Matematika, 2004, Jilid 20, bil. 1, hlm. 19–30

Numerical Comparisons for Various Estimators of Slope Parameters for Unreplicated Linear Functional Model

Abdul Ghapor Hussin
Centre for Foundation Studies in Science, University of Malaya, Kuala Lumpur
Email: ghapor@um.edu.my

Abstract A variety of consistent estimators of the slope parameter for unreplicated linear functional relationship model have been proposed over the years. This paper gives some numerical comparisons among them. As a result we are able to make specific recommendations regarding a choice among them.

Keywords Consistent estimator, slope parameter, unreplicated linear functional relationship model.

Abstrak *Berbagai penganggar konsistent telah dicadangkan untuk parameter kecerunan bagi hubungan fungsian linear tanpa replika. Artikel ini memberikan perbandingan berangka bagi penganggar-penganggar tersebut. Adalah diharapkan kita dapat mencadangkan pemilihan yang sesuai baginya.*

Katakunci *Penganggar konsistent, parameter kecerunan, hubungan fungsian linear tanpa replika.*

1 Introduction

The errors-in-variables model (EIVM) differs from the ordinary or classical linear regression model is that the true independent variables or the explanatory variables are not observed directly, but are masked by measurements error. In many practical situations, actual data are of this quality, especially in economics and other social sciences as well as in environmental science as shown by Hussin [6]. It is well known that the presence of errors of measurements in the independent or explanatory variables make the ordinary least squares estimators inconsistent and biased in large as well as small samples.

Suppose the variables X and Y are related by $Y = \alpha + \beta X$, where both the X and Y are observed with errors. This model comes under the errors-in-variables model. If X is a mathematical variables, this termed as a linear functional relationship model, and if X is a random variables, then this is termed a linear structural relationship model between X and Y. Kendall [7], [8] has formalised this distinction and Dolby [1] introduced an ultrastructural relationship model, which is a combination of the functional and structural relationship.

This paper discussed some of the known consistent estimators of the slope parameter, that is $\hat{\beta}$ for the unreplicated functional model. Many authors have proposed a consistent estimators for the above problem based on grouping and instrumental variable, for example the two-group method of Wald [10], the three-group method, the weighted regression, Housener-Brennan's method (Housner [3]) and Durbin's ranking method (Durbin [2]). The maximum likelihood estimation requires us to know in advance the ratio of the error variances and the suggestion was given by Lindley [9]. If both variables X and Y represent similar characteristics measured in the same unit, it can be assume that the ratio of the error variances is unity, but in practice this assumption is not always true. In the following section we present the model considered and established the notation. Section 3 outlines various known consistent estimators for β. In section 4 we present some results based on the simulation studies to evaluate the performance of each one of the estimators by looking at the observed mean squared errors and observed mean squared differences.

2 Unreplicated Functional Relationship Model

Suppose X and Y are two mathematical variables connected by a linear relation of the form $Y = \alpha + \beta X$. For any $i = 1, \ldots, n$, (x_i, y_i) are observations in particular but unknown value (X_i, Y_i) of (X, Y), both are subjected to errors, so $x_i = X_i + \delta_i$ and $y_i = Y_i + \varepsilon_i$ where $Y_i = \alpha + \beta X_i$. Here we assumed that δ_i and $\varepsilon_i, i = 1, \ldots, n$ are mutually independent and normally distributed random variables with zero means and finite variance σ^2 and τ^2 respectively and independent of (X_i, Y_i). That is

$$\delta_i \sim N(0, \sigma^2) \text{ and } \varepsilon_{\mathrm{i}} \sim N(0, \tau^2).$$

There are $(n + 4)$ parameters to be estimated, i.e. $\alpha, \beta, \sigma^2, \tau^2, X_1, \ldots, X_n$. The log likelihood function is given by

$$\begin{aligned} &\mathrm{L}(\alpha, \beta, X_1, ..., X_n, \sigma^2, \tau^2; x_1, ...x_n, y_1, ..., y_n) \\ &\quad = -n \log(2\pi) - \frac{1}{2} n(\log \sigma^2 + \log \tau^2) - \sum \frac{(x_i - X_i)^2}{2\sigma^2} - \sum \frac{(y_i - \alpha - \beta X_i)^2}{2\tau^2} \end{aligned}$$

Detail parameters estimation and the application of this model can be found in Hussin [4], [5].

3 Some Estimation of β for the Unreplicated Functional Model

Various solutions have been suggested for the estimator of β for the problem of unreplicated functional model. This section gives outlines of some of them.

(i) The two-group method

Wald [10] proposed this method to find consistent estimator for β. He computed the arithmetic means $(\bar{x}_1, \bar{y}_1)$ for the lower group of observations and $(\bar{x}_2, \bar{y}_2)$ for the higher group after ranking the observations in ascending order on the basis of the values of x_i and then dividing them into two equal sub-groups. If the total number of observations is odd then the central observation is omitted. Then he estimated

$$\hat{\beta} = \frac{(\bar{y}_2 - \bar{y}_1)}{(\bar{x}_2 - \bar{x}_1)},$$

where $\bar{y}$ and $\bar{x}$ are the means of all the sample observations. This method yields a consistent estimate of β even though it is not the most efficient since its variance is not the smallest possible.

(ii) The three-group method

This method has been proposed based on the same idea as the two-group method. The observations are first arranged in ascending order on the basis of the x_i values and then divided into three equal groups (or approximately equal if the number of observations is not exactly divisible by 3). The middle group is ignored from the analysis. The arithmetic means then computed for the lowest group $(\bar{x}_1, \bar{y}_1)$ and $(\bar{x}_3, \bar{y}_3)$ for the highest group. Hence the estimated slope parameter, $\hat{\beta}$ is given by the formula

$$\hat{\beta} = \frac{(\bar{y}_3 - \bar{y}_1)}{(\bar{x}_3 - \bar{x}_1)}.$$

This method in general gives a consistent estimate of β. Furthermore, the method is more efficient, than the two-group method.

(iii) Weighted regression

The weighted regression may be outlined as follows.

(a) Obtain an estimate of β from the regression $Y = f(X)$, that is

$$\hat{Y} = \hat{a}_o + \hat{b}_o \hat{X}.$$

(b) Obtain an estimate of β from the regression $X = f(Y)$, that is

$$\hat{X} = \hat{a}_1 + \hat{b}_1 \hat{Y}.$$

Take as the final estimate of β the geometric mean of the above two estimates, that is

$$\hat{\beta} = \left(\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \frac{\sum (y_i - \bar{y})^2}{\sum (x_i - \bar{x})(y_i - \bar{y})} \right)^{1/2} = \left(\frac{\sum (y_i - \bar{y})^2}{\sum (x_i - \bar{x})^2} \right)^{1/2}.$$

The weighted regression method is based on the implicit assumption that the ratio of the variances of the errors is equal to the ratio of the variances of the observed variables, that is

$$\frac{\tau^2}{\sigma^2} = \frac{Var(y)}{Var(x)}.$$

This assumption is necessary in order to make $\hat{\beta}$ a consistent estimate.

(iv) Housner-Brennan's method

Another consistent estimate of β was proposed by Housner and Brennan [3]. The observations are first arranged in ascending order on the basis of the x's values, i.e. we have $x_1 \leq x_2 \leq ... \leq x_n$. The estimate of β is given by

$$\hat{\beta} = \frac{\sum_{i=1}^{n} i(y_i - \bar{y})}{\sum_{i=1}^{n} i(x_i - \bar{x})}.$$

Housner and Brennan also show that, this estimate is more efficient compared to estimate obtained by

(i) minimising the sum of squares of the y deviation only

(ii) minimising the sum of squares of the orthogonal deviation and

(iii) the two-group method.

(v) Durbin's 'ranking' method

Durbin [2] has suggested, the estimate of β is given by

$$\hat{\beta} = \frac{\sum (x_i - \bar{x})^2 (y_i - \bar{y})}{\sum (x_i - \bar{x})^3},$$

where the x's and y's are ranked in ascending order, on the basis of the X values. Durbin has proved that the variance of $\hat{\beta}$ obtained from this method has a smaller variance than the two-group and three-group methods.

(vi) The maximum likelihood approach and with λ assumed known

Let $\lambda = \dfrac{\tau^2}{\sigma^2}$ be the ratio of the error variances and assumed known, then

$$\hat{\beta} = \frac{(s_{yy} - \lambda s_{xx}) + \left[(s_{yy} - \lambda s_{xx})^2 + 4\lambda s_{xy}^2\right]^{1/2}}{2 s_{xy}}$$

where

$$s_{xx} = \sum (x_i - \bar{x})^2,$$

s_{yy} and s_{xy} defined similarly.

4 Simulation Results

In order to evaluate the performance of each of the estimator, we carried out a simulation studies. Suppose for each data set $k, k = 1, 2, ..., n_s$, (n_s is number of simulations, that is each simulation we consider a new data set), we calculated the mean squared error (MSE), that is

$$\frac{1}{n_s} \sum (\hat{\beta} - \beta)$$

and the mean squared difference (MSD), that is

$$\frac{1}{n_s} \sum (\hat{\beta} - \hat{\beta}_{k_o})^2,$$

and their standard errors, where:

$\hat{\beta}$, estimates by various method considered in Section 3,

$\hat{\beta}_{k_o}$, estimates when λ is known (by maximum likelihood estimation), using λ equal to the value used in simulating the data, and

β, the parameter that we choose in simulating the data.

$\hat{\beta}_{k_o}$ and β are also known the baselines for comparisons.

500 simulation have been carried out from each of 3 different configuration of parameter values, we call it Dataset 1, Dataset 2 and Dataset 3. For each of them the X_i have been generated from $N(15, 16.00)$ with 50 observations, but for the analyses the X_i were regarded as fixed (but unknown) parameters. Suppose for each simulation we have:

(a) $\hat{\beta}_{k,reg}$, estimates where, $\hat{\beta}_{k,reg} = \frac{1}{2}$ {slope reg. of y on x+ slope reg. of x on y},

(b) $\hat{\beta}_{k,\lambda=1}$, estimates by assuming $\lambda = 1$. This is the most common practice and the easiest way to estimate parameters for the unreplicated functional model,

(c) $\hat{\beta}_{k,2}$, estimates by two-group method,

(d) $\hat{\beta}_{k,3}$, estimates by three-group method,

(e) $\hat{\beta}_{k,HB}$, estimates by Housner-Brennan method,

(f) $\hat{\beta}_{k,WR}$, estimates by weighted regression, and

(g) $\hat{\beta}_{k,D}$, estimates by Durbin 'ranking' method.

Simulation results for three different data set are given in Tables 1 to 3 and graphical illustration in Figures 1 to 3.

Tables 1 to 3 and Figures 1 to 3 suggest that estimate by maximum likelihood estimation assuming $\lambda = 1$ and estimate by weighted regression give smaller MSE and MSD compared to other method considered for 3 different dataset. As an example in Table 1, the MSE for method by assuming $\lambda = 1$ in MLE is 5.825×10^{-3} with standard error 3.536×10^{-4}, and that the MSE for method by weighted regression is 5.514×10^{-3} with standard error 3.349×10^{-4}. Examination of the difference in mean values relation to their standard error (for example using a t-test) indicates clearly the superiority of these estimates.

5 Conclusions

The research was motivated by the practical question on how to look at the relationship between the two variables of unreplicated data or to estimate the slope parameter, that is β, when the explanatory variable, X measured with error. Various estimates have been proposed and we are trying to compared them numerically. Based on our simulation studies we found that the method by assuming $\lambda = 1$ in maximum likelihood estimation and the method by weighted regression are favourable compared to other methods considered. We also found that the three group method also works quite well in all 3 different dataset. Although we have not dealt explicitly with estimators of α, our conclusions can be easily extended to the usual estimators of α given by $\alpha = \bar{y} - \beta\bar{x}$, where it is clear that α is consistent if β is consistent.

Table 1: Results for Dataset 1 $\alpha = 0.0, \beta = 1.2, \sigma^2 = 0.64$ and $\tau^2 = 1.44$

Estimates	**MSE**	**MSD**
$\hat{\beta}_{k,reg}$	4.964×10^{-2} (2.391×10^{-4})	5.389×10^{-2} (1.021×10^{-3})
$\hat{\beta}_{k,\lambda=1}$	5.825×10^{-3} (3.536×10^{-4})	3.028×10^{-4} (5.044×10^{-6})
$\hat{\beta}_{k,2}$	9.939×10^{-3} (5.341×10^{-4})	6.788×10^{-3} (3.202×10^{-4})
$\hat{\beta}_{k,3}$	6.695×10^{-3} (3.818×10^{-4})	4.035×10^{-3} (1.945×10^{-4})
$\hat{\beta}_{k,HB}$	1.022×10^{-2} (5.462×10^{-4})	5.642×10^{-3} (1.547×10^{-4})
$\hat{\beta}_{k,WR}$	5.514×10^{-3} (3.349×10^{-4})	2.222×10^{-4} (4.993×10^{-6})
$\hat{\beta}_{k,D}$	9.251×10^{-2} (1.193×10^{-2})	8.342×10^{-2} (1.108×10^{-2})

Table 2: Results for Dataset 2 $\alpha = 0.5, \beta = 0.8, \sigma^2 = 0.81$ and $\tau^2 = 0.64$

Estimates	**MSE**	**MSD**
$\hat{\beta}_{k,reg}$	3.196×10^{-2}	3.382×10^{-2}
	(2.158×10^{-4})	(8.469×10^{-4})
$\hat{\beta}_{k,\lambda=1}$	1.621×10^{-3}	1.982×10^{-5}
	(9.415×10^{-5})	(3.511×10^{-7})
$\hat{\beta}_{k,2}$	2.172×10^{-3}	1.475×10^{-3}
	(1.233×10^{-4})	(7.812×10^{-5})
$\hat{\beta}_{k,3}$	1.395×10^{-3}	8.011×10^{-4}
	(8.780×10^{-5})	(4.569×10^{-5})
$\hat{\beta}_{k,HB}$	2.445×10^{-3}	1.374×10^{-3}
	(1.387×10^{-4})	(4.152×10^{-5})
$\hat{\beta}_{k,WR}$	1.533×10^{-3}	1.711×10^{-5}
	(8.850×10^{-5})	(6.550×10^{-7})
$\hat{\beta}_{k,D}$	3.164×10^{-2}	2.877×10^{-2}
	(4.791×10^{-3})	(4.568×10^{-3})

Table 3: Results for Dataset 3 $\alpha = 0.0, \beta = 1.0, \sigma^2 = 0.81$ and $\tau^2 = 1.00$

Estimates	**MSE**	**MSD**
$\hat{\beta}_{k,reg}$	2.072×10^{-3} (3.623×10^{-5})	4.762×10^{-3} (2.559×10^{-4})
$\hat{\beta}_{k,\lambda=1}$	2.502×10^{-3} (1.574×10^{-4})	2.671×10^{-5} (4.890×10^{-6})
$\hat{\beta}_{k,2}$	3.220×10^{-3} (1.816×10^{-4})	2.428×10^{-3} (1.270×10^{-4})
$\hat{\beta}_{k,3}$	2.343×10^{-3} (1.404×10^{-4})	1.661×10^{-3} (9.034×10^{-5})
$\hat{\beta}_{k,HB}$	3.565×10^{-3} (1.981×10^{-4})	2.135×10^{-3} (6.484×10^{-5})
$\hat{\beta}_{k,WR}$	2.277×10^{-3} (1.433×10^{-4})	2.609×10^{-5} (9.916×10^{-7})
$\hat{\beta}_{k,D}$	4.249×10^{-2} (4.853×10^{-3})	3.814×10^{-2} (4.578×10^{-3})

References

[1] G. R. Dolby, *The ultra-structural relation: A synthesis of the functional and structural relations,* Biometrika, 63 (1976), 39–50.

[2] J. Durbin, *Errors in Variables,* Rev. Int. Statist. Inst. Vol. 22 (1954), 23–32.

[3] G. W. Housner & J. F. Brennan , *Estimation of linear trend,* Ann. Math. Statist., 19 (1948), 380–388.

[4] A. G. Hussin, *Pseudo-replication in functional relataionships with environmental application,* PhD Thesis, University of Sheffield, 1997.

[5] A. G. Hussin, *Application of linear circular functional model in remote sensing data,* Proceding International Conference of ECM (1999).

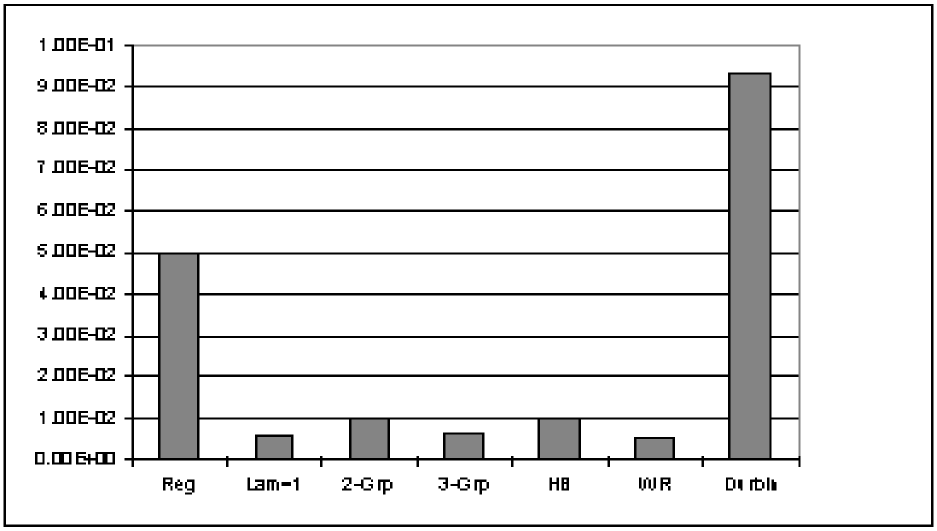
[6] A. G. Hussin, *A possible application of EIVM in environmental science,* Menemui Matematik, 20 (1998), 28–35.

[7] M. G. Kendall, *Regression: Structure and functional relationship I,* Biometrika 38 (1951), 11–25.

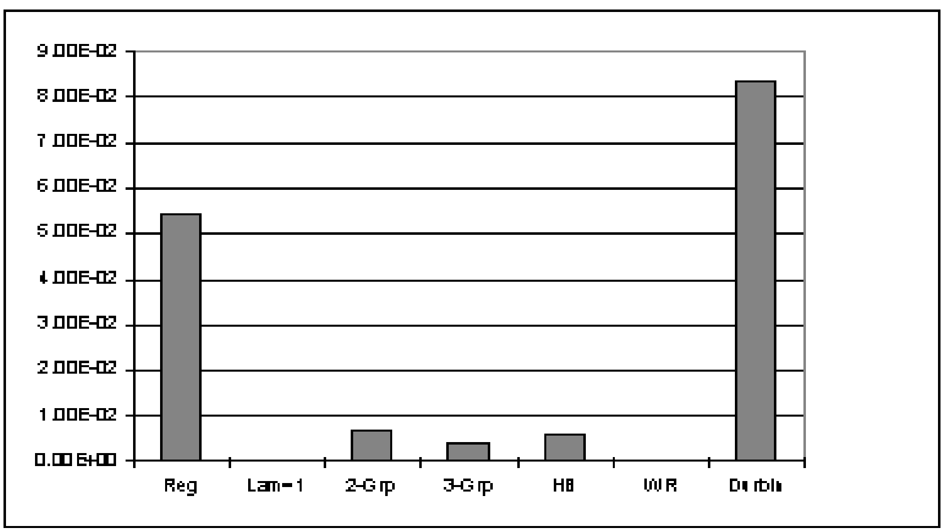
[8] M. G. Kendall, *Regression: Structure and functional relationship II,* Biometrika 39 (1952), 96–108.

[9] V. D. Lindley, *Regression lines and the linear functional relationship,* J. Roy. Statistical Soc. Supp. 9 (1947), 218–244.

[10] A. Wald, *The fitting of straight lines if both variables are subject to error,* Ann. Math. Statist., 11 (1940), 284–300.

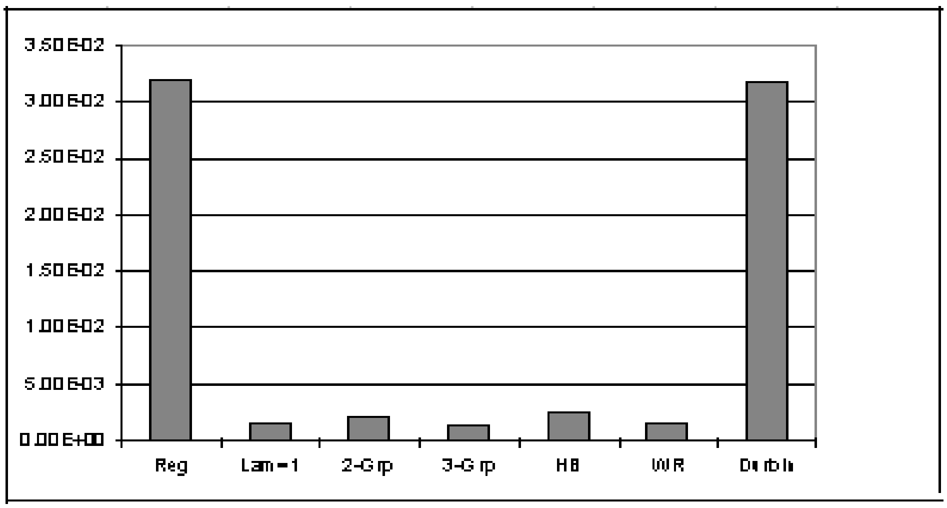


a) Observed mean square errors, i.e. $\frac{1}{500}\sum(\hat{\beta}-\beta)^2$

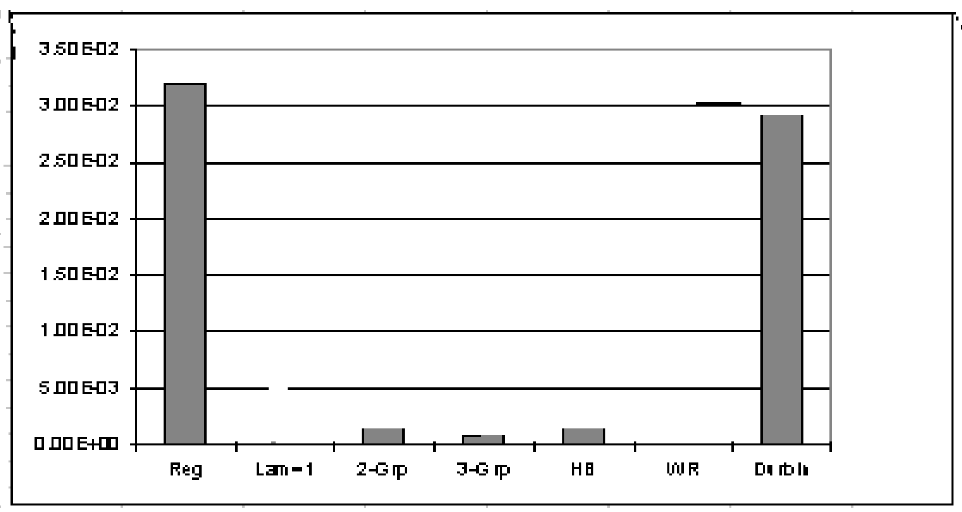


b) Observed mean square differences, i.e. $\frac{1}{500}\sum(\hat{\beta}-\hat{\beta}_{k_0})^2$

Figure 1: Comparisons for Dataset 1, $\alpha = 0.0$, $\beta = 1.2$, $\sigma^2 = 0.64$ and $\tau^2 = 1.44$.

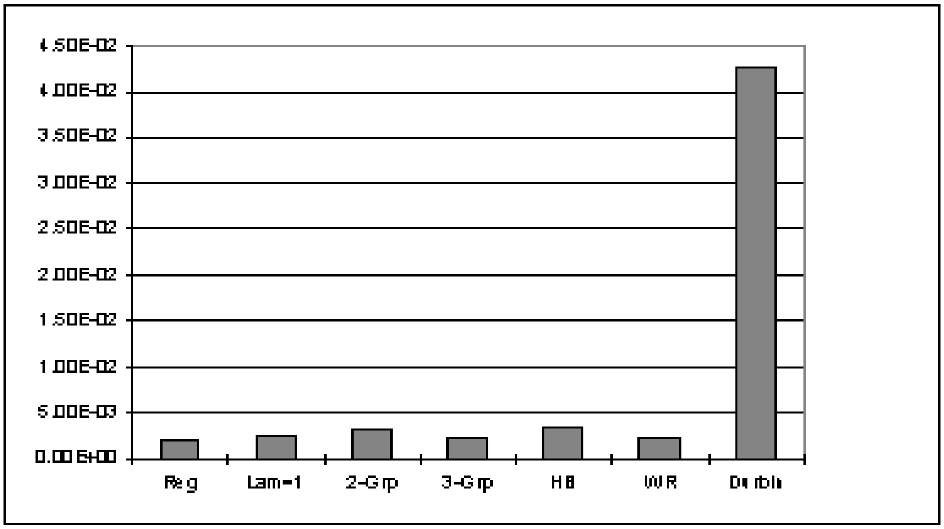


a) Observed mean square errors, i.e. $\frac{1}{500}\sum(\hat{\beta}-\beta)^2$

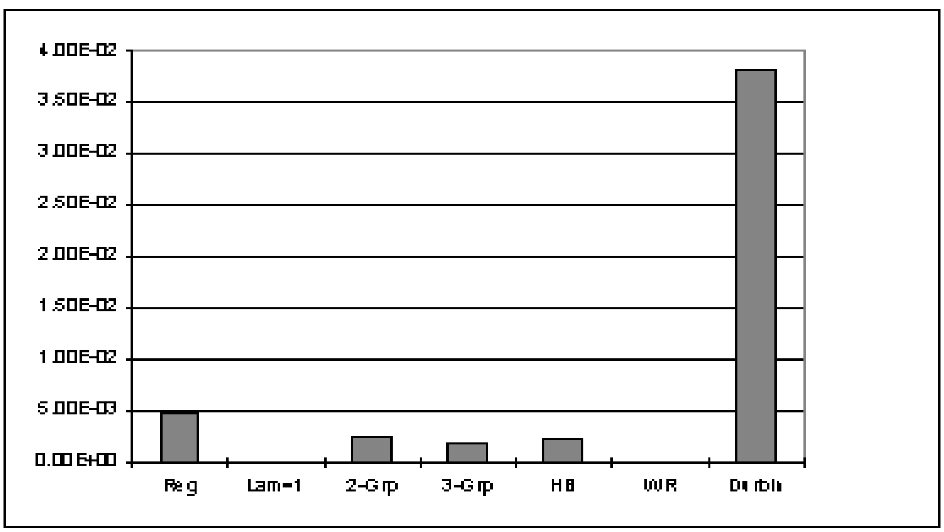


b) Observed mean square differences, i.e. $\frac{1}{500}\sum(\hat{\beta}-\hat{\beta}_{k_0})^2$

Figure 2: Comparisons for Dataset 1, $\alpha = 0.5$, $\beta = 0.8$, $\sigma^2 = 0.81$ and $\tau^2 = 0.64$.



a) Observed mean square errors, i.e. $\frac{1}{500}\sum(\hat{\beta}-\beta)^2$



b) Observed mean square differences, i.e. $\frac{1}{500}\sum(\hat{\beta}-\hat{\beta}_{k_0})^2$

Figure 3: Comparisons for Dataset 1, $\alpha = 0.0$, $\beta = 1.0$, $\sigma^2 = 0.81$ and $\tau^2 = 1.00$.