method in IRT models (see Chen, Hou, & Dodd, 1998 for a summary). Because both PCC and EAP are becoming increasingly popular in social science research, understanding the behavior of the newly developed EAP estimator for PCC from a systematic comparison with the ML estimator would be beneficial. In this article, a methodological framework of both ML and EAP estimators will be introduced and the performance of the two estimators under various conditions will be compared, especially in the case of small sample size, using a Monte Carlo simulation study.

Polychoric Correlation and the Maximum Likelihood (ML) Estimation

Traditionally, ML has been the only estimator for estimating PCC, and the procedures are summarized here (see Olsson, 1979 for details). Two ordinal variables are observed with r and s possible categories. Given that the two corresponding continuous latent constructs follow the crucial assumption of bivariate normal distribution, the log-likelihood function of any sample is:

$$\ln L = \ln C + \sum_{i=1}^{s} \sum_{j=1}^{r} n_{ij} \ln \pi_{ij} \qquad (1.1)$$

where C is a constant and π_{ij} is the probability that a given observation falls into the contingency table cell (i, j) between two ordinal variables,

$$\pi_{ij} =$$
$$\Phi_2(a_i, b_j) - \Phi_2(a_{i-1}, b_j) - \Phi_2(a_i, b_{j-1}) + \Phi_2(a_{i-1}, b_{j-1})$$
$$(1.2)$$

where a and b are the threshold parameters for the categories $i = 1, 2, \ldots, s$ and $j = 1, 2, \ldots, r$, with $a_0 = b_0 = -\infty$ and $a_s = b_r = +\infty$, and Φ_2 is

the bivariate standard normal cumulative density function (CDF) with correlation ρ.

The threshold and correlation parameters can be estimated by: 1) taking partial derivatives of the log-likelihood function with respect to the parameters (thresholds and correlation), 2) setting these equations equal to zero, and 3) solving these equations for the parameters of interest using the numerical iterative procedure such as the Newton-Raphson method (Olsson, 1979). This method attempts to estimate all parameters of interest simultaneously, and was referred to as the "full ML method" or the "one-step ML method" by Olsson (1979).

Olsson also presented the two-step ML method for PCC, which estimates threshold values first:

$$a_i = \Phi_1^{-1}(P_i.) \text{ and } b_j = \Phi_1^{-1}(P._j) \qquad (1)$$

where P_{ij} is the observed proportion in cell (i, j), $P_i.$ and P_j are observed cumulative marginal proportions of the contingency table, and Φ_1 is the univariate normal CDF.

These threshold values are then substituted into the log-likelihood function, Equation (1), and the correlation parameter is estimated similar to the one-step ML method illustrated above. Olsson (1979) further showed that the difference of estimation between the one-step and two-step ML methods is negligible. Therefore, the two-step method is used in this study for the purpose of computational convenience.

Several issues of the ML methods are worthwhile to be noted here. As mentioned earlier, ML methods are iteratively searching the maximum of the log-likelihood function using the gradients (the first and second derivative of the log-likelihood function). Therefore, in general, ML estimators present the following disadvantages: 1) it is possible to get a non-converged solution; 2) there is no guarantee of getting the global maximum; 3) consequently, the ML estimates depend on a starting value; 4) above disadvantages tend to get worse as sample sizes decrease. Because it is very common for one to analyze small sample sizes (e.g., less than 100) in social and behavioral applied research,

these disadvantages of ML estimation have occasionally frustrated researchers who want to estimate PCC over the last several decades.

Expected A Posteriori Estimation

The EAP estimation method for PCC proposed by Choi et al. (in press) also adopts the assumption of bivariate underlying normal distribution, and uses the same procedure, Equation (3), in the above two-step ML method to estimate threshold values. However, when it estimates PCC, it follows Bayes' theorem: posterior distribution $\propto$ likelihood function $\times$ prior distribution. In other words, subjective belief about what the true correlation is likely to be can be incorporated into the estimation procedure through the prior distribution. Here is a brief development of the EAP method for PCC (more details of the EAP estimator of IRT model are available in Bock & Aitken, 1981):

$$\Pr(\rho | x) = \frac{\Pr(\rho)\Pr(x | \rho)}{\int_{-1}^{1} \Pr(\rho)\Pr(x | \rho)d\rho} \qquad (2)$$

where $\Pr(\rho | x)$ is the posterior distribution of given x which is frequency data of two variables, $\Pr(\rho)$ is a prior distribution of ρ, and $\Pr(x | \rho)$ is the same likelihood function L in the ML method. Then, the EAP (i.e., the mean of the posterior distribution) estimate of PCC can be simply expressed as:

$$\rho_{EAP} = E[\Pr(\rho | x)] = \int_{-1}^{1} \rho \Pr(\rho | x)d\rho. \quad (3)$$

For the purpose of numerical computation of the integration, the above two equations can be re-expressed as:

$$\Pr(\rho | x) = \frac{\Pr(\rho)\Pr(x | \rho)}{\sum_{i=1}^{k} \Pr(\rho_i)\Pr(x | \rho_i)} \qquad (4)$$

and

$$\hat{\rho}_{EAP} = \sum_{i=1}^{k} \rho_i \Pr(\rho_i | x) \qquad (5)$$

where k is the number of equally spaced quadrature points from -1 to +1.

As described in the above development, the EAP method is based on both Bayes' theorem and the non-iterative numerical integration. Consequently, the EAP method has the following advantages over the ML method: 1) there is no non-convergence issue (i.e., the estimates always exist); 2) there is no risk of a local maxima problem; 3) the estimates do not depend on a starting value; 4) the capability of including the a priori knowledge/belief on the parameter into the estimation process using prior distribution.

Methodology

Beyond the methodological advantages of the EAP over the ML illustrated above, it would be useful to investigate the empirical behavior of the two estimators over various conditions especially for a small sample sizes. In this study, a Monte Carlo simulation was used to examine the effect of sample size, population correlation magnitude, and number of categories on the PCC estimation for both EAP and ML estimators. The procedures can be summarized in the following way: (1) bivariate normal data was randomly generated from specific population correlation magnitude (ρ) and sample size (n); (2) the generated interval data was categorized over number of categories based on the threshold scheme and; (3) PCC was estimated by the EAP and ML estimation methods; and (4) the above procedures were repeated 1,000 times (i.e., iteration number in = 1,000).

Data Generation

The sample size variable was n = 30, 50, 100, and 500 observations. These numbers were chosen to reflect from small to moderate sample size that might be commonly encountered in the social sciences. The population correlation variable was chosen with ρ = 0, 0.1, 0.3, 0.5, and 0.7 magnitudes, ranging from null to moderate high.

The categorization rule (threshold scheme) used in this simulation study was the normal category option, which was also called the equal category width option within the range from -3 to 3 in standard normal distribution (Bollen & Barb, 1981). Therefore, the

distribution of categorized data gets closer to normal as the number of categories is increased. The number of category for both ordinal variables was $r = s = 2, 3, 5,$ and 7.

EAPPCC, a MATLAB subroutine (Choi et al., in press) was adopted for both EAP and ML estimators. Also, entire Monte Carlo simulation was implemented by a specialized program code in MATLAB (The MathWorks Inc., 2007), and the MVNRND function in MATLAB was used to generate bivariate normal data with specified population correlation (ρ) magnitude and sample size (n).

Estimation Options

In this study, the following estimators were considered: the ML, the EAP with Uniform ($-1, 1$) prior (EAPU), the EAP with Normal (0, 1/3) prior (EAPN), the EAP with Beta (5, 5) prior (EAPB5), and the EAP with Beta (1.5, 1.5) prior (EAPB1.5). The latter three were compared for prior sensitivity onto estimates.

Note that the range of Beta and Normal prior distribution have been adjusted to [-1, 1] for the purpose of constructing the appropriate priors for the correlation. As shown in Figure 1, the shape of Beta (5, 5) is very similar to that of Normal (0, 1/3). From the perspective of informativeness of prior, above two priors are more informative than Beta (1.5, 1.5) prior whereas the uniform distribution is least informative, specifically non-informative. Note that for comparing the performance of ML and EAP estimator, the EAPU (EAP with Uniform prior) was compared to the ML.

A two-step ML method was employed for the ML method and the EAP method adopted 100 quadrature points ($k = 100$). Chen et al. (1998) showed that any quadrature points of 20 or above were substantially the same. Therefore, 100 quadrature points would be a fair balance considering both accuracy requirement and computational load in this study.

Figure 1: Probability density functions for four prior distributions: Uniform (-1, 1), Beta (1.5, 1.5), Beta (5, 5), Normal (0, 1/3)
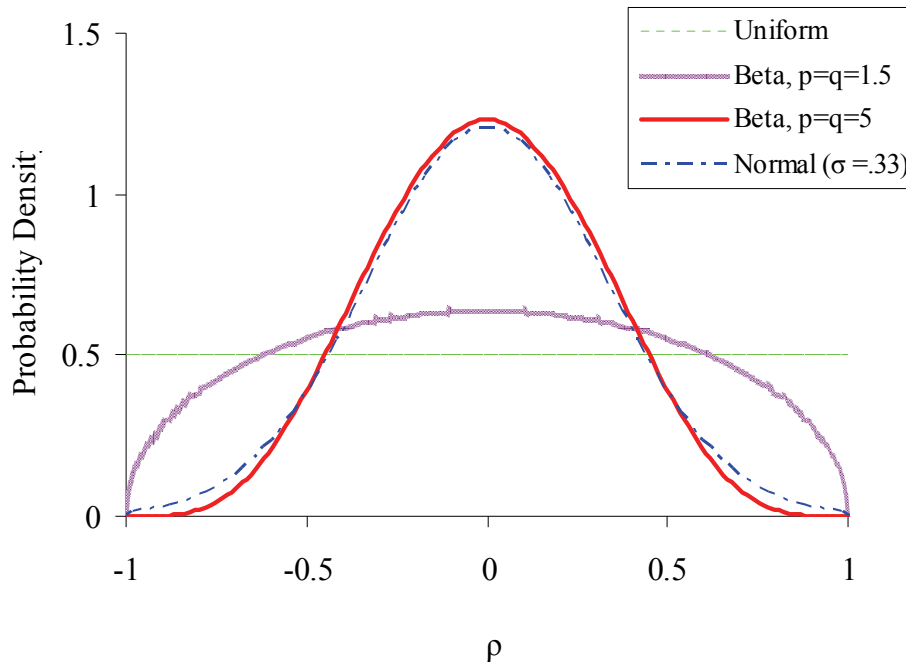
First, the convergence of the ML estimator was examined using the convergence rate (*CR*):

$$CR = pcin / in \times 100 \%\qquad(6)$$

where *pcin* was the properly converged iteration number, and *in* was the total iteration number attempted (i.e., *in* = 1,000).

Second, for evaluating the bias of the ML and EAPU (EAP with uniform distribution as prior) estimates, mean relative bias (*MRB*) was employed as the major statistics, and its general form was (Bandalos, 2006):

$$MRB = \sum_{i=1}^{pcin}\left[(\hat{\rho}_i - \rho)/\rho\right]/pcin\qquad(7)$$

where $\hat{\rho}_i$ was the *i*th iteration PCC estimate. In case of $\rho = 0$, mean bias (*MB*) was used instead of *MRB*, to avoid the issue of dividing by zero:

$$MB = \sum_{i=1}^{pcin}\left[\hat{\rho}_i - \rho\right]/pcin.\qquad(8)$$

In case average *MRB* values over different ρ magnitudes were needed (e.g., Figures 2 and 4), the *MB* values of $\rho = 0$ were excluded from averaging *MRB* values. Additionally, mean value (*M*) of estimates were also presented and analyzed:

$$M = \sum_{i=1}^{pcin}\hat{\rho}_i / pcin.\qquad(9)$$

Third, for evaluating the variability of the ML and EAPU estimates, root mean squared error (*RMSE*) and standard deviation of mean values (*SD*) were examined with the following definitions:

$$RMSE = \left[\sum_{i=1}^{pcin}(\hat{\rho}_i - \rho)^2 / pcin\right]^{1/2},\quad(10)$$

and

$$SD = \left[\sum_{i=1}^{pcin}(\hat{\rho}_i - M)^2 / pcin\right]^{1/2}.\quad(11)$$

MRB and *SD* were used for examining the prior sensitivity of EAP estimators.

Results

Convergence Rate for the ML Estimator

In this study, the convergence of both EAP and ML estimators was assessed; as shown in previous studies (e.g., Flora & Curran, 2004), not all iterations were converges for the ML estimator in this study as well. From the pilot study, it was found that the convergence rate was very low (< 30%) in most cases when a fixed initial value (0) was used. Therefore, PPMC was used with categorized data as the initial value for the ML method in this simulation study.

As indicated in Table 1, the average convergence rates were below 100% in all scenarios for the ML. Furthermore, as the sample size, number of categories, or ρ magnitude decrease, the rates tended to become worse. In contrast, as expected, EAP estimates could be obtained for all iterations for all conditions (i.e., *pcin* = 1,000 for the EAP).

Table 1: Convergence Rates of the ML Estimator

n	30	50	100	500	
%	98.1	98.7	99.1	99.4	
r = s	2	3	5	7	
%	97.5	99.1	99.5	99.4	
ρ	0	0.1	0.3	0.5	0.7
%	97.8	98.1	98.7	99.1	99.5

Note. Values were averaged over other conditions

Bias of ML and EAPU Estimates

Statistics regarding the bias of estimators (*M* and *MRB*) are presented in Table 2, and are also summarized and depicted in Figure 2.

First, for $\rho = 0$, the differences between the ML and EAP (specifically the EAPU) estimators were negligible, and *M* for both estimates were very close to zero (i.e., |*M*| < 0.01 for most cases).

341

Table 2: Simulation Results of the ML and EAPU Estimates

	ML				EAPU			
	M	*SD*	*MRB*[a]	*RMSE*	*M*	*SD*	*MRB*[a]	*RMSE*
n	$\rho = 0, r = s = 2$				$\rho = 0, r = s = 2$			
30	-0.013	0.295	-0.013	0.295	-0.011	0.249	-0.011	0.249
50	0.001	0.224	0.001	0.224	0.001	0.201	0.001	0.202
100	-0.003	0.161	-0.003	0.161	-0.003	0.152	-0.003	0.152
500	0.002	0.074	0.002	0.069	0.002	0.072	0.002	0.068
n	$\rho = 0, r = s = 5$				$\rho = 0, r = s = 5$			
30	0.007	0.231	0.007	0.231	0.006	0.202	0.006	0.202
50	0.007	0.165	0.007	0.165	0.007	0.150	0.007	0.150
100	0.000	0.111	0.000	0.110	0.000	0.105	0.000	0.105
500	0.004	0.050	0.004	0.051	0.004	0.050	0.004	0.050
n	$\rho = 0.1, r = s = 2$				$\rho = 0.1, r = s = 2$			
30	0.102	0.291	2.115	0.291	0.087	0.247	-13.291	0.247
50	0.096	0.224	-3.980	0.225	0.086	0.202	-13.581	0.203
100	0.091	0.150	-9.227	0.151	0.086	0.142	-14.056	0.143
500	0.103	0.073	3.334	0.070	0.102	0.071	1.924	0.069
n	$\rho = 0.1, r = s = 5$				$\rho = 0.1, r = s = 5$			
30	0.109	0.217	9.145	0.217	0.095	0.191	-4.611	0.191
50	0.097	0.163	-2.837	0.163	0.089	0.149	-11.277	0.150
100	0.102	0.112	1.488	0.112	0.096	0.107	-3.552	0.107
500	0.102	0.050	1.460	0.050	0.100	0.050	0.297	0.050
n	$\rho = 0.3, r = s = 2$				$\rho = 0.3, r = s = 2$			
30	0.282	0.265	-6.102	0.267	0.240	0.226	-20.096	0.234
50	0.300	0.205	0.080	0.205	0.271	0.186	-9.587	0.188
100	0.302	0.150	0.515	0.150	0.286	0.143	-4.631	0.143
500	0.299	0.069	-0.476	0.066	0.295	0.069	-1.761	0.065
n	$\rho = 0, r = s = 3$				$\rho = 0, r = s = 3$			
30	-0.016	0.275	-0.016	0.276	-0.014	0.236	-0.014	0.236
50	0.002	0.214	0.002	0.215	0.002	0.194	0.002	0.195
100	0.001	0.140	0.001	0.140	0.001	0.132	0.001	0.132
500	0.001	0.062	0.001	0.062	0.001	0.061	0.001	0.061
n	$\rho = 0, r = s = 7$				$\rho = 0, r = s = 7$			
30	0.002	0.208	0.002	0.208	0.002	0.182	0.002	0.182
50	-0.001	0.155	-0.001	0.155	-0.001	0.142	-0.001	0.141
100	0.003	0.113	0.003	0.113	0.003	0.108	0.003	0.107
500	0.002	0.048	0.002	0.048	0.002	0.048	0.002	0.048

[a] Mean bias (*MB*) was used instead of *MRB* in case of $\rho = 0$

Table 2: Simulation Results of the ML and EAPU Estimates (continued)

	ML				EAPU			
	M	*SD*	*MRB*[a]	*RMSE*	*M*	*SD*	*MRB*[a]	*RMSE*
n	$\rho = 0.1, r = s= 3$				$\rho = 0.1, r = s= 3$			
30	0.112	0.270	11.719	0.271	0.096	0.231	-4.081	0.231
50	0.104	0.203	3.799	0.201	0.094	0.184	-6.248	0.182
100	0.108	0.136	7.582	0.137	0.102	0.129	1.688	0.129
500	0.103	0.058	2.509	0.058	0.101	0.058	1.211	0.058
n	$\rho = 0.1, r = s= 7$				$\rho = 0.1, r = s= 7$			
30	0.082	0.192	-18.201	0.193	0.071	0.168	-28.592	0.171
50	0.104	0.147	3.537	0.147	0.095	0.135	-5.346	0.135
100	0.106	0.111	5.876	0.111	0.101	0.105	0.714	0.105
500	0.099	0.047	-0.563	0.047	0.098	0.047	-1.691	0.047
n	$\rho = 0.3, r = s= 3$				$\rho = 0.3, r = s= 3$			
30	0.323	0.255	7.648	0.256	0.280	0.222	-6.835	0.223
50	0.316	0.184	5.209	0.185	0.287	0.170	-4.226	0.170
100	0.299	0.132	-0.243	0.132	0.284	0.127	-5.286	0.127
500	0.297	0.055	-1.025	0.055	0.294	0.055	-2.140	0.055
n	$\rho = 0.3, r = s= 5$				$\rho = 0.3, r = s= 5$			
30	0.301	0.200	0.462	0.200	0.266	0.180	-11.369	0.183
50	0.297	0.151	-1.169	0.151	0.273	0.142	-9.041	0.144
100	0.304	0.104	1.263	0.103	0.290	0.100	-3.199	0.100
500	0.300	0.046	-0.006	0.046	0.297	0.046	-0.982	0.046
n	$\rho = 0.5, r = s= 2$				$\rho = 0.5, r = s= 2$			
30	0.492	0.235	-1.602	0.235	0.422	0.206	-15.541	0.220
50	0.488	0.176	-2.329	0.176	0.444	0.163	-11.237	0.173
100	0.499	0.128	-0.161	0.128	0.475	0.123	-4.947	0.125
500	0.501	0.058	0.236	0.056	0.495	0.058	-0.988	0.056
n	$\rho = 0.5, r = s= 5$				$\rho = 0.5, r = s= 5$			
30	0.515	0.172	3.003	0.172	0.463	0.164	-7.332	0.168
50	0.501	0.130	0.235	0.130	0.468	0.127	-6.339	0.131
100	0.508	0.088	1.693	0.088	0.491	0.088	-1.778	0.088
500	0.499	0.041	-0.164	0.041	0.496	0.041	-0.884	0.041
n	$\rho = 0.7, r = s= 2$				$\rho = 0.7, r = s= 2$			
30	0.685	0.192	-2.217	0.193	0.598	0.178	-14.637	0.205
50	0.692	0.142	-1.180	0.143	0.636	0.138	-9.079	0.151
100	0.698	0.099	-0.241	0.099	0.669	0.098	-4.400	0.102
500	0.700	0.044	-0.009	0.043	0.693	0.044	-0.945	0.044

[a] Mean bias (*MB*) was used instead of *MRB* in case of $\rho = 0$

Table 2: Simulation Results of the ML and EAPU Estimates (continued)

	ML				EAPU			
	M	*SD*	*MRB*[a]	*RMSE*	*M*	*SD*	*MRB*[a]	*RMSE*
n	$\rho = 0.7, r = s = 5$				$\rho = 0.7, r = s = 5$			
30	0.709	0.124	1.299	0.124	0.655	0.128	-6.458	0.136
50	0.711	0.092	1.622	0.092	0.679	0.095	-2.997	0.097
100	0.705	0.064	0.714	0.065	0.689	0.066	-1.559	0.067
500	0.701	0.028	0.119	0.028	0.698	0.029	-0.330	0.029
n	$\rho = 0.3, r = s = 7$				$\rho = 0.3, r = s = 7$			
30	0.310	0.187	3.472	0.188	0.275	0.170	-8.252	0.171
50	0.303	0.146	1.003	0.146	0.280	0.137	-6.707	0.138
100	0.301	0.102	0.353	0.102	0.288	0.099	-3.927	0.099
500	0.301	0.045	0.301	0.044	0.298	0.044	-0.644	0.044
n	$\rho = 0.5, r = s = 3$				$\rho = 0.5, r = s = 3$			
30	0.503	0.214	0.571	0.214	0.440	0.193	-12.068	0.202
50	0.507	0.162	1.445	0.162	0.466	0.154	-6.707	0.157
100	0.502	0.112	0.445	0.112	0.481	0.110	-3.861	0.112
500	0.500	0.051	0.049	0.051	0.496	0.051	-0.863	0.051
n	$\rho = 0.5, r = s = 7$				$\rho = 0.5, r = s = 7$			
30	0.513	0.156	2.530	0.157	0.464	0.151	-7.108	0.155
50	0.510	0.122	2.018	0.122	0.479	0.120	-4.156	0.122
100	0.502	0.083	0.341	0.083	0.486	0.083	-2.896	0.084
500	0.501	0.037	0.214	0.037	0.498	0.037	-0.456	0.037
n	$\rho = 0.7, r = s = 3$				$\rho = 0.7, r = s = 3$			
30	0.719	0.169	2.641	0.170	0.640	0.161	-8.573	0.172
50	0.706	0.122	0.791	0.122	0.659	0.121	-5.906	0.127
100	0.695	0.090	-0.665	0.090	0.672	0.090	-4.028	0.094
500	0.701	0.038	0.178	0.038	0.697	0.038	-0.491	0.039
n	$\rho = 0.7, r = s = 7$				$\rho = 0.7, r = s = 7$			
30	0.705	0.116	0.646	0.116	0.657	0.120	-6.143	0.128
50	0.703	0.089	0.366	0.089	0.674	0.093	-3.677	0.096
100	0.703	0.057	0.436	0.057	0.689	0.058	-1.542	0.059
500	0.700	0.026	0.057	0.026	0.698	0.026	-0.329	0.026

[a] Mean bias (*MB*) was used instead of *MRB* in case of $\rho = 0$

Figure 2: *MRB* across Different Sample Sizes, Correlation Magnitude, and Number of Categories
(Values averaged over other conditions; Mean Bias (*MB*) was Used in Case of $\rho = 0$)
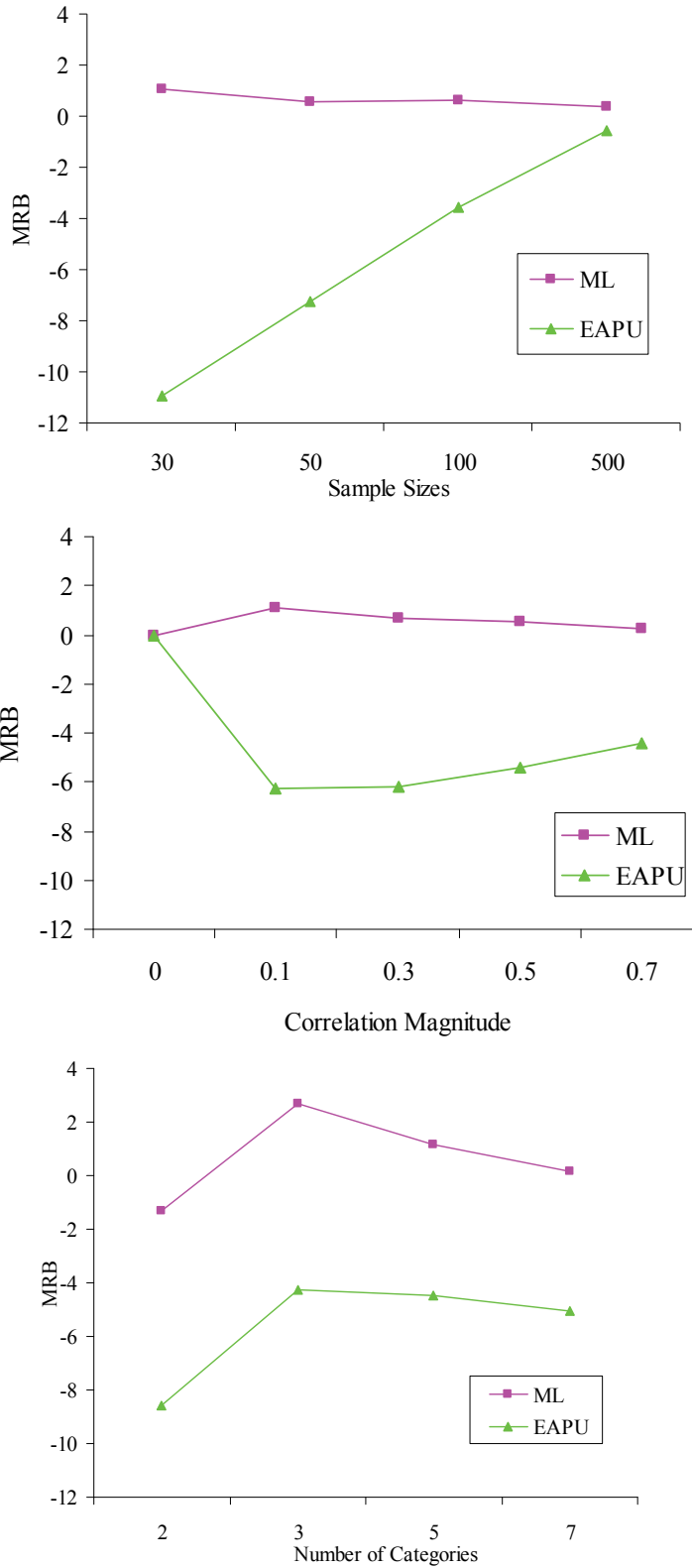
Second, as sample size increased, the difference between the two estimators disappeared, and both estimators performed better (i.e., smaller magnitude of *MRB*). Also, it seemed the EAP estimator was more sensitive to the change of sample size than the ML estimator according to Table 2 and Figure 2.

Third, for small sample sizes (*n* = 30 and 50) with non-zero ρ magnitude, *MRB* patterns of the two estimators substantially differed. For the ML case, *MRB* values can be largely positive or negative. Accordingly, the average *MRB* values (depicted in Figure 2) appeared to be smaller than most individual biases. Namely, the average *MRB* in Figure 2 were attenuated compared with the actual bias of the ML estimator in terms of *MRB*. For EAPU case, the *MRB* values tended to be largely negative for most cases, which supported the consistency between average and individual *MRB* values in Table 2 and Figure 2. These observations suggested the existence of systematic underestimation for the EAP estimator.

Fourth, the EAP estimator performed slightly better as ρ magnitude increased except in the null case. Also, the EAP estimator performed poorly for two categories as compared to higher numbers of categories, as shown in Figure 2. Again, the average *MRB* values for the ML in Figure 2 were less meaningful because they represent averages of the negative and positive values in each simulation cell *MRB* values.

Variability of ML and EAPU Estimates

Detailed *RMSE* and *SD* values of ML and EAP estimates are presented in Table 2, and are summarized and depicted in Figure 3.

First, in terms of *RMSE* and *SD*, the EAP outperformed the ML estimator in most cases. However, the differences in *RMSE* and *SD* among estimators are negligible for many cases.

Second, in small sample sizes (*n* = 30 and 50), the EAP estimator clearly outperformed the ML estimator (Figure 3). As sample size increased, the difference of *RMSE* values between two estimators disappeared, and the variability or fluctuation became increasingly

smaller, which suggested that both estimators were asymptotically efficient.

Third, a similar pattern was found over ρ magnitude. In small magnitude, the EAP estimator evidently outperformed the ML estimator. As the magnitude increased, the difference between the two estimators disappeared, and the variability of estimates decreased for both estimators. However, *RMSE* values were more sensitive to the change of sample size than that of ρ magnitude. This can be observed by comparing different charts in Figure 3.

Fourth, for number of categories, the EAP estimator also appeared to outperform the ML estimator in all cases. However, the differences in *RMSE* values between five and seven categories were very small for both estimators.

Prior Sensitivity for EAP with Different Prior Distributions

Statistics of EAP estimates with different prior distributions are presented in Table 3, and those are also summarized and depicted in Figures 4 through 5.

First, the EAP estimators whose prior distributions were more informative increasingly biased toward the mean of prior distribution. As shown in Table 3, the mean of estimates (*M*) for both EAPN and EAPB5 estimators, whose prior distributions are most informative (Figure 1), are very close to zero for most conditions. Also, *MRB* values for both EAPN and EAPB5 were extremely negatively biased (approximately -92) for all non zero population correlation cases. The *M* and *MRB* values for EAPB1.5, whose prior distribution is moderately informative in this study, are always between the above informative prior cases and the least informative prior case, EAPU. The *MRB* values in Figure 4 present essentially the same results.

Second, for estimators whose prior distributions were more informative, the variability of estimates in terms of *SD* was less. Both EAPN and EAPB5 cases showed the smallest *SD* values, whereas the *SD* values of the EAPB1.5 estimator were larger when compared with the above two estimators, but smaller when compared with the EAPU case (Figure 5). These results were also consistent with the

Figure 3: *RMSE* across Different Sample Sizes, Correlation Magnitude, and
Number of Categories
(Values were averaged over other conditions.)
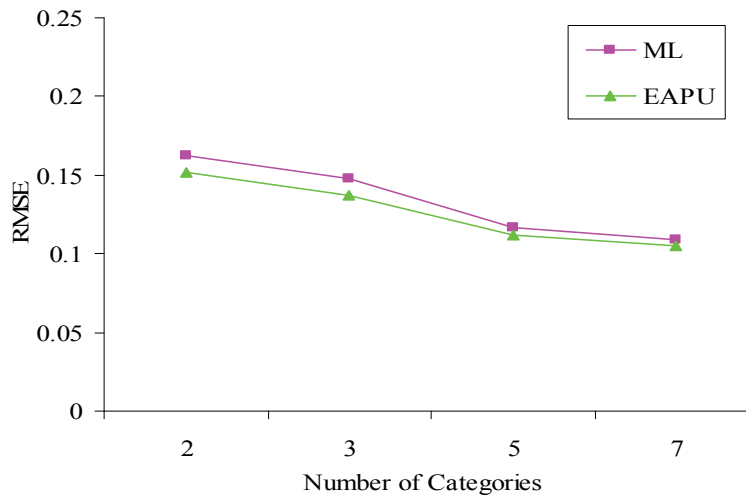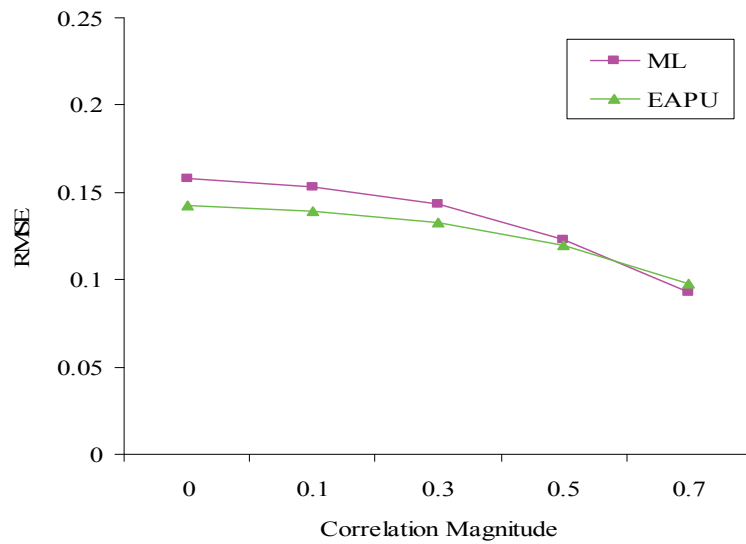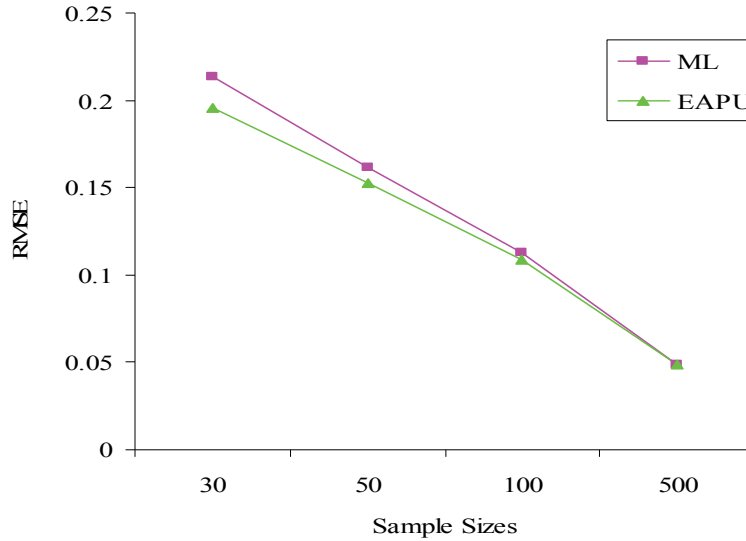
Table 3: Simulation Results of EAPU Estimates with Different Prior
Distributions

	EAPU				EAPB1.5			
	M	*SD*	*MRB*[a]	*RMSE*	*M*	*SD*	*MRB*[a]	*RMSE*
n	$\rho = 0$				$\rho = 0$			
30	-0.004	0.217	-0.004	0.217	-0.001	0.089	-0.001	0.089
50	0.002	0.172	0.002	0.172	0.001	0.069	0.001	0.069
100	0.000	0.124	0.000	0.124	0.000	0.049	0.000	0.049
500	0.002	0.058	0.002	0.057	0.001	0.022	0.001	0.022
n	$\rho = 0.1$				$\rho = 0.1$			
30	0.087	0.209	-12.644	0.210	0.036	0.085	-63.600	0.107
50	0.091	0.168	-9.113	0.168	0.038	0.067	-62.341	0.092
100	0.096	0.121	-3.801	0.121	0.039	0.048	-61.026	0.077
500	0.100	0.056	0.435	0.056	0.040	0.022	-60.496	0.065
n	$\rho = 0.3$				$\rho = 0.3$			
30	0.265	0.200	-11.638	0.203	0.112	0.082	-62.550	0.205
50	0.278	0.159	-7.390	0.160	0.115	0.065	-61.684	0.196
100	0.287	0.117	-4.261	0.117	0.116	0.047	-61.302	0.190
500	0.296	0.053	-1.382	0.052	0.118	0.022	-60.675	0.184
n	$\rho = 0.5$				$\rho = 0.5$			
30	0.447	0.179	-10.512	0.186	0.191	0.077	-61.865	0.319
50	0.464	0.141	-7.110	0.146	0.194	0.061	-61.179	0.312
100	0.483	0.101	-3.370	0.102	0.199	0.044	-60.226	0.304
500	0.496	0.047	-0.798	0.046	0.203	0.024	-59.395	0.300
n	$\rho = 0.7$				$\rho = 0.7$			
30	0.637	0.147	-8.967	0.160	0.275	0.071	-60.667	0.430
50	0.662	0.112	-5.415	0.118	0.284	0.055	-59.471	0.419
100	0.680	0.078	-2.882	0.081	0.289	0.039	-58.658	0.412
500	0.697	0.034	-0.524	0.034	0.297	0.023	-57.632	0.406

Note. Values were averaged across different number of categories

Table 3: Simulation Results of EAPU Estimates with Different Prior
Distributions (continued)

	EAPB5				EAPN			
	M	*SD*	*MRB*[a]	*RMSE*	*M*	*SD*	*MRB*[a]	*RMSE*
n	$\rho = 0$				$\rho = 0$			
30	-0.001	0.017	-0.001	0.017	-0.001	0.016	-0.001	0.016
50	0.000	0.013	0.000	0.013	0.000	0.012	0.000	0.012
100	0.000	0.010	0.000	0.010	0.000	0.009	0.000	0.009
500	0.000	0.004	0.000	0.004	0.000	0.004	0.000	0.004
n	$\rho = 0.1$				$\rho = 0.1$			
30	0.007	0.016	-92.914	0.095	0.006	0.015	-93.578	0.095
50	0.007	0.013	-92.636	0.094	0.007	0.012	-93.352	0.094
100	0.007	0.009	-92.400	0.093	0.007	0.009	-93.164	0.093
500	0.007	0.005	-92.328	0.092	0.007	0.004	-93.114	0.093
n	$\rho = 0.3$				$\rho = 0.3$			
30	0.022	0.016	-92.712	0.278	0.020	0.014	-93.393	0.280
50	0.022	0.012	-92.585	0.278	0.020	0.011	-93.306	0.280
100	0.022	0.009	-92.538	0.278	0.020	0.008	-93.286	0.280
500	0.023	0.005	-92.409	0.277	0.021	0.004	-93.185	0.280
n	$\rho = 0.5$				$\rho = 0.5$			
30	0.037	0.014	-92.688	0.464	0.033	0.013	-93.368	0.467
50	0.037	0.011	-92.603	0.463	0.033	0.010	-93.319	0.467
100	0.038	0.008	-92.453	0.463	0.034	0.007	-93.205	0.466
500	0.039	0.005	-92.293	0.462	0.035	0.004	-93.076	0.466
n	$\rho = 0.7$				$\rho = 0.7$			
30	0.052	0.013	-92.623	0.648	0.047	0.011	-93.304	0.653
50	0.053	0.009	-92.470	0.647	0.048	0.009	-93.192	0.652
100	0.053	0.007	-92.370	0.647	0.048	0.006	-93.123	0.652
500	0.054	0.004	-92.213	0.646	0.049	0.004	-92.998	0.651

Note. Values were averaged across different number of categories

Figure 4: *MRB* of the EAP Estimator across Different Sample Sizes and
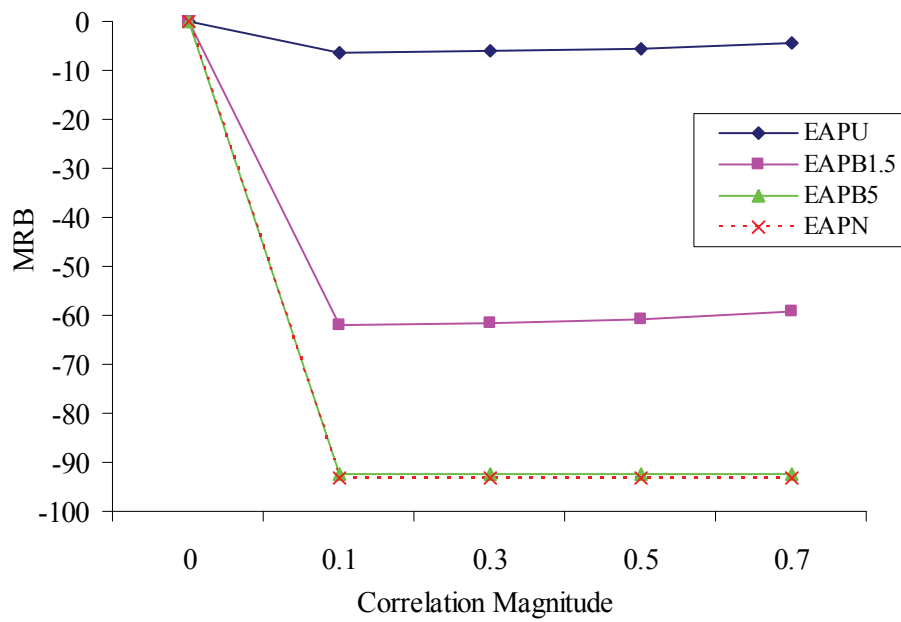Correlation Magnitude
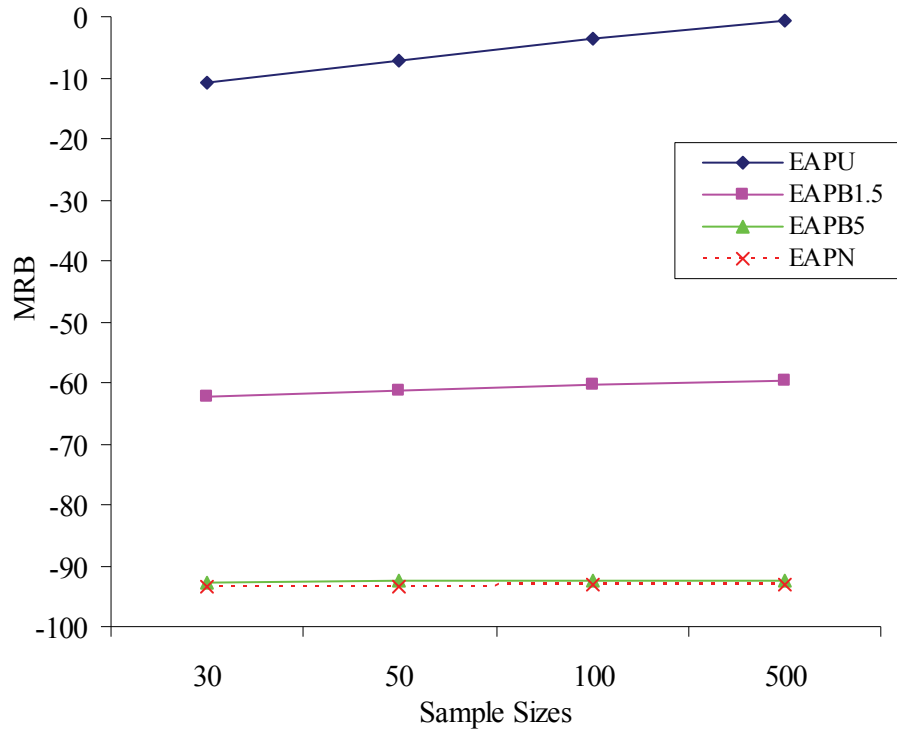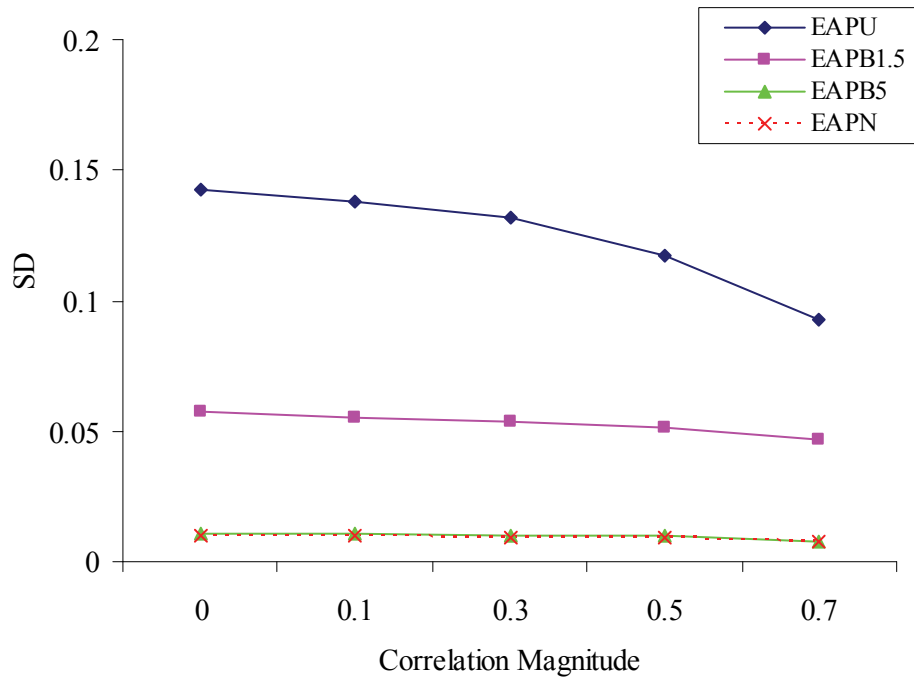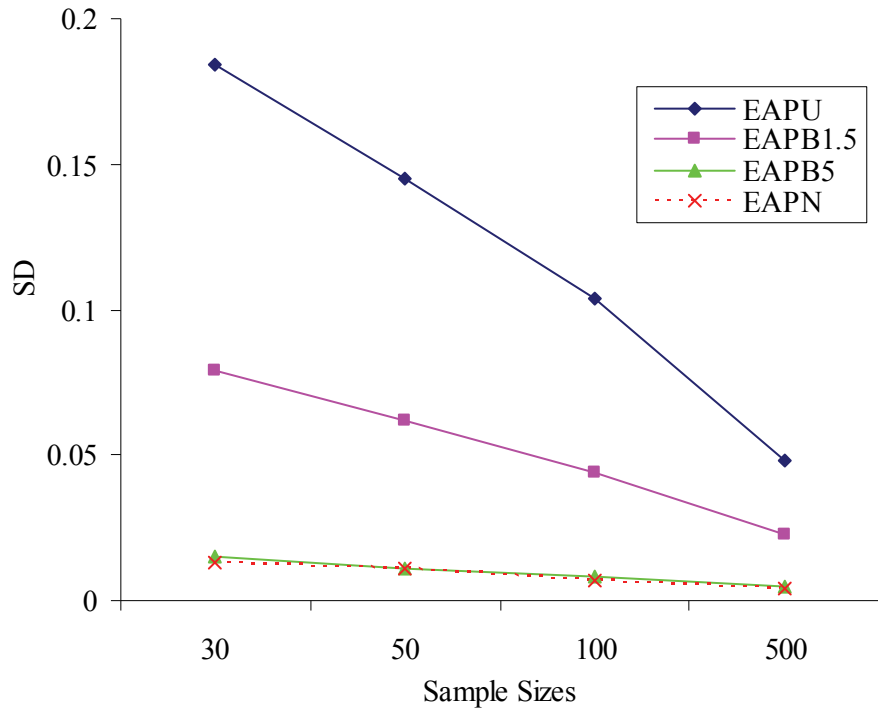(Values were averaged over other conditions)

Figure 5: *SD* of the EAP Estimator across Different Sample Sizes and
Correlation Magnitude
(Values were averaged over other conditions)

informativeness order of different prior distributions (Figure 1).

Third, the performance of the EAP estimators in terms of biases (*MRB*) and variability (*SD*) got better as sample size or ρ increase. However, as the prior distribution got more informative, both *MRB* and *SD* are less sensitive to the change of sample size or ρ magnitude, whereas the EAPU estimates were most sensitive to those changes (Figures 4 through 5). Furthermore, for $\rho = 0$ magnitude, biases of all estimators disappeared.

Conclusion

This study attempted to evaluate and compare the behaviors of ML and EAP estimators for PCC focusing on small sample size cases. The convergence rate of the ML estimator improves when using the PPMC for categorized data as the initial value rather than a fixed number, 0. However, non-converged cases for the ML remain an issue. In contrast, for the EAPU estimator, there is no non-convergent issue. These results could prove promising and useful to the applied researcher who is planning to estimate PCC and perhaps already arrived at a non-converged solution from the ML estimator.

For small sample sizes, the ML estimator can substantially underestimate or overestimate the ρ magnitude, whereas the EAP estimator always underestimates the ρ magnitude as shown in Figure 6. This, shrinkage effect, is a well known problem of EAP, similar to other Bayesian estimates. Because of the nature of EAP, a weighted average of posterior distribution which is a function of prior, EAP estimates is generally biased toward the mean of the prior distribution. Because all priors used in this research have zero as a mean, EAP estimates were biased toward zero for all cases. Important issues to be addressed include 1) the shrinkage effect is more apparent as the ρ magnitude increased or sample size decreased; and 2) the shrinkage effect disappears when $\rho = 0$.

Although this shrinkage effect is obviously a negative aspect of the EAP estimator, the systematic underestimation pattern could be wisely utilized to arrive at a conservative estimate of the true value. As

shown in Figure 2 and Table 2, the true value is most likely 0% to 15% higher than the EAPU estimate. Applied researchers should note that the ML estimator cannot provide such information.

For the variability of estimates in terms of *RMSE* or *SD*, the EAP estimator generally outperforms the ML estimator. The results are more apparent when the sample size, ρ magnitude, or number of categories is small. Because the EAP estimator is a weighted average over a prior distribution, it tends to provide more stable estimates than the ML estimator in those conditions.

For the EAP estimator with small sample sizes, the use of two categories would not be recommended due to relatively large bias and variability of estimates. Meanwhile, the use of three, five, or seven categories does not provide much difference in *MRB* values. Although *RMSE* gets smaller as the number of categories increases, the difference in *RMSE* between two categories and higher numbers of categories is not as imminent as that of *MRB* (i.e., the average *RMSE* with two categories – the average *RMSE* with seven categories < 0.05). In sum, especially when sample sizes are small (50 or below) and number of categories are not large (five or below), the EAP estimator can be recommended over the ML because the EAP is free from the convergence issue and provides smaller estimate variability. Also, as the sample size increases, both the shrinkage effect and the difference between the ML and EAP estimators disappear.

EAP estimators with more informative prior distribution could result in stronger bias toward the mean of the prior distribution, and provide less variation of estimates in terms of *SD*. For EAP estimators with relatively informative prior distributions like EAPN or EAPB5, both bias and variation of estimates are insensitive to the change of the sample size or ρ magnitude. For applied researchers with a strong a priori belief of the true correlation, EAP can provide a more stable estimate. Researchers can also include such information into an estimation procedure by adopting an informative prior distribution in the EAP estimator. This advantage of the EAP is beneficial especially

when the research involves small sample sizes in practice.

The following suggestions for future research on the estimation of PCC are based on the findings from this study. First, other Bayesian estimate, e.g., maximum a posteriori (MAP) estimate might be considered for analysis. As it relies on the mode, rather than the mean, of the posterior distribution, MAP could have some advantages against either the EAP or ML estimates. Second, the violation of

normality under different settings should be considered. As it has been shown that modest violation to normality is not critical for the ML estimator under relatively large sample size (Flora & Curran, 2004; Olsson, 1979; Quiroga, 1992), the situations for small sample size or for other estimators are not fully understood yet. Third, as this study focused on the point estimate, investigations on the interval estimates over different estimators would be needed, and should be addressed in future research.

Figure 6: *M* across Different Sample Sizes
(Values were averaged over number of categories)
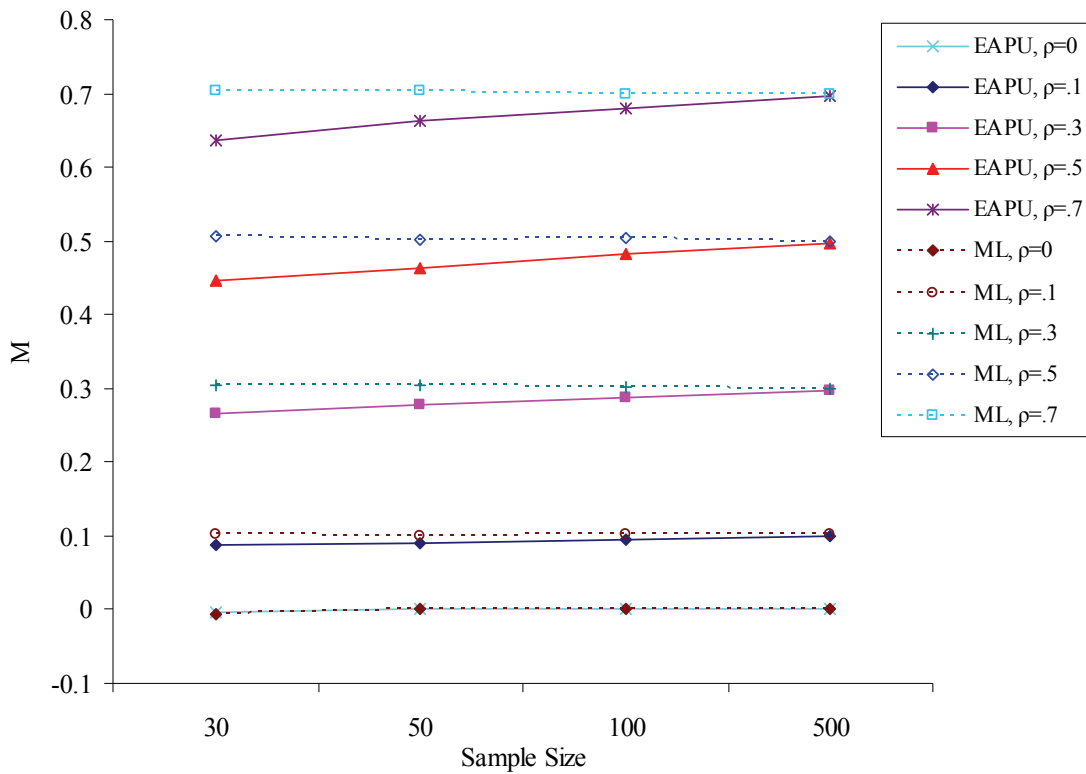
References

Babakus, E., & Ferguson, C. E. (1988). On choosing the appropriate measure of association when analyzing rating scale data. *Academy of Marketing Science*, *16*(1), 95-103.

Bandalos, D. L. (2006). The use of Monte Carlo studies in structural equation modeling research. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 385-427). Charlotte, NC: Information Age Publishing Inc.

Bock, R. D., & Aitken, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika*, *46*, 443-459.

Bollen, K. A., & Barb, K. H. (1981). Pearson's r and coarsely categorized measures. *American Sociological Review, 46*, 232-239.

Casella, G., & Berger, R. L. (2002). *Statistical inference* (*2nd Ed.*). Pacific Grove, CA: Duxbury.

Chen, S.-K., Hou, L., & Dodd, B. G. (1998). A comparison of maximum likelihood estimation and expected a posteriori estimation in CAT using the partial credit model. *Educational and Psychological Measurement*, *58*(4), 569-595.

Choi, J., Chen, J., & Kim, S. (in press). EAPPCC: A Matlab subroutine for estimating polychoric correlation matrices using an expected a posteriori estimation method. *Applied Psychological Measurement*.

Flora, D. B. (2002). *Evaluation of categorical variable methodology for confirmatory factor analysis with Likert-type data.* Unpublished Doctoral dissertation, University of North Carolina at Chapel Hill, Chapel Hill.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*(4), 466-491.

Jöreskog, K. G. (2002-2005). *Structural equation modeling with ordinal variables using LISREL*. Retrieved 10/07/2007, from http://www.ssicentral.com/lisrel/techdocs/ordinal.pdf.

Mislevy, R. J., & Stocking, M. L. (1989). A consumer's guide to LOGIST and BILOG. *Applied Psychological Measurement*, *13*(1), 57-75.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, *44*, 443-460.

Quiroga, A. M. (1992). *Studies of the polychoric correlation and other correlation measures for ordinal variables*. Unpublished Doctoral dissertation, Acta Universitatis Upsaliensis.

Rigdon, E. E., & Ferguson, C. E. (1991). The performance of the polychoric correlation coefficient and selected fitting functions in confirmatory factor analysis with ordinal data. *Journal of Marketing Research*, *28*(4), 491-497.

SAS Institute Inc. (2004). *SAS/STAT 9.1 user's guide*. Cary, NC: SAS Institute Inc.

The MathWorks Inc. (2007). Documentation for MathWorks Products, R2007b [Electronic Version]. Retrieved 10/07/2007 from http://www.mathworks.com/access/helpdesk/help/helpdesk.html.