Centre of Excellence for Science and Innovation Studies

CESIS Electronic Working Paper Series

Paper No. 259

On Liu Estimators for the Logit Regression Model

Kristofer Månsson
B. M. Golam Kibria

October 2011

The Royal Institute of technology
Centre of Excellence for Science and Innovation Studies (CESIS)
http://www.cesis.se

On Liu Estimators for the Logit Regression Model

By

Kristofer Månsson[1] B. M. Golam Kibria[2] and Ghazi Shukur[1,3]

[1] Department of Economics, Finance and Statistics, Jönköping University
P.O. Box 1026, SE- 551 11 Jönköping, Sweden

[2] Department of Mathematics and Statistics, Florida International University,
Miami, Florida, USA

[3] Department of Economics and Statistis, Linnaeus University,
Växjö Sweden.

Abstract

In innovation analysis the logit model used to be applied on available data when the dependent variables are dichotomous. Since most of the economic variables are correlated between each other practitioners often meet the problem of multicollinearity. This paper introduces a shrinkage estimator for the logit model which is a generalization of the estimator proposed by Liu (1993) for the linear regression. This new estimation method is suggested since the mean squared error (MSE) of the commonly used maximum likelihood (ML) method becomes inflated when the explanatory variables of the regression model are highly correlated. Using MSE, the optimal value of the shrinkage parameter is derived and some methods of estimating it are proposed. It is shown by means of Monte Carlo simulations that the estimated MSE and mean absolute error (MAE) are lower for the proposed Liu estimator than those of the ML in the presence of multicollinearity. Finally the benefit of the Liu estimator is shown in an empirical application where different economic factors are used to explain the probability that municipalities have net increase of inhabitants.

Key words: Estimation; MAE; MSE; Multicollinearity; Logit; Liu; Innovation analysis.

JEL Classification: C18; C35; C39

1. Introduction

Consider the situation when the dependent variable is $Be(\pi_i)$, where $\pi_i = \frac{\exp(x_i'\beta)}{1+\exp(x_i'\beta)}$ where x_i is the i*th* row of X which is an $n\times(p+1)$ data matrix with *p* explanatory variables, and β is a $(p+1)\times 1$ vector of coefficients. In this situation the parameters of the model should be estimated using the maximum likelihood (ML) method by applying the following iterative weighted least square (IWLS) algorithm:

$$\hat{\beta}_{ML} = \left(X'\hat{W}X\right)^{-1} X'\hat{W}\hat{z}, \qquad (1.1)$$

where $\hat{z}$ is a vector where the *i*th element equals $\hat{z}_i = \log(\hat{\pi}_i) + \frac{y_i - \hat{\pi}_i}{\hat{\pi}_i(1-\hat{\pi}_i)}$ and $\hat{W}$ is a diagonal matrix with i*th* diagonal element equals $(\hat{\pi}_i)(1-\hat{\pi}_i)$. This estimator approximately minimizes the weighted sum of squared error (WSSE). However, several sources of instability for the ML estimator exists. One may have the problem of separation where a linear combination of the regressors is perfectly predictive of the dependent variable. This problem discussed by Albert and Anderson (1984) lead to non-existence of the ML estimator. The authors also showed that in case of almost perfect separation the ML estimates are instable. Another source of instability which is the focus of this paper arises when the regressors are collinear. In that situation the weighted matrix of cross-products, $X'WX$, is ill-conditioned which leads to instability and high variance of the ML estimator.

Shrinkage estimator is a commonly applied solution to the general problem caused by multicollinearity. For the linear model a lot of research has been conducted and Hoerl and Kennard (1970) suggested the well-know ridge regression estimator. This estimator has then been extended to the logit model by Schaeffer et al. (1984) and further developments were made by Månsson and Shukur (2011) where some different new ridge parameters for logit ridge regression were suggested. However, the disadvantage of this method is that the estimated parameters are complicated non-linear functions of the ridge parameter *k* which can take on values between zero and infinity. Therefore, Liu (1993) suggested another estimator where the parameters obtained from this estimator has the benefit of being a linear function of the shrinkage parameter *d*. Due to this advantage over the ridge regression, the Liu estimator has been used by various researchers. Among them Akdeneiz and Kaciranlar (1995), Kaciranlar (2003) and Alheety and Kibria (2009) and very recently Kibria (2011) are notable. This estimator can also be extended to logit models. Now, by noting that the IWLS algorithm in equation (1.1) approximately minimizes the weighted sum of square error (WSSE), then one can apply the following estimator

$$\hat{\beta}_d = \left(X'\hat{W}X + I \right)^{-1} \left(X'\hat{W}X + dI \right) \hat{\beta}_{ML} = Z\hat{\beta}_{ML} . \tag{1.2}$$

The purpose of this paper is to apply the Liu estimator in order to solve the problems caused by multicollinearity. The Liu estimator is assumed to perform better than ML when the regressors are highly inter-correlated since $\hat{\beta}_{ML}$ is, on average, too long in that situation and $\hat{\beta}_d$ shrinks the length of the vector $\hat{\beta}_{ML}$. This paper will also suggest some methods of estimating the shrinkage parameter *d*. The performance of ML and the Liu estimator will be studied using Monte Carlo simulations where factors such as the number of regressors, the sample size and the degree of correlation are varied. In order to judge the performance of the estimator the mean squared error (MSE) and mean absolute error (MAE) are used. The result shows that the Liu estimator always outperforms ML in the presence of multicollinearity. The benefits of the Liu estimator will also be shown in an empirical application where different economic factors are used to explain the probability that municipalities have a net increase of inhabitants.

This paper is organized as follows: In Section 2, the statistical methodology is described. In section 3, the design of the experiment and a result discussion are provided. Then in section 4 an empirical application is provided. Finally, in Section 5, some concluding remarks are provided.

2. Statistical methodology

2. 1. The Statistical properties of the ML and Liu estimators

The Liu estimator for the logit model is a biased shrinkage estimator and a direct generalization of the one proposed for linear regression model by Liu (1993). The shrinkage parameter *d* may take on values between zero and one and when *d* equals to one then

$\hat{\beta}_d = \hat{\beta}_{ML}$. When d is less than one we have $\left\|\hat{\beta}_d\right\| \le \left\|\hat{\beta}_{ML}\right\|$. Since $\hat{\beta}_{ML}$ is, on average, too long in the presence of multicollinearity, $\hat{\beta}_d$ is assumed to perform better than $\hat{\beta}_{ML}$ in such situation. This may also be shown by studying the MSE properties of the two estimators. The MSE of the ML estimator equals:

$$MSE\left(\hat{\beta}_{ML}\right) = E\left(L_{ML}^2\right) = E\left(\hat{\beta}_{ML} - \beta\right)'\left(\hat{\beta}_{ML} - \beta\right) = tr\left(X'WX\right)^{-1} = \sum_{j=1}^{J} \frac{1}{\lambda_j} \quad (2.1)$$

where λ_j is the jth eigenvalue of the $X'WX$ matrix. When looking at the MSE it can easily be seen that it becomes inflated in the presence of multicollinearity since some eigenvalues will be small when $X'WX$ is ill-conditioned. On the other hand, the MSE of the Liu estimator is:

$$\begin{aligned} &MSE\left(\hat{\beta}_d\right) = E\left(L_d^2\right) = E\left(\hat{\beta}_d - \beta\right)'\left(\hat{\beta}_d - \beta\right) = \\ &E\left[\left(\hat{\beta}_{ML} - \beta\right)'Z'Z\left(\hat{\beta}_{ML} - \beta\right)\right] + \left(Z\beta - \beta\right)'\left(Z\beta - \beta\right) = \\ &tr\left[\left(\hat{\beta}_{ML} - \beta\right)'\left(\hat{\beta}_{ML} - \beta\right)Z'Z\right] + k^2\beta'\left(X'WX + kI\right)^{-2}\beta = \\ &\sum_{j=1}^{J} \frac{\left(\lambda_j + d\right)^2}{\lambda_j\left(\lambda_j + 1\right)^2} + \left(d-1\right)^2 \sum_{j=1}^{J} \frac{\alpha_j^2}{\left(\lambda_j + 1\right)^2} = \gamma_1\left(d\right) + \gamma_2\left(d\right) \end{aligned} \quad (2.2)$$

where α_j^2 is defined as the jth element of $\gamma\beta$ and γ is the eigenvector defined such that $X'WX = \gamma'\Lambda\gamma$ where Λ equals $diag\left(\lambda_j\right)$. For the Liu estimator one wants to find a value of d so that the decrease of the variance ($\gamma_1\left(d\right)$) is greater than the increase caused by adding the squared bias ($\gamma_2\left(d\right)$). In order to show that such a value of d less than one exists so that $MSE\left(\hat{\beta}_d\right) < MSE\left(\hat{\beta}_{ML}\right)$ we start taking the first derivative of equation (2.2) with respect to d:

$$g'\left(d\right) = 2\sum_{j=1}^{J} \frac{\lambda_j + d}{\lambda_j\left(\lambda_j + 1\right)^2} + 2\left(d-1\right)\sum_{j=1}^{J} \frac{\alpha_j^2}{\left(\lambda_j + 1\right)^2}$$

and then by inserting the value one in equation (2.3) we get:

$$g'\left(d\right) = 2\sum_{j=1}^{J} \frac{1}{\lambda_j\left(\lambda_j + 1\right)}, \quad (2.4)$$

which is greater than zero since $\lambda_j > 0$. Hence, there exists a value of *d* between zero and one so that $MSE(\beta_d) < MSE(\beta_{ML})$. Furthermore, the optimal value of any individual parameter d_j can be found by setting equation (2.4) to zero and solve for d_j. Then it may be shown that

$$d_j = \frac{\alpha_j^2 - 1}{\frac{1}{\lambda_j} + \alpha_j^2}, \tag{2.5}$$

corresponds to the optimal value of the shrinkage parameter. Hence, the optimal value of d_j is negative when α_j^2 is less than one and positive when it is greater than one. However, just as in Liu (1993) the shrinkage parameter will be limited to values between zero and one.

2.2 Estimating the shrinkage parameter

In order to estimate the optimal value of *d* in equation (2.5) several methods will be proposed. The idea behind these proposed estimators are obtained from the work of Hoerl and Kennard (1970), Kibria (2003) and Khalaf and Shukur (2005) where several different methods of estimating the shrinkage parameter for linear ridge regression have been proposed. As in those papers, the shrinkage parameter, d_j, will be estimated by a single value $\hat{d}$. The first estimator which is based on the work by Hoerl and Kennard (1970) is the following:

$$D1 = \max\left(0, \frac{\hat{\alpha}_{\max}^2 - 1}{\frac{1}{\hat{\lambda}_{\max}} + \hat{\alpha}_{\max}^2}\right),$$

where we define $\hat{\alpha}_{\max}^2$ and $\hat{\lambda}_{\max}$ to be the maximum element of $\hat{\alpha}_j^2$ and $X'\hat{W}X$ respectively. Replacing the values of the unknown parameters with the maximum value of the unbiased estimators is an idea taken from Hoerl and Kennard (1970). However, for the Liu estimator another maximum operator is also used that will ensure that the estimated value of the shrinkage parameter is not negative. Furthermore, the following estimators, which are based on the ideas in Kibria (2003), are proposed:

$$D2 = \max\left(0, median\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right), \quad D3 = \max\left(0, \frac{1}{p}\sum_j^J \frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right).$$

Using the average value and the median is very common when estimating the shrinkage parameter for the ridge regression and the D2 and D3 estimators have direct counterparts in equation (13) and (15) of Kibria (2003). Finally, the following estimators are proposed:

$$D4 = \max\left(0, \max\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right),\ D5 = \max\left(0, \min\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right).$$

For these estimators other quantiles than the median is used which was successfully applied by Khalaf and Shukur (2005).

3. The Monte Carlo simulation

3. 1 The Design of the Experiment

The main focus of this paper is to compare the MSE properties of the ML and Liu estimators when the regressors are highly intercorrelated. Hence, the core factor varied in the design of the experiment is the degree of correlation (ρ^2) between the regressors. Therefore, the following formula which enables us to vary the strength of the correlation is used to generate the explanatory variables:

$$x_{ij} = \left(1-\rho^2\right)^{(1/2)} z_{ij} + \rho z_{ip} \qquad i = 1,2,\ldots n,\ j = 1,2,\ldots p \qquad (3.1)$$

where z_{ij} pseudo-random numbers from the standard normal distribution. We consider four different values of ρ^2 corresponding to 0.75, 0.85, 0.95 and 0.99. The n observations for the dependent variable are obtained from the $Be(\pi_i)$ distribution where

$$\pi_i = \frac{\exp\left(x_i'\beta\right)}{1+\exp\left(x_i'\beta\right)}\ . \qquad (3.2)$$

The parameter values of β are chosen so that $\beta'\beta = 1$. We use 50, 75, 100 and 150 degrees of freedoms (*df*=*n*-*p*) and models consisting of two and four explanatory variables. The experiment is replicated 2000 times by generating new pseudo-random numbers. Then the MSE is calculated as follows:

$$MSE = \frac{\sum_{i=1}^{2000} \left(\hat{\beta} - \beta\right)' \left(\hat{\beta} - \beta\right)}{2000}. \qquad (3.3)$$

and the MAE as:

$$MAE = \frac{\sum_{i=1}^{2000} \left|\hat{\beta} - \beta\right|}{2000}. \qquad (3.4)$$

3.2 Result Discussion

The simulated MSE and MAE for all of the estimators for different *n* and ρ are presenetd in Tables 1 and 2 for p= 2 and 4 respectively. From the tables, we can see at a glance that the degree of correlation inflates the MSE and MAE. This increase is particularly large for ML and it is more severe when applying MSE as performance criteria instead of MAE. For the Liu estimators, the inflation of the MSE and MAE is less severe than for ML. However, there is big difference between the peformance of the Liu estimatos depending on which shrinkage parameter is applied. The least robust option among the different proposed methods of estimating the shrinkage parameter is the D4. The performance of the D1 to D3 estimators are almost equaivalent. However, the most robust option is the D5 estimator. This shrinakge parameter has always either the lowest value of both measures of performance or it is close to the estimator that minimizes the MSE and MAE. Moreover, one can see that as the number of explanatory variables increases the MSE and MAE increases. This increase is more sever for the ML than the Liu estimator and it is also larger if MSE is used to judge the performance of the estimator instead of MAE. Finally, when considering all of the results it is clear to see that increasing the sample size has a positive effect especially for ML. This is expected since $\hat{\beta}_{ML}$ is a consistent estimator.

4. Empirical Application

The different estimation methods will be illustrated using a dataset taken from the Statistics Sweden.[1] A logit regression model is estimated where the dependent variable is defined as follows:

[1] The homepage is www.scb.se

$$y_i = \begin{cases} 1 & \text{if the net population change is positive in municipality i} \\ & \text{0 otherwise} \end{cases}.$$

This dependent variable is explained by the following regressors, the number of unemployed people (x_1), the number of build appartments (x_2), the amount of bankrupt firms (x_3) and the population (x_4), respectively. We will estimate a logit model for the full sample and for the urban regions in Sweden.[2] The full sample consists of 290 observations and the subsample consists of 84 municipalitits. The bivariate correlation (for the full population) between the regressors can be found in Table 3:

Table 3: Correlation matrix

	x_1	x_2	x_3	x_4
x_1	1			
x_2	0.8854	1		
x_3	0.9426	0.9430	1	
x_4	0.9663	0.9010	0.9367	1

From Table 3 one can see that the bivariate correlations are high (all are greater than 0.88) and that we therefore might have a sever multicollinearity problem. The logit regression model is estimated in R using the IWLS algotirhm[3] and the Liu estimators is applied with the shrinkage parameter D5 since this is the one that minimizes the estimated MSE and MAE. In order to estimate the standard errors of the different paramaters bootstrap technique is applied. The results can be found in Table 4. We can see that the number of unemployed people and bankrupt firms have a negative impact while the other two variables have a positive impact on the probability of a municipality to have a net increase of inhabitants. This is expected since a higher value of unemployed people and bankrupted firms indicate a poor economic

[2] The urban regions are defined as the municipalitites belonging to the Functional analysis (FA) regions Stockholm, Göteborg and Malmö.

[3] We are using the function glm() in order to estimate the logit model which is part of the standard routines in R. However, any software will work fine since the Liu method does not require any changes of the existing routines of estimating the logit regression model. The Liu estimator only requires that one is able to extract the result the maximum likelihood estimators of the coefficients and the variance-covariance matrix which is defined as $\left(X'\hat{W}X\right)^{-1}$.

performance. The positive effect of the population variable indicates that more people are moving to urban regions. The estimated standard errors is decreased for all variables, but the most substantial reduction can be found for x_2 (i.e. the number of build appartments). For this variable the reduction of the estimated parameter is also substantial. This indicates that the multicollinearity problem leads to an estimated value that is lager than it should be. Hence, the positive impact of building new appartments is most likely exagerated when applying ML. When looking at the t-statstics one can see that these values using Liu method are larger than those for the ML which further shows the superiority of the Liu estimator since the p-values become lower. Once again it is for the variable x_2 the largest increase of the t-statstic can be found.

For the subsample compared with the full sample the sign of the variables x_1 and x_2 are changed. The positve impact of increasing the number of unemployed people may be due to the fact that many immigrants choose to settle down in the areas in the large cities with high unemployement rates. The negative impact of variable x_2 may be explanied by the fact that not enough appartments are constructed where a lot of the people are choosing to move. When looking at the standard errors one can see a much larger reduction of the standard errors for the subsample than the full sample. The reduction of the bootstrapped standard errors is especially remarkable for variables x_1 and x_2. The increase of the t-statistics is also larger for the subsmaple than the full sample. In this case the increas of x_4, may be noticed since this variable becomes statistcally significant when the Liu estimator is applied.

Table 4: The results from the logit regression analysis

	Full sample							
	ML				Liu			
	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
$\hat{\beta}$	-0.1990	3.3276	-0.4574	0.0149	-0.2098	1.1497	-0.3566	0.0150
$se(\hat{\beta})$	0.1256	1.3760	0.6174	0.0043	0.1060	0.2506	0.4509	0.0037
$t_{\hat{\beta}}$	-1.5844	2.4183	-0.7408	3.4651	-1.9792	4.5878	-0.7909	4.0541
	Large city regions							
	ML				Liu			
	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
$\hat{\beta}$	2.2983	-1.2418	-0.0409	0.0184	0.2426	-0.1333	-0.1175	0.0162
$se(\hat{\beta})$	15.904	31.149	0.9088	0.1078	0.3393	0.6864	0.1801	0.0062
$t_{\hat{\beta}}$	0.1445	-0.0399	-0.0450	0.1707	0.7150	-0.1942	-0.6524	2.6129

5. Some Concluding Remarks

In this paper a new Liu estimator for the logit model has been proposed. The MSE and MAE of this new estimator and the traditional ML method are calculated by using Monte Carlo simulations. In the design of the experiment, factors such as the degree of correlation, the sample size and the number of explanatory variables are varied. The result from the simulation study clearly showed that the MSE and MAE of the ML method become inflated in the presence of multicollinearity. This problem is particularly severe when the sample size is small and the correlation between two explanatory variables is high. The results from the Monte Carlo study also evident that the new Liu estimator is much more robust to increase of correlation and it has superior MSE and MAE properties over the ML in all of the evaluated situations. The best option to estimate the logit model in the presence of multicollinearity is to apply the Liu estimator together with the shrinkage parameter D5. The benefit of this method is clearly shown in an empirical application where one can see a substantial decrease of the standard errors and an increase of the t-statistics, especially for the subsample. We hope the findings of the paper will be useful for the practitioners.

Acknowledgements

This paper was written while Dr. Kibria was visiting Professor Ghazi Shukur at the Department of Economics and Statistics, Linnaeus University, and Department of Finance, Economics and Statistics, Jönköping International Business School, Sweden, during May-June 2011. He acknowledges the excellent research support provided by Linnaeus University and Jönköping International Business School, Sweden.

References

Albert, A. and Anderson, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* **71**, 1-10.

Akdeniz, F. and S. Kaciranlar (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. *Communications in Statistics-Theory and Methods*, **24**, 1789-1797.

Alheety, M. I. and B. M. G. Kibria (2009). On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors. *Surveys in Mathematics and its Applications*, **4**, 155-167.

Hoerl, A.E. and R.W. Kennard (1970). Ridge regression: biased estimation for non-orthogonal Problems. *Technometrics*, **12**, 55-67.

Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, **22**, 393-402.

Månsson, K. and Shukur, G. (2011), **"**On Ridge Parameters in Logistic Regression", *Communications in Statistics, Theory and Methods*, **40**, Issue 18, 3366-3381.

Kaciranlar, S. (2003). Liu estimator in the general linear regression model. *Journal of Applied Statistical Science*, **13**, 229-234.

Khalaf, G. and Shukur, G. (2005). Choosing Ridge Parameter for Regression Problems. *Communications in Statistics, Theory and Methods*, **34**, Issue 5, 1177-1182.

Kibria B.M.G. (2003). Performance of some new ridge regression estimators *Communications in Statistics, Theory and Methods*, **32**, 419-435.

Kibria, B. M. G. (2011). On Some Liu and Ridge Type Estimators and their Properties Under the Ill-conditioned Gaussian Linear Regression Model. To Appear in *Journal of Statistical Computation and Simulation.*

Schaefer, R.L., Roi, L. D. and Wolfe, R. A. (1984). A ridge logistic estimator. *Communications in Statistics- Theory and Methods*, **13**, 99-113.

Muniz, G. and Kibria, B. M. G. (2009). On some ridge regression estimators: An Empirical Comparison. *Communications in Statistics-Simulation and Computation* 38:621-630.

Table 1: The simulated MSE and MAE for different n and ρ^2 and p=2

	Estimated MSE						Estimated MAE					
	ML	D1	D2	D3	D4	D5	ML	D1	D2	D3	D4	D5
ρ^2=0.75												
50	1.041	0.452	0.341	0.341	0.478	0.301	1.087	0.695	0.629	0.629	0.711	0.605
75	0.639	0.327	0.286	0.286	0.338	0.270	0.865	0.612	0.582	0.582	0.618	0.572
100	0.442	0.258	0.237	0.237	0.262	0.230	0.725	0.551	0.534	0.534	0.554	0.530
150	0.280	0.189	0.181	0.181	0.190	0.179	0.582	0.477	0.470	0.470	0.478	0.469
200	0.205	0.150	0.147	0.147	0.150	0.147	0.498	0.428	0.426	0.426	0.429	0.426
ρ^2=0.85												
50	1.913	0.748	0.575	0.575	0.882	0.443	1.454	0.824	0.735	0.735	0.881	0.669
75	1.003	0.420	0.352	0.352	0.453	0.312	1.104	0.685	0.640	0.640	0.706	0.615
100	0.810	0.383	0.337	0.337	0.400	0.312	0.978	0.654	0.623	0.623	0.664	0.609
150	0.465	0.256	0.240	0.240	0.261	0.233	0.749	0.551	0.539	0.539	0.554	0.535
200	0.350	0.212	0.205	0.205	0.213	0.204	0.652	0.509	0.504	0.504	0.510	0.503
ρ^2=0.95												
50	5.453	1.785	1.433	1.433	2.803	0.815	2.525	1.177	1.075	1.075	1.497	0.787
75	3.057	0.980	0.806	0.806	1.352	0.541	1.940	0.928	0.852	0.852	1.085	0.703
100	2.530	0.900	0.721	0.721	1.152	0.509	1.753	0.892	0.819	0.819	0.997	0.709
150	1.438	0.522	0.456	0.456	0.601	0.379	1.315	0.708	0.672	0.672	0.749	0.628
200	1.093	0.407	0.363	0.363	0.443	0.326	1.165	0.669	0.643	0.643	0.691	0.620
ρ^2=0.99												
50	27.40	9.476	8.969	8.969	18.441	4.505	5.636	2.491	2.573	2.573	3.975	1.476
75	15.38	4.934	4.767	4.767	9.245	2.550	4.363	1.863	1.885	1.885	2.815	1.157
100	12.14	4.014	3.587	3.587	7.219	1.860	3.875	1.685	1.644	1.644	2.472	1.028
150	7.660	2.389	2.097	2.097	4.172	1.098	3.114	1.325	1.275	1.275	1.851	0.853
200	5.391	1.620	1.414	1.414	2.616	0.833	2.608	1.109	1.047	1.047	1.454	0.766

Table 2: The simulated MSE and MAE for different n and ρ^2 and p=4

	Estimated MSE						Estimated MAE					
	ML	D1	D2	D3	D4	D5	ML	D1	D2	D3	D4	D5
ρ^2=0.75												
50	4.233	1.854	1.020	0.877	2.428	0.625	3.062	1.936	1.463	1.391	2.136	1.259
75	2.190	1.098	0.699	0.653	1.204	0.624	2.287	1.569	1.310	1.283	1.629	1.265
100	1.524	0.844	0.621	0.601	0.889	0.591	1.914	1.393	1.243	1.232	1.421	1.226
150	0.936	0.583	0.505	0.503	0.594	0.501	1.525	1.193	1.134	1.133	1.201	1.132
200	0.659	0.447	0.415	0.415	0.450	0.414	1.279	1.050	1.024	1.025	1.053	1.024
ρ^2=0.85												
50	7.635	3.164	1.964	1.601	4.615	0.761	4.039	2.367	1.737	1.610	2.794	1.275
75	3.975	1.790	1.004	0.871	2.207	0.687	3.075	1.940	1.487	1.419	2.126	1.320
100	2.615	1.251	0.774	0.719	1.416	0.670	2.509	1.662	1.364	1.334	1.749	1.306
150	1.574	0.847	0.631	0.615	0.892	0.607	1.970	1.404	1.257	1.249	1.433	1.245
200	1.187	0.694	0.570	0.564	0.715	0.561	1.716	1.292	1.204	1.200	1.306	1.199
ρ^2=0.95												
50	26.48	9.661	7.841	6.720	18.98	1.977	7.237	3.733	2.864	2.660	5.339	1.307
75	12.86	4.886	3.222	2.558	7.988	0.872	5.407	2.913	2.161	1.971	3.777	1.279
100	8.701	3.459	1.946	1.535	5.034	0.708	4.511	2.528	1.831	1.667	3.070	1.286
150	5.220	2.102	1.164	0.964	2.727	0.716	3.560	2.084	1.558	1.458	2.359	1.326
200	3.872	1.693	0.979	0.835	2.026	0.714	3.052	1.885	1.470	1.401	2.043	1.338
ρ^2=0.99												
50	144.7	50.01	57.38	51.74	121.78	15.755	16.58	7.815	7.549	7.429	13.98	2.554
75	68.14	23.18	20.30	17.95	51.79	4.515	12.36	5.853	5.208	4.999	9.851	1.827
100	47.87	16.10	13.68	11.79	34.46	3.024	10.55	5.078	4.417	4.125	8.197	1.649
150	28.98	9.606	7.280	6.222	19.16	1.912	8.303	4.013	3.245	3.007	6.101	1.442
200	20.74	6.917	4.908	4.003	12.80	1.255	7.013	3.442	2.709	2.447	4.950	1.340