Performance of Some Ridge Parameters for Probit Regression:

with Application on Swedish Job Search Data

Håkan Locking[1], Kristofer Månsson[2] and Ghazi Shukur[1,2]

[1] *Department of Economics and Statistics, Linnaeus University, Sweden.*

[2] *Department of Economics, Finance and Statistics, Jönköping University, Sweden.*

Abstract

In ridge regression the estimation of the ridge parameter is an important issue. This paper generalizes some methods for estimating the ridge parameter for probit ridge regression (PRR) model based on the work of Kibria et al. (2011). The performance of these new estimators are judged by calculating the mean square error (MSE) using Monte Carlo simulations. In the design of the experiment we chose to vary the sample size and the number of regressors. Furthermore, we generate explanatory variables that are linear combinations of other regressors, which is a common situation in economics. In an empirical application regarding Swedish job search data we also illustrate the benefits of the new method.

Keywords: probit regression; maximum likelihood; multicollinearity; ridge regression; MSE; job search.

Mathematics Subject Classification: Primary 62J07; Secondary 62J02

1. Introduction

In this paper we investigate the effect of having explanatory variables, that are a linear combination of other regressors, on the probit regression model. This problem is very common in the area of microeconometrics and it leads to high variance and instability when estimating the unknown vector of coefficients by applying the traditional maximum likelihood (ML) method. A popular solution to this type of problem is ridge regression introduced for the linear regression model by Hoerl and Kennard (1970a,b). The authors showed in that paper that the ridge regression estimator has better mean squared error (MSE) properties than ordinary least squares (OLS) when the explanatory variables are collinear. Ridge regression estimator for other models such as the logit and probit has then, based on the result from Hoerl and Kennard (1970a,b), been derived for the non-linear logit and Poisson models by Schaeffer et al. (1984), Månsson and Shukur (2011a,b), among others.

The purpose of this paper is to develop probit ridge regression (PRR) by generalizing some methods of estimating the ridge parameter evaluated in Kibria et al. (2011) so they can be used for this estimation method. In order to be able to judge the performance of the different methods of estimating k we calculate the mean squared error (MSE) using Monte Carlo simulations. In the design of the experiment we chose to vary the sample size, the number of explanatory variables and the degree of correlation. Furthermore, we chose to generate explanatory variables that are linear combinations of other regressors and we evaluate the effect of both continuous regressors and dummy variable. Hence, in the simulation study we replicate an empirically relevant situation which is usually not considered when different ridge parameters are evaluated. The result from the simulation study shows that the PRR always outperforms the ML in the presence of highly correlated linear combinations of the regressors. Then, in an empirical application the benefit of using PRR instead of ML is illustrated to practitioners. We show that using this new estimation method we obtain estimators of the unknown vector of coefficients with much lower variances than the ML method.

The paper is organized as follows: in Section 2, we describe the statistical methodology. The design of the experiment and simulated results are provided in Section 3. In Section 4 we provide an empirical example while in section 5 we give a brief summary and conclusions.

2. Methodology

This section defines the probit regression model and describes the PRR and the traditional ML estimation methods.

2.1 The Probit Ridge Regression Estimator

Consider the following regression model:

$$y_i{*} = x_i{'}\beta + u_i \tag{2.1}$$

where $y_i{*}$ is an latent variable, x_i is the i*th* row of X which is an $n\times(p+1)$ data matrix with p explanatory variables, β is a $(p+1)\times 1$ vector of coefficients and u_i is an error term assumed to be normally distributed. The latent variable is not observable in reality; instead we may analyze the following dummy variable:

$$y_i = \begin{cases} 1 & \text{if } y_i{*} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

which is distributed as $Be(\pi_i)$ where $\pi_i = \Phi(x_i{'}\beta)$ and Φ is the distribution function of the standard normal distribution. In this situation the probit regression model should be used which is estimated by ML by applying the subsequent iterative weighted least square (IWLS) algorithm discussed in Cameron and Trivedi (1998):

$$\beta_{ML} = \left(X{'}\hat{W}X\right)^{-1}\left(X{'}\hat{W}\hat{z}\right) \tag{2.3}$$

where $\hat{W} = diag\left[\frac{\left(\phi(x_i{'}\beta)\right)^2}{\Phi(x_i{'}\beta)\left(1-\Phi(x_i{'}\beta)\right)}\right]$ and $\hat{z}$ is a vector where the *i*th element equals $\hat{z}_i = \log(\hat{\pi}_i) + \frac{y_i - \hat{\pi}_i}{\hat{\pi}_i(1-\hat{\pi}_i)}$. The MSE of the ML estimator corresponds to:

$$E\left(L_{ML}^2\right) = tr\left(X'\hat{W}X\right)^{-1} = \sum_{j=1}^{J} \frac{1}{\lambda_j}, \quad (2.4)$$

where λ_j is the *j*th eigenvalue of the $X'\hat{W}X$ matrix. When the explanatory variables are collinear some eigenvalues will be small which inflate the MSE. In this situation the following PRR estimator might be a better alternative:

$$\beta_{RR} = \left(X'\hat{W}X + kI\right)^{-1} X'\hat{W}X\beta_{ML}. \quad (2.5)$$

The MSE of this estimator equals:

$$E\left(L_{ML}^2\right) = \sum_{j=1}^{J} \frac{\lambda_j}{\left(\lambda_j + k\right)^2} + k^2 \sum_{j=1}^{J} \frac{\alpha_j^2}{\left(\lambda_j + k\right)^2}, \quad (2.6)$$

where the first term corresponds to the variance and the second term equals the squared bias. The PRR estimator will have a lower MSE than the ML estimate if we find a value of *k* such that the reduction in the variance term is greater than the increase of the squared bias.

2.2 Suggested estimators of the ridge parameter

There is not a definite rule of how to estimate the ridge parameter *k*. However, many suggestions have been given for the linear regression model and some of them will be generalized in this paper so they are applicable for PRR. The first one that we suggest is based on the classical ridge parameter proposed by Hoerl and Kennard (1970a,b):

$$K1 = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2},$$

where we define $\hat{\alpha}_{\max}^2$ to be the maximum element of $\gamma\beta_{ML}$ and $\hat{\sigma}^2$ corresponds to the sum of square deviance residuals divided by the degrees of freedom ($n-p-1$). In Schaeffer et al. (1984) a modified version of this estimator was proposed:

$$K2 = \frac{1}{\hat{\alpha}_{\max}^2}.$$

Furthermore, two ridge regression estimators will be proposed based on Kibria (2003):

$K3 = \frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{l} \hat{\alpha}_j^2\right)^{\frac{1}{l}}}$, and $K4 = \underset{i}{median}\left(m_{ji}^2\right)$, where $m_j = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}}$.

We then propose the following ridge parameter evaluated by Kibria et al. (2011):

$$K5 = \max\left(\frac{1}{m_j}\right),\ K6 = \max\left(m_j\right),\quad K7 = \prod_{j=1}^{p}\left(\frac{1}{m_j}\right)^{\frac{1}{p}},\quad K8 = \prod_{j=1}^{p}\left(m_j\right)^{\frac{1}{p}}$$

$$K9 = median\left(\frac{1}{m_j}\right), K10 = median\left(m_j\right), K11 = \max\left(\frac{1}{q_j}\right),\quad K12 = \max\left(q_j\right),\ K13 = \prod_{j=1}^{p}\left(\frac{1}{q_j}\right)^{\frac{1}{p}},$$

$$K14 = \prod_{j=1}^{p}\left(q_j\right)^{\frac{1}{p}},\quad K15 = median\left(\frac{1}{q_j}\right),\ K16 = median\left(q_j\right),$$

where $q_j = \frac{\lambda_{\max}}{(n-p)\hat{\sigma}^2 + \lambda_{\max}\hat{\alpha}_j^2}$ and $\lambda_{\max}$ is defined as the maximum eigenvalue of $X'\hat{W}X$.

2.3 Judging the performance of the estimators

To investigate the performance of the PRR and ML method we calculate the MSE using the following equation:

$$MSE = \frac{\sum_{i=1}^{R} SE_i}{R} = \frac{\sum_{i=1}^{R}\left(\hat{\beta}-\beta\right)_i'\left(\hat{\beta}-\beta\right)_i}{R}, \tag{2.7}$$

where $\hat{\beta}$ is the estimator of β obtained from ML or PRR and R equals 2000 which corresponds to the number of replicates used in the Monte Carlo simulation.

3. The Monte Carlo simulation

In this section we describe the design of the experiment and discuss the result of the simulation study.

3.1 The Design of the Experiment

Following Kibria (2003) we generate p explanatory variables using the following equation,

$$x_{ij} = (1-\theta)^{(1/2)} z_{ij} + \rho_1 z_{ij+1} + \rho_2 z_{ij+2} + \cdots + \rho_p z_{ij+p}, i = 1,2,...n, \tag{3.1}$$

where $\theta = \left(\sum_{j=1}^{p} \rho_j \right)^2$ represents to which degree the explanatory variable is determined by the other regressors, and z_{ij} are pseudo-random numbers from the standard normal distribution. When dummy variables are used instead we consider the x_{ij} to be latent variables and we make the explanatory variables binary by applying equation (2.2). The dependent latent variable is then generated using the following formula:

$$y_i{*} = \beta_1 x_{i1} + \ldots + \beta_l x_{ip} + u_i \tag{3.2}$$

where u_i are pseudo-random numbers from the standard normal distribution. This latent variable is also going to be made binary by using equation (2.2).

The factors we chose to vary in the Monte Carlo experiment are the degree of correlation, the number of observations and the number of explanatory variables. Three different values of θ corresponding to 0.85, 0.95 and 0.99 are considered. We study sample sizes with 100, 250, 500 and 1000 observations and equations with 5 and 10 regressors. We will generate models consisting of only 5 or 10 continuous regressors. Furthermore, models consisting of a mixture between continuous and discrete random variables will be considered. In the mixture models 40 % of the regressors will be dummy variables and 60 % continuous variables.

3.2 Result Discussion

The estimated MSEs of the different estimation methods can be found in Tables 1 and 2. The factors that have an impact on the estimated MSE are to what degree the explanatory variables are determined by the other regressors, the number of observations and the number of explanatory variables. Increasing θ while holding *n* and *p* fixed leads, in general, to a higher estimated MSE for ML and PRR when applying most of the different ridge parameters. The least robust option of estimating k is to use either the K1 and K2 that are proposed by Hoerl and Kennard (1970a,b) and Schaeffer et al. (1984), respectively. Other ridge parameters that are better than these two but still do not work well in the presence of multicollinearity are the those based on q_j (i.e, K12, K14 and K16). However, for PRR when the ridge parameter is estimated using either the inverse of m_j (ridge parameters K5, K7 and K9) or the inverse of q_j (i.e, K11, K13 and K15) the estimated MSE occasionally decreases. The ridge parameters that are calculated based on the inverse of q_j are the ones with the lowest estimated MSE for all different values of θ when the sample size is low. However, in contrast to most of the other ridge parameters, the estimated MSE of K11, K13 and K15 increases with the sample size. Hence, when the number of observations is large the ridge parameters K3 and K6 should be preferred. These results hold for both 5 and 10 linear combinations and both when we have only continuous variables and a mixture between discrete and continuous variables.

Table 1: Estimated MSE when all regressors are continuous

Estimated MSE when $p=5$																	
	ML	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16
θ=0.85																	
100	1.461	1.148	1.018	0.361	0.442	0.548	0.291	0.917	0.474	0.864	0.533	0.181	1.365	0.172	1.371	0.171	1.369
250	0.599	0.508	0.460	0.199	0.247	0.417	0.206	0.517	0.299	0.504	0.327	0.117	0.572	0.126	0.574	0.129	0.573
500	0.214	0.196	0.186	0.104	0.120	0.192	0.118	0.206	0.152	0.204	0.158	0.119	0.209	0.126	0.209	0.127	0.209
1000	0.103	0.098	0.095	0.064	0.073	0.098	0.078	0.101	0.085	0.101	0.087	0.079	0.102	0.081	0.102	0.082	0.102
θ=0.95																	
100	1.617	1.258	1.098	0.368	0.473	0.570	0.297	0.989	0.486	0.923	0.558	0.183	1.501	0.174	1.508	0.172	1.505
250	0.653	0.548	0.489	0.205	0.254	0.443	0.207	0.558	0.310	0.542	0.340	0.117	0.621	0.126	0.623	0.129	0.623
500	0.235	0.214	0.201	0.108	0.126	0.207	0.122	0.224	0.160	0.223	0.166	0.123	0.229	0.130	0.229	0.131	0.229
1000	0.114	0.108	0.104	0.069	0.077	0.107	0.080	0.111	0.092	0.111	0.094	0.084	0.112	0.087	0.112	0.087	0.112
θ=0.99																	
100	1.809	1.320	1.109	0.344	0.439	0.496	0.286	0.950	0.432	0.876	0.499	0.182	1.623	0.172	1.635	0.170	1.630
250	0.781	0.620	0.537	0.196	0.245	0.442	0.200	0.620	0.295	0.595	0.330	0.104	0.725	0.112	0.729	0.114	0.727
500	0.293	0.256	0.234	0.109	0.128	0.242	0.126	0.272	0.171	0.269	0.182	0.118	0.281	0.126	0.282	0.128	0.281
1000	0.141	0.130	0.123	0.069	0.079	0.129	0.085	0.136	0.101	0.135	0.106	0.090	0.137	0.094	0.137	0.095	0.137
Estimated MSE when $p=10$																	
θ=0.85																	
100	6.898	4.908	3.709	0.679	0.923	0.789	0.374	1.891	0.842	1.670	1.004	0.225	4.404	0.210	4.678	0.210	4.582
250	2.335	1.729	1.309	0.275	0.367	0.813	0.268	1.401	0.523	1.292	0.606	0.232	1.775	0.276	1.818	0.284	1.801
500	0.811	0.654	0.525	0.138	0.175	0.515	0.173	0.687	0.298	0.665	0.330	0.257	0.697	0.285	0.704	0.290	0.701
1000	0.376	0.323	0.272	0.089	0.108	0.304	0.116	0.351	0.186	0.347	0.200	0.194	0.345	0.207	0.347	0.209	0.346
θ=0.95																	
100	10.37	7.450	5.490	0.820	1.184	0.663	0.348	2.059	0.936	1.749	1.143	0.222	6.076	0.195	6.591	0.194	6.415
250	3.805	2.851	2.101	0.349	0.482	0.909	0.263	1.937	0.647	1.760	0.760	0.196	2.730	0.250	2.828	0.262	2.788
500	1.307	1.047	0.815	0.159	0.213	0.727	0.176	1.060	0.386	1.016	0.441	0.288	1.089	0.340	1.105	0.349	1.098
1000	0.632	0.539	0.441	0.107	0.141	0.483	0.127	0.580	0.270	0.569	0.300	0.272	0.570	0.300	0.573	0.304	0.572
θ=0.99																	
100	17.18	11.68	8.157	1.077	1.571	0.506	0.351	1.593	0.983	1.335	1.202	0.244	6.804	0.196	8.106	0.194	7.663
250	6.118	4.218	2.905	0.429	0.598	0.696	0.263	1.764	0.650	1.533	0.772	0.169	3.311	0.216	3.598	0.227	3.489
500	2.072	1.514	1.071	0.183	0.250	0.707	0.171	1.326	0.396	1.223	0.461	0.240	1.470	0.294	1.519	0.304	1.499
1000	0.983	0.766	0.571	0.110	0.144	0.566	0.125	0.813	0.274	0.783	0.311	0.241	0.802	0.281	0.814	0.288	0.809

Table 2: Estimated MSE when there is a mix between continuous and discrete regressors

Estimated MSE when *p=5*																	
	ML	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16
θ=0.85																	
100	1.152	0.898	0.795	0.303	0.365	0.524	0.284	0.804	0.397	0.764	0.447	0.170	1.085	0.165	1.089	0.164	1.087
250	0.467	0.396	0.361	0.170	0.200	0.351	0.194	0.419	0.243	0.410	0.264	0.115	0.449	0.123	0.450	0.125	0.450
500	0.172	0.158	0.150	0.091	0.103	0.157	0.114	0.166	0.126	0.165	0.130	0.106	0.168	0.110	0.168	0.111	0.168
1000	0.082	0.078	0.076	0.054	0.059	0.079	0.069	0.081	0.069	0.080	0.071	0.066	0.081	0.067	0.081	0.068	0.081
θ=0.95																	
100	1.299	0.978	0.852	0.311	0.389	0.537	0.285	0.848	0.412	0.798	0.471	0.169	1.192	0.162	1.196	0.161	1.194
250	0.529	0.445	0.400	0.181	0.216	0.383	0.201	0.464	0.268	0.453	0.293	0.116	0.507	0.124	0.508	0.126	0.508
500	0.187	0.171	0.159	0.092	0.106	0.170	0.114	0.180	0.134	0.179	0.140	0.109	0.183	0.113	0.183	0.114	0.183
1000	0.090	0.086	0.083	0.057	0.063	0.087	0.067	0.089	0.076	0.089	0.077	0.071	0.089	0.072	0.089	0.072	0.089
θ=0.99																	
100	1.380	0.935	0.795	0.282	0.347	0.504	0.262	0.812	0.373	0.764	0.424	0.181	1.174	0.173	1.179	0.172	1.177
250	0.619	0.445	0.388	0.167	0.198	0.377	0.191	0.470	0.249	0.457	0.274	0.108	0.522	0.114	0.523	0.115	0.523
500	0.248	0.177	0.161	0.091	0.103	0.176	0.112	0.189	0.134	0.187	0.140	0.102	0.193	0.106	0.193	0.107	0.193
1000	0.115	0.090	0.085	0.057	0.065	0.090	0.066	0.093	0.077	0.093	0.080	0.070	0.094	0.071	0.094	0.071	0.094
Estimated MSE when *p=10*																	
θ=0.85																	
100	3.927	3.128	2.515	0.489	0.684	0.925	0.382	1.968	0.747	1.774	0.902	0.244	3.340	0.234	3.392	0.233	3.373
250	1.308	1.109	0.921	0.244	0.302	0.808	0.283	1.101	0.478	1.065	0.540	0.235	1.207	0.264	1.213	0.269	1.210
500	0.445	0.404	0.356	0.134	0.161	0.391	0.183	0.427	0.263	0.423	0.283	0.243	0.429	0.256	0.429	0.258	0.429
1000	0.207	0.195	0.181	0.084	0.099	0.196	0.120	0.204	0.151	0.203	0.158	0.156	0.203	0.161	0.203	0.161	0.203
θ=0.95																	
100	4.722	4.384	3.463	0.593	0.850	0.855	0.379	2.154	0.818	1.881	1.001	0.243	4.416	0.227	4.565	0.225	4.513
250	1.979	1.621	1.284	0.275	0.372	0.945	0.283	1.498	0.558	1.413	0.652	0.215	1.748	0.250	1.763	0.256	1.757
500	0.664	0.583	0.496	0.141	0.178	0.532	0.184	0.619	0.318	0.609	0.351	0.269	0.625	0.294	0.626	0.298	0.626
1000	0.302	0.278	0.249	0.092	0.114	0.275	0.124	0.294	0.192	0.292	0.206	0.197	0.292	0.205	0.292	0.206	0.292
θ=0.99																	
100	5.703	4.341	3.316	0.584	0.833	0.838	0.386	2.119	0.795	1.844	0.975	0.246	4.456	0.230	4.576	0.229	4.535
250	2.117	1.713	1.370	0.340	0.452	0.907	0.289	1.458	0.530	1.375	0.620	0.211	1.775	0.241	1.848	0.246	1.842
500	0.669	0.580	0.479	0.139	0.174	0.528	0.178	0.619	0.306	0.608	0.339	0.254	0.627	0.276	0.629	0.279	0.628
1000	0.319	0.291	0.255	0.090	0.112	0.288	0.127	0.309	0.193	0.306	0.208	0.197	0.307	0.206	0.308	0.207	0.307

4. Empirical Application

A standard job search model predicts that optimal search behaviors generate a reservation wage and the worker will accept any offer above his reservation wage. Firms, on the other hand, create vacancies to maximize profit which generate an exogenous flow of offers to the workers. However, the probability of receiving a job offer will be influenced by the effort an unemployed person exerts. The hiring situation is also characterized by imperfect information, so the employer have to relay on attributes in there attempt to value the job-seekers. Such attributes could be age, chosen search channel, education and all individual attributes the firm could observe.

Böheim and Taylor (2002) found that direct contact is the most effective search method in Britain and that other search methods were not significant. In another UK study, Frijters et al. (2005) using a panel of unemployed men during 1997-2001 found that search channels such as direct contact, social networks and agencies are more effective than using job centers and newspaper. Using data from 1981, 1983 and 1986 for Canada, Osberg (1993) found significant effects for different search methods depending on sample and year, and in most of the estimates there were only 1-2 search methods for each year that were significant. The results using US data show the same thing, only the search trough newspapers was significant out of 5 studied methods in Holzer (1988). The above studies report that job seekers use about 3 methods on average, out of 5-6 studied alternatives. As Böheim and Taylor (2002) point out, job search does not appear to be a single, uniform activity for the unemployed seeking work. Thus researchers have numerous nominal variables that are not mutually exclusive so we can expect a large degree of multicollinearity between the including variables, especially as other variables are included in the studies as well or put in another way, the data could contain to little variation to be able to answer detailed questions about the search effectiveness of individual channels.

This study uses a dataset earlier used by Bolinder (1999) containing a random selection of 1806 registered unemployed in the beginning of 1996 in Sweden. The data was kindly supplied by Mattias Strandh, Umeå University. The outcome is if the respondent has got a job or not during a 2 year period after the first contact. Our search channels include using newspaper advertising, using friends or own contact. Each variable is graded from never used it, sometimes used it to using it often. We do not use search through public employment service as a variable as it is mandatory for gaining access to unemployment benefits. Thus our reference point includes search through public employment services.

The other exogenous variables, observed in the initial period, used in this study are Time spent in search, Number of contacts with employers, Work experience in desired job divided up into 2 categories, Gender, Age, Age Squared, Education divide up into 3 categories, Citizenship, Civil status, Handicap, Length of unemployment spell, Earlier work classification divided up into 6 categories and if the individual has worked before or not. We also include a variable that measures attitude or motivation towards work. The variable is defined as a summation of categorical values on Importance of working, Like to work even if you have money, Dislikes being unemployed, Become boring if you don't have a job. To have a job is among the most important things in life. Thus, we expect Attitude toward work to have a positive impact on the likelihood to get a job. To allow for nonlinear effect from the categorical variable we include it squared as well.

The results can be found in Tables 3 and 4 are for values of the estimated coefficients together with the vector bootstrapped standard errors (in parenthesis). Some estimators of k parameters have very high values and push all coefficients to zero while others give a very low value of k so they do not adjust the coefficients. The results are broadly consistent with the simulations study, the suggested parameterization of the method K1, K13 and K15 reduce the standards errors, although method K5, K7, K9 and K11 are similar in this sample. The average reduction in standard errors are between 0 and 57% for individual coefficients for the suggested method K1, K13, K15 and the unweighted average reduction is about 40%. However, the number of significant coefficients at the 5% level does not change that much. The K13 is a clear exception from this conclusion in this sample. The K13 produces estimated parameters that are very close to those from the ML (which are consistent in large samples) and at the same time it heavily reduces the standard errors of the coefficient so that these estimated parameters become statistically significant.

The overall results are broadly in line with earlier cited studies. The results show a large improvement in the precision of the estimated effects on the different search channels, and searching through friends are the outstanding channel, if the objective is to find a job. More astonishing is that an extensive search on your own and through newspaper seems to be counter productive, when we control for time spend in search. Thus indicate that regulations that require obliged search and employer contact attempts for workers on unemployment benefits are not an effective method to improve their job chances. The results instead suggest that social networking is effective as a means to get a job during the high unemployment period of this study. Moreover, the results emphasize that highly educated and persons with skilled blue collar or high white collar work experience have a large advantage in the job search market.

Table 3: Impact of search strategies and human capital variables on the probability of obtaining a job during 2006-2008 conditioned on being unemployed 2006. Results for Maximum Likelihood and Probit Ridge Regression methods.

	ML	K1	K2	K3	K4	K5	K6	K7	K8
Medium Education	0.28416 (0.09701)	0.28039 (0.04582)	0.10371 (0.09646)	0.00451 (0.00504)	0.01025 (0.00619)	0.28032 (0.09639)	0.00036 (0.02464)	0.28383 (0.09695)	0.13253 (0.05891)
High Education	0.50942 (0.16123)	0.50974 (0.04409)	0.17535 (0.15974)	0.00549 (0.00339)	0.01247 (0.00423)	0.50974 (0.16001)	0.00053 (0.02128)	0.50946 (0.16112)	0.23732 (0.06873)
Age	0.03147 (0.02710)	0.02691 (0.01524)	0.00349 (0.02655)	-0.00374 (0.00606)	-0.00163 (0.00700)	0.02682 (0.02516)	-0.00238 (0.00958)	0.03106 (0.02697)	0.00218 (0.01641)
Newspaper (sometimes)	-0.06619 (0.09809)	-0.06635 (0.04745)	-0.02685 (0.09762)	-0.00251 (0.00536)	-0.00492 (0.00628)	-0.06636 (0.09768)	-0.00033 (0.02491)	-0.06621 (0.09804)	-0.03281 (0.06067)
Newspaper (often)	-0.03054 (0.12626)	-0.02929 (0.04864)	0.01107 (0.12544)	0.00138 (0.00408)	0.00276 (0.00532)	-0.02927 (0.12560)	0.00018 (0.01950)	-0.03043 (0.12619)	0.00909 (0.06449)
Own contact (sometimes)	-0.06081 (0.09402)	-0.06089 (0.04464)	-0.01753 (0.09355)	-0.00229 (0.00431)	-0.00436 (0.00647)	-0.06089 (0.09362)	-0.00031 (0.01714)	-0.06082 (0.09398)	-0.02159 (0.05680)
Own contact (often)	-0.10782 (0.12327)	-0.10797 (0.04574)	-0.00753 (0.12246)	0.00192 (0.00422)	0.00363 (0.00531)	-0.10797 (0.12262)	0.00024 (0.01937)	-0.10784 (0.12320)	-0.01960 (0.06243)
Friends (sometimes)	0.06383 (0.09892)	0.06258 (0.04830)	0.00120 (0.09844)	-0.00211 (0.00460)	-0.00398 (0.00687)	0.06256 (0.09854)	-0.00027 (0.01672)	0.06373 (0.09888)	0.00875 (0.06047)
Friends (often)	0.29573 (0.12377)	0.29353 (0.04805)	0.11694 (0.12305)	0.00618 (0.00426)	0.01330 (0.00558)	0.29348 (0.12312)	0.00065 (0.01969)	0.29554 (0.12370)	0.14699 (0.06474)
Time spent in search (Hours)	0.00642 (0.00552)	0.00655 (0.00491)	0.00617 (0.00551)	0.00478 (0.00404)	0.00502 (0.00425)	0.00656 (0.00552)	0.00281 (0.00433)	0.00643 (0.00552)	0.00642 (0.00503)
Number of Contacts with employer	0.00088 (0.01597)	0.00071 (0.01489)	0.00133 (0.01596)	0.00295 (0.00880)	0.00325 (0.01042)	0.00071 (0.01596)	0.00111 (0.01211)	0.00086 (0.01597)	0.00089 (0.01516)
Female	0.03261 (0.08507)	0.03120 (0.04690)	-0.00690 (0.08479)	-0.00289 (0.00525)	-0.00537 (0.00658)	0.03117 (0.08476)	-0.00042 (0.02170)	0.03248 (0.08504)	-0.00207 (0.05718)
Single	-0.27021 (0.09230)	-0.27340 (0.05023)	-0.18004 (0.09193)	-0.01217 (0.00720)	-0.02572 (0.00664)	-0.27345 (0.09172)	-0.00131 (0.03802)	-0.27049 (0.09226)	-0.21041 (0.06369)
Foreign Citizenship	-0.00451 (0.16826)	0.00081 (0.04416)	-0.01160 (0.16675)	-0.00083 (0.00260)	-0.00185 (0.00375)	0.00091 (0.16691)	-0.00008 (0.01300)	-0.00403 (0.16813)	-0.01065 (0.06513)
Handicap	-0.33498 (0.16665)	-0.33354 (0.04520)	-0.10795 (0.16528)	-0.00417 (0.00270)	-0.00924 (0.00365)	-0.33351 (0.16552)	-0.00042 (0.01479)	-0.33486 (0.16653)	-0.14271 (0.06736)
Some earlier work experience in the desired job	0.15106 (0.11333)	0.14895 (0.04405)	0.04222 (0.11263)	0.00046 (0.00396)	0.00161 (0.00580)	0.14891 (0.11266)	-0.00001 (0.01410)	0.15088 (0.11327)	0.05725 (0.05747)
Good earlier work experience in the desired job	0.09647 (0.12089)	0.09763 (0.04607)	0.04500 (0.12018)	0.00320 (0.00464)	0.00673 (0.00589)	0.09765 (0.12029)	0.00030 (0.01996)	0.09658 (0.12083)	0.05540 (0.06085)

Length of unemployment spell	-0.00066 (0.00020)	-0.00066 (0.00020)	-0.00070 (0.00020)	-0.00074 (0.00020)	-0.00073 (0.00020)	-0.00066 (0.00020)	-0.00076 (0.00020)	-0.00066 (0.00020)	-0.00070 (0.00020)
Worked before	0.03046 (0.15436)	0.01875 (0.04825)	-0.05019 (0.15321)	-0.00371 (0.00357)	-0.00779 (0.00464)	0.01853 (0.15216)	-0.00039 (0.01754)	0.02943 (0.15415)	-0.05773 (0.06716)
Skilled Blue collar worker	0.41636 (0.11436)	0.41309 (0.04840)	0.18281 (0.11381)	0.01118 (0.00615)	0.02397 (0.00591)	0.41303 (0.11361)	0.00112 (0.03058)	0.41609 (0.11427)	0.22036 (0.06473)
Low white collar worker	-0.03674 (0.12527)	-0.03667 (0.04695)	-0.05817 (0.12445)	-0.00374 (0.00329)	-0.00797 (0.00498)	-0.03668 (0.12445)	-0.00041 (0.01486)	-0.03672 (0.12518)	-0.06650 (0.06315)
Mid white collar worker	0.06175 (0.16829)	0.05963 (0.04602)	0.00972 (0.16666)	0.00025 (0.00265)	0.00066 (0.00400)	0.05959 (0.16676)	0.00002 (0.01246)	0.06158 (0.16813)	0.01128 (0.06615)
High white collar worker	0.51352 (0.26063)	0.50565 (0.03564)	0.06872 (0.25629)	0.00218 (0.00210)	0.00487 (0.00248)	0.50548 (0.25662)	0.00022 (0.01249)	0.51289 (0.26020)	0.09878 (0.06061)
Executive or had own business	0.75596 (0.23278)	0.74672 (0.03246)	0.11502 (0.22981)	0.00339 (0.00278)	0.00766 (0.00229)	0.74652 (0.22881)	0.00034 (0.02165)	0.75523 (0.23228)	0.16605 (0.06845)
Attitude to work	0.06473 (0.06069)	0.05231 (0.03192)	-0.03354 (0.05909)	-0.00987 (0.00759)	-0.01600 (0.00970)	0.05208 (0.05477)	-0.00203 (0.01747)	0.06363 (0.06028)	-0.03332 (0.03487)
Age Squared	-0.00047 (0.00034)	-0.00041 (0.00020)	-0.00011 (0.00033)	0.00000 (0.00010)	-0.00003 (0.00011)	-0.00041 (0.00031)	-0.00003 (0.00014)	-0.00046 (0.00033)	-0.00010 (0.00022)
Length of unemployment spell Squared	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)
Attitude to work Squared	-0.00089 (0.00167)	-0.00056 (0.00094)	0.00160 (0.00163)	0.00090 (0.00033)	0.00109 (0.00038)	-0.00056 (0.00152)	0.00057 (0.00055)	-0.00086 (0.00166)	0.00161 (0.00102)

Note: The standard errors are in parenthesis.

Table 4:

	K9	K10	K11	K12	K13	K14	K15	K16
Medium Education	0.28376 (0.09693)	0.14606 (0.06216)	0.28327 (0.09683)	0.00000 (0.00021)	0.28415 (0.02464)	0.00545 (0.09695)	0.28415 (0.05891)	0.00689 (0.09693)
High Education	0.50947 (0.16111)	0.26720 (0.07406)	0.50952 (0.16094)	0.00000 (0.00013)	0.50942 (0.02128)	0.00661 (0.16112)	0.50942 (0.06873)	0.00835 (0.16111)
Age	0.03097 (0.02695)	0.00155 (0.01669)	0.03037 (0.02637)	-0.00001 (0.00069)	0.03146 (0.00958)	-0.00341 (0.02697)	0.03146 (0.01641)	-0.00287 (0.02695)
Newspaper (sometimes)	-0.06621 (0.09803)	-0.03580 (0.06421)	-0.06624 (0.09799)	0.00000 (0.00029)	-0.06619 (0.02491)	-0.00293 (0.09804)	-0.06619 (0.06067)	-0.00356 (0.09803)
Newspaper (often)	-0.03041 (0.12617)	0.00750 (0.06944)	-0.03025 (0.12610)	0.00000 (0.00021)	-0.03054 (0.01950)	0.00162 (0.12619)	-0.03054 (0.06449)	0.00198 (0.12617)
Own contact (sometimes)	-0.06083 (0.09397)	-0.02391 (0.05932)	-0.06084 (0.09393)	0.00000 (0.00013)	-0.06081 (0.01714)	-0.00267 (0.09398)	-0.06081 (0.05680)	-0.00321 (0.09397)
Own contact (often)	-0.10785 (0.12319)	-0.02617 (0.06669)	-0.10788 (0.12312)	0.00000 (0.00022)	-0.10782 (0.01937)	0.00223 (0.12320)	-0.10782 (0.06243)	0.00269 (0.12319)
Friends (sometimes)	0.06370 (0.09887)	0.01272 (0.06376)	0.06354 (0.09882)	0.00000 (0.00011)	0.06383 (0.01672)	-0.00246 (0.09888)	0.06383 (0.06047)	-0.00295 (0.09887)
Friends (often)	0.29550 (0.12369)	0.16087 (0.06894)	0.29522 (0.12361)	0.00000 (0.00020)	0.29573 (0.01969)	0.00737 (0.12370)	0.29573 (0.06474)	0.00917 (0.12369)
Time spent in search (Hours)	0.00643 (0.00552)	0.00654 (0.00506)	0.00645 (0.00552)	0.00003 (0.00102)	0.00642 (0.00433)	0.00484 (0.00552)	0.00642 (0.00503)	0.00491 (0.00552)
Number of Contacts with employer	0.00086 (0.01597)	0.00070 (0.01521)	0.00084 (0.01597)	0.00001 (0.00105)	0.00088 (0.01211)	0.00305 (0.01597)	0.00088 (0.01516)	0.00315 (0.01597)
Gender	0.03246 (0.08503)	0.00037 (0.05981)	0.03227 (0.08498)	0.00000 (0.00027)	0.03261 (0.02170)	-0.00335 (0.08504)	0.03260 (0.05718)	-0.00400 (0.08503)
Single	-0.27056 (0.09226)	-0.22273 (0.06666)	-0.27098 (0.09211)	-0.00001 (0.00042)	-0.27022 (0.03802)	-0.01446 (0.09226)	-0.27022 (0.06369)	-0.01791 (0.09226)
Foreign Citizenship	-0.00392 (0.16811)	-0.00944 (0.07128)	-0.00321 (0.16790)	0.00000 (0.00009)	-0.00450 (0.01300)	-0.00100 (0.16813)	-0.00450 (0.06513)	-0.00126 (0.16811)
Handicap	-0.33484 (0.16651)	-0.15950 (0.07359)	-0.33467 (0.16638)	0.00000 (0.00011)	-0.33497 (0.01479)	-0.00500 (0.16653)	-0.33497 (0.06736)	-0.00627 (0.16651)
Some earlier work experience in the desired job	0.15084 (0.11326)	0.06449 (0.06105)	0.15057 (0.11316)	0.00000 (0.00015)	0.15106 (0.01410)	0.00062 (0.11327)	0.15106 (0.05747)	0.00088 (0.11326)
Good earlier work experience in the desired job	0.09660 (0.12082)	0.06033 (0.06464)	0.09676 (0.12074)	0.00000 (0.00024)	0.09648 (0.01996)	0.00381 (0.12083)	0.09648 (0.06085)	0.00472 (0.12082)
Length of unemployment spell	-0.00066	-0.00069	-0.00066	-0.00065	-0.00066	-0.00074	-0.00066	-0.00073

	(0.00020)	(0.00020)	(0.00020)	(0.00018)	(0.00020)	(0.00020)	(0.00020)	(0.00020)
Worked before	0.02919 (0.15412)	-0.06058 (0.07253)	0.02766 (0.15364)	0.00000 (0.00017)	0.03044 (0.01754)	-0.00441 (0.15415)	0.03043 (0.06716)	-0.00545 (0.15412)
Skilled Blue collar	0.41603 (0.11426)	0.23693 (0.06849)	0.41562 (0.11416)	0.00001 (0.00031)	0.41635 (0.03058)	0.01333 (0.11427)	0.41635 (0.06473)	0.01658 (0.11426)
Low white collar	-0.03672 (0.12517)	-0.06919 (0.06730)	-0.03669 (0.12507)	0.00000 (0.00010)	-0.03674 (0.01486)	-0.00445 (0.12518)	-0.03674 (0.06315)	-0.00552 (0.12517)
Mid white collar	0.06154 (0.16810)	0.01179 (0.07195)	0.06127 (0.16792)	0.00000 (0.00007)	0.06175 (0.01246)	0.00031 (0.16813)	0.06174 (0.06615)	0.00041 (0.16810)
High white collar	0.51275 (0.26012)	0.11537 (0.06827)	0.51179 (0.25970)	0.00000 (0.00008)	0.51350 (0.01249)	0.00261 (0.26020)	0.51350 (0.06061)	0.00328 (0.26012)
Executive or own business	0.75506 (0.23221)	0.19401 (0.07477)	0.75394 (0.23182)	0.00000 (0.00014)	0.75594 (0.02165)	0.00408 (0.23228)	0.75594 (0.06845)	0.00514 (0.23221)
Sumatt	0.06338 (0.06024)	-0.03304 (0.03554)	0.06175 (0.05836)	0.00000 (0.00059)	0.06471 (0.01747)	-0.01106 (0.06028)	0.06470 (0.03487)	-0.01271 (0.06024)
Age Squared	-0.00046 (0.00033)	-0.00009 (0.00022)	-0.00045 (0.00033)	-0.00006 (0.00004)	-0.00047 (0.00014)	-0.00001 (0.00033)	-0.00047 (0.00022)	-0.00001 (0.00033)
Length of unemployment spell Squared	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)	0.00000 (0.00000)
Sumatt Squared	-0.00085 (0.00165)	0.00161 (0.00104)	-0.00081 (0.00161)	0.00037 (0.00012)	-0.00089 (0.00055)	0.00094 (0.00166)	-0.00089 (0.00102)	0.00099 (0.00165)

Note: The standard errors are in parenthesis.

5. Conclusions

In this paper we generalize some new methods of estimating the ridge parameter *k*, evaluated for linear regression by Kibria et al. (2011), to be applicable for probit ridge regression (PPR). These new methods of estimating *k* for PRR are evaluated by means of Monte Carlo simulations along with the traditional ML method. In the simulation study we focus on the problem that several explanatory variables in the regression model are determined by linear combinations of other regressors. To judge the performance of the different estimation methods in this circumstance we estimate the MSE. We show that the degree of which an explanatory variable is determined by other regressors is important and increasing this factor yields an immense increase of the estimated MSE of ML. Instead of applying the ML we may recommend using PRR and estimate *k* using K11, K13 or K15 when the number of observations is low and the K3 and K6 estimators for large sample sizes, although they showed to heavily shrink the estimated parameter toward zero in the empirical study. In the empirical application we show that the problem of explanatory variables being a function of other regressors is an empirical relevant issue in microeconometrics and we also illustrate the PRR method. In the application we find that the average reduction in standard errors are between 0 and 57% for individual coefficients for the suggested method K1, K13, K15 and the unweighted average reduction is about 40%. More specifically, the K13 has shown to outperform the others in the empirical study in the sense that it produces parameter estimates that are very close to those of the ML method and at the same time have the smallest variances. The results show that married, highly educated and persons with skilled blue collar or high white collar work experience have a large advantage in the job search market and the most effective search method is to use friends.

Referenser:

Alkhamisi, M. A., Khalaf, G. and Shukur, G. (2006). Some modifications for Choosing Ridge Parameter. *Communications in Statistics- Theory and Methods*, **35**: 1-16.

Bolinder M (1999) Sökbeteendets betydelse för chanson att hitta ett jobb. Arbetsmarknad & Arbetsliv, år g 5, nr 1.

Böheim R. and Taylor M. P. (2002). Job search methods, intensity and success in Britain in the 1990s. Working paper No 0206, Department of economics, Johannes Kepler University of Linz.

Cameron, A. C. and Trivedi P. K. (1998). *Regression analysis of count data*. Cambridge University Press, New York.

Frijters, P., Shields, M. A. and Price, S. W. (2005), Job Search Methods and Their Success: A Comparison of Immigrants and Natives in the UK. The Economic Journal, 115: F359–F376.

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

Holzer, H. (1988). Search method use by unemployed youth, *Journal of Labor Economics*, **6**, 1-20.

Hoerl, A. E. and Kennard, R. W. (1970b). Ridge Regression: Application to Non-Orthogonal Problems. *Technometrics*, **12**, 69-82.

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., Månsson, K. and Shukur, G. (2011). Performance of some logistic ridge regression parameters. To appear in *Computational Economics*.

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

Månsson, K. and Shukur, G (2011b). A Poisson Ridge Regression estimator. Economic Modelling, **28**, Issue 4, 1475-1481.

Osberg, L. (1993). Fishing in different pools: job search strategies and job-finding success in Canada in the early 1980s, *Journal of Labor Economics*, **11**, 348-386.

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