IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 1, January 2020
ISSN (Online) 2348 – 7968 | Impact Factor (2019) – 6.248
www.ijiset.com

Dealing Heteroscedasticity Problem in Regression Modeling Using ML-Fisher Scoring Algorithm: Simulation Study

I Gede Nyoman Mindra Jaya [1*], Neneng Sunengsih[2]

[1] Department Statistics, Padjadjaran University, Indonesia

[2] Department Statistics, Padjadjaran University, Indonesia

*email: mindra@unpad.ac.id

Abstract
The violation of the homoscedasticity assumption might lead to a misleading conclusion in regression modeling. When the variance error is a function of independent variables, the regression model may not be accurate, and the hypothesis testing might be invalid. We evaluated the maximum likelihood estimator using the fisher scoring approach, which is usually used to handle the heteroscedasticity problem in regression modeling. We also evaluate the performance of the ordinary least square estimator to see the serious effect of the heteroskedasticity problem on the accuracy, precision, and power of parameter estimation using the Monte Carlo simulation study. The maximum likelihood with the fisher scoring approach gives outperform estimates than the ordinary least square.

Keywords: *Heteroscedasticity, Fisher Scoring, Ordinary least square, Maximum likelihood,*

1. Introduction

Homoskedasticity is one of Gauss Markov's assumptions that usually hard to be fulfilled for the real data. This assumption assumes that the error term for each unit of the observation has a similar variance ([1] [2] [3]). However, because of some conditions such as data collecting problem, outlier, omitted variable, and model specification may cause heteroscedasticity problem [2].

The violation of the homoscedasticity assumption because of some conditions which may lead to the inefficiency of the estimator. This assumption may not affect unbiased and consistent properties of maximum likelihood or ordinary least square estimators ([2]). Although for some condition, the heteroscedasticity problem does not have consequences on unbiased properties, this condition can produce a wrong regression model. However, in the case of omitted variables, and those variables have a strong correlation with the error term, the heteroscedasticity problem may lead to bias estimates and strongly influence the inference of the regression parameters. It might produce the misleading conclusion of the hypothesis testing. This condition shows that we have to pay attention to the violation of the homoscedasticity assumption to get the better model.

Many authors have considered variance modeling to have a correct standard error and confidence interval for mean in regression modeling [2]. There are several solutions have been proposed when the heteroscedasticity problem is related to the independent variables [1]. The variance error σ_i^2 is modeled as (i) $\sigma_i^2 = (\boldsymbol{z}_i'\boldsymbol{\gamma})$, (ii) $\sigma_i^2 = (\boldsymbol{z}_i'\boldsymbol{\gamma})^2$, (iii) $\sigma_i^2 = (\boldsymbol{z}_i'\boldsymbol{\gamma})^p$, (iv) $\sigma_i^2 = \exp\big((\boldsymbol{z}_i'\boldsymbol{\gamma})\big)$. In case (iv), the log of the variance is a linear function of independent variables, which leads to a multiplicative heteroscedastic model [2]. There is the other way to model heteroscedasticity. However, this paper is focused on the model (i), (ii) and (iv).

In regression modeling, there is some estimator that usually used, such as ordinary least square (OLS), Maximum Likelihood, Bayesian method and machine learning ([4] [5] [6] [7]). In this paper, we evaluate the biases and power tests of the first two methods to regression parameter inference.

The structure of the remainder of this paper is as follows. Section 2 presents the methodology. Section 3 applies simulation design. Section 4 presents the result and discussion, and section 5 presents the conclusions.

2. Methodology

2.1. Modeling heteroscedasticity problem

Assume we are modeling linear regression model with heteroscedasticity problem as below:

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i;\ \varepsilon_i \sim N(0, \sigma_i^2), i = 1, \ldots, n \quad (1)$$

where y_i denotes the response variable for *i*-th unit observation, $\mathbf{x}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iK})'$ denotes the K vector of the covariates with regression coefficient $\boldsymbol{\beta} =$

ISSN (Online) 2348 – 7968 | Impact Factor (2019) – 6.248
www.ijiset.com

$(\beta_0, \beta_1, \dots, \beta_K)'$. The error term ε_i is assumed follows normal distribution with un-constant variance. The variance σ_i^2 defined as $\sigma_i^2 = g(\boldsymbol{z}_i, \boldsymbol{\gamma})$ with $\boldsymbol{z}_i = (\boldsymbol{z}_{i1}, \dots, \boldsymbol{z}_{iL})$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_L)$ are the others covariates and its regression coefficients on the σ_i^2 with $g(.)$ is appropriate real function.

2.2. Ordinary least square (OLS)

OLS is the simple estimator that usually used to estimate regression parameters. The principal of OLS method is minimize the sum square error; $\widehat{\boldsymbol{\beta}} = \arg.\min(\sum_i^n \varepsilon_i^2)$. The OLS estimator is [8]:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} \tag{2}$$

OLS is the Best Linear Unbiased Estimator when all Gauss Markov assumption is satisfied. However, the violation of the homoskedasticity assumption leads to the inefficiency of parameter estimate and inconsistency of the standard error estimate. The inconsistency of the covariance matrix of the estimated regression coefficients, the tests of hypotheses, (t-test, F-test) are no longer valid. The weighted least square regression (WLS) to overcome the heteroscedasticity problem. However, defining the weight matrix becomes a new problem.

2.3. Maximum likelihood estimation using Fisher scoring algorithm

Maximum likelihood algorithm is a flexible algorithm to be used estimate the regression parameter when the heteroscedastic problem is found. The likelihood function is given by [9]:

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma}|\mathbf{y}) \approx \prod_{i=1}^{n} \left[\frac{1}{(\sigma_i^2)^{1/2}} \times \exp\left(-\frac{1}{2\sigma_i^2}(\mathrm{y_i} - \mathbf{x}_\mathrm{i}'\boldsymbol{\beta})^2\right)\right] \tag{3}$$

and the log likelihood function can be written as:

$$\ln L(\boldsymbol{\beta}, \boldsymbol{\gamma}|\mathbf{y}) = -\frac{1}{2}\sum_{i=1}^{n} \ln \sigma_i^2 - \frac{1}{2}\sum_{i=1}^{n} \frac{1}{\sigma_i^2}(\mathrm{y_i} - \mathbf{x}_\mathrm{i}'\boldsymbol{\beta})^2 \tag{4}$$

The parameter estimate of $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma})'$ are obtained by $\widehat{\boldsymbol{\theta}} = \mathrm{argmax}(\ln L(\boldsymbol{\beta}, \boldsymbol{\gamma}|\mathbf{y}))$. However, there is no analytical solution for this function. Optimization method such as fisher scoring method can be used to obtain the parameter estimate of $\boldsymbol{\theta}$.

The Fisher scoring method is based on the bloc diagonal Fisher information matrix to get maximum likelihood of the parameters interest [1]. An iterative algorithm of Fisher scoring method is given below [2]:

a) Given the initial value of $\boldsymbol{\beta}^{(\mathbf{0})}$ and $\boldsymbol{\gamma}^{(\mathbf{0})}$ for the parameter. Usually we choose parameter estimate from OLS regression model.

b) $\boldsymbol{\beta}^{(\boldsymbol{k}+\mathbf{1})}$ is obtained from $\boldsymbol{\beta}^{(k+1)} = \left(\mathbf{x}'\boldsymbol{W}^{(\boldsymbol{k})}\mathbf{x}\right)^{-\mathbf{1}}\mathbf{x}'\boldsymbol{W}^{(\boldsymbol{k})}\mathbf{y}$, where $\boldsymbol{W}^{(\boldsymbol{k})} = \mathrm{diag}\left(w_i^{(k)}\right), w_i^{(k)} = 1/(\sigma_i^2)^{(k)}$ and $(\sigma_i^2)^{(k)} = \exp\left(\boldsymbol{z}_i'\boldsymbol{\gamma}^{(k)}\right)$

c) $\boldsymbol{\gamma}^{(k+1)}$ is obtained from $\boldsymbol{\gamma}^{(k+1)} = \left(\mathbf{z}'\boldsymbol{W}^{(\boldsymbol{k})}\mathbf{z}\right)^{-\mathbf{1}}\mathbf{z}'\boldsymbol{W}^{(\boldsymbol{k})}\tilde{\mathbf{y}}$, where $\boldsymbol{W} = \frac{1}{2}\mathbf{I}_n$, with $\mathbf{I}_n$ the $n \times n$ identity matrix, and $\tilde{\mathbf{y}}$ is a n-dimensional vector with i-th component:

$$\tilde{\mathrm{y}}_i = \eta_i + \frac{1}{\sigma_i^2}(\mathrm{y_i} - \mathbf{x}_\mathrm{i}'\boldsymbol{\beta})^2 - 1 \tag{5}$$

where $\eta_i = g(z_i'\boldsymbol{\gamma})$

d) Steps (b) and (c) will be repeated iteratively until the pre-specified stopping criterion is satisfied.

3. Simulation Design

We develop Monte Carlo simulation study to evaluate the bias and power of test of OLS, ML, and Bayesian approach in inferencing regression coefficient when the heterogeneity problem is found. There are two type heteroscedasticity functions were considered: $g(z_i'\boldsymbol{\gamma}) = (z_i'\boldsymbol{\gamma})^p = (\mathrm{x}_1)^p; p = \{-2, -1, -0, 1, 2\}$. The systematic components for these models is defined follows Cordeiro (2008) [9].

$$\mu = \beta_0 + \beta_1 \mathrm{x}_1 \tag{6}$$

and for the variance

$$\sigma^2 = x_1^p \tag{7}$$

The true value of the parameters for the simulations were taken as $\beta_0 = 1$, and $\beta_1 = 1$. We generated the independent variables from Uniform distribution (0,1) and we assume that all independents are fixed throughout the simulation with equal sample size. The number of observation was set at $n = 10, 20, 30, 50, 100, 300$ and 500 with $m = 1000$ number of iterations for each cases. The calculations were performed using R-statistical package.

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 1, January 2020
ISSN (Online) 2348 – 7968 | Impact Factor (2019) – 6.248
www.ijiset.com

4. Result and Discussion

Table 1-2 give the sample means, power test, bias and mean square error of parameter estimate with ordinary least square and maximum likelihood estimator respectively for different number of sample (n) and power (p) of heteroscedasticity.

At first glance, it can be seen that for the estimated parameter has a negative bias or tends to underestimate while for the beta parameter estimate one has a positive bias or tends to overestimate.

Table 1. OLS Estimation

n	p	β_0	β_1	Power β_0	Power β_1	Bias β_0	Bias β_1	RMSEA β_0	RMSEA β_1
10	-2	0.80	1.28	0.15	0.11	-0.20	0.28	0.20	0.28
10	-1.5	0.99	1.01	0.28	0.21	-0.01	0.01	0.01	0.01
10	-1	1.03	0.95	0.07	0.02	0.03	-0.05	0.03	0.05
10	-0.5	1.01	0.97	0.28	0.14	0.01	-0.03	0.01	0.03
10	0	1.01	0.98	0.41	0.12	0.01	-0.02	0.01	0.02
10	0.5	1.00	1.01	0.40	0.19	0.00	0.01	0.00	0.01
10	1	0.99	1.02	0.46	0.18	-0.02	0.02	0.02	0.02
10	1.5	1.01	0.96	0.60	0.29	0.01	-0.04	0.01	0.04
10	2	0.99	1.02	0.62	0.25	-0.01	0.02	0.01	0.02
20	-2	1.14	0.78	0.19	0.07	0.14	-0.22	0.14	0.22
20	-1.5	0.86	1.22	0.33	0.28	-0.14	0.22	0.14	0.22
20	-1	1.03	0.97	0.30	0.18	0.03	-0.03	0.03	0.03
20	-0.5	0.98	1.04	0.39	0.23	-0.02	0.04	0.02	0.04
20	0	0.97	1.04	0.49	0.27	-0.03	0.04	0.03	0.04
20	0.5	1.00	1.01	0.58	0.23	0.00	0.01	0.00	0.01
20	1	1.01	0.96	0.88	0.30	0.01	-0.04	0.01	0.04
20	1.5	1.00	1.00	0.94	0.48	0.00	-0.01	0.00	0.01
20	2	1.00	0.99	0.90	0.41	0.00	-0.01	0.00	0.01
30	-2	0.67	1.45	0.31	0.17	-0.33	0.45	0.33	0.45
30	-1.5	0.72	1.41	0.35	0.28	-0.28	0.41	0.28	0.41
30	-1	1.06	0.92	0.42	0.29	0.06	-0.08	0.06	0.08
30	-0.5	1.01	0.99	0.44	0.27	0.01	-0.01	0.01	0.01
30	0	1.00	1.01	0.72	0.39	0.00	0.01	0.00	0.01
30	0.5	1.00	1.01	0.94	0.44	0.00	0.01	0.00	0.01
30	1	0.99	1.02	0.96	0.54	-0.01	0.02	0.01	0.02
30	1.5	1.00	1.00	0.98	0.63	0.00	0.00	0.00	0.00
30	2	1.00	0.99	1.00	0.74	0.00	-0.01	0.00	0.01
50	-2	0.91	1.13	0.44	0.30	-0.09	0.13	0.09	0.13
50	-1.5	1.08	0.88	0.43	0.29	0.08	-0.12	0.08	0.12
50	-1	1.02	0.97	0.44	0.21	0.02	-0.03	0.02	0.03
50	-0.5	1.02	0.96	0.65	0.35	0.02	-0.04	0.02	0.04
50	0	1.01	0.98	0.94	0.53	0.01	-0.02	0.01	0.02
50	0.5	1.00	1.00	1.00	0.78	0.00	0.00	0.00	0.00
50	1	1.00	1.00	1.00	0.75	0.00	0.00	0.00	0.00
50	1.5	1.01	0.99	1.00	0.85	0.01	-0.01	0.01	0.01
50	2	1.00	1.00	1.00	0.91	0.00	0.00	0.00	0.00
100	-2	1.01	0.98	0.35	0.25	0.01	-0.03	0.01	0.03
100	-1.5	0.77	1.35	0.42	0.27	-0.23	0.35	0.23	0.35
100	-1	0.99	1.00	0.55	0.35	-0.01	0.00	0.01	0.00
100	-0.5	0.99	1.02	0.82	0.52	-0.02	0.02	0.02	0.02
100	0	0.99	1.01	1.00	0.80	-0.01	0.01	0.01	0.01
100	0.5	1.00	1.00	1.00	0.97	0.00	-0.01	0.00	0.01
100	1	1.00	1.00	1.00	0.97	0.00	0.00	0.00	0.00
100	1.5	1.00	0.99	1.00	0.99	0.00	-0.01	0.00	0.01
100	2	1.00	0.99	1.00	0.99	0.00	-0.01	0.00	0.01
300	-2	1.10	0.85	0.39	0.27	0.10	-0.15	0.10	0.15
300	-1.5	1.00	1.00	0.48	0.37	0.00	0.00	0.00	0.00
300	-1	0.99	1.01	0.87	0.57	-0.01	0.01	0.01	0.01

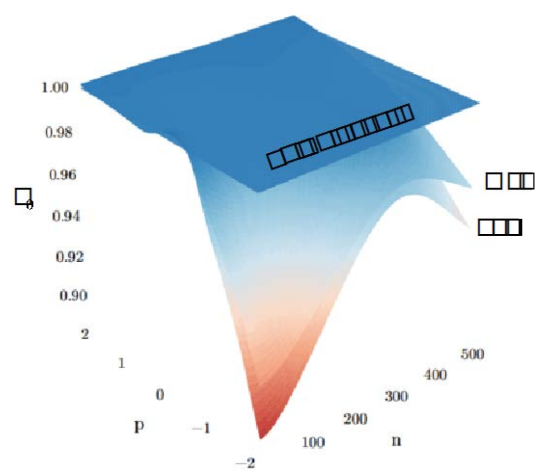
n	p	β_0	β_1	Power β_0	Power β_1	Bias β_0	Bias β_1	RMSEA β_0	RMSEA β_1
300	-0.5	0.99	1.02	1.00	0.93	-0.01	0.02	0.01	0.02
300	0	1.00	0.99	1.00	1.00	0.00	-0.01	0.00	0.01
300	0.5	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
300	1	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
300	1.5	1.00	0.99	1.00	1.00	0.00	-0.01	0.00	0.01
300	2	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	-2	0.82	1.27	0.38	0.28	-0.18	0.27	0.18	0.27
500	-1.5	1.00	1.00	0.45	0.28	0.00	0.00	0.00	0.00
500	-1	0.99	1.02	0.91	0.67	-0.01	0.02	0.01	0.02
500	-0.5	1.00	1.00	1.00	0.99	0.00	0.00	0.00	0.00
500	0	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	0.5	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	1	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	1.5	1.00	1.00	1.00	1.00	0.00	-0.01	0.00	0.01
500	2	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00

Table 2. ML-Fisher Scoring Estimation

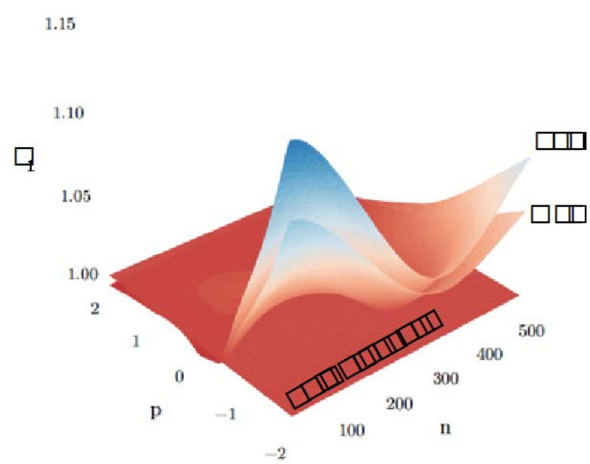
n	p	β_0	β_1	Power β_0	Power β_1	Bias β_0	Bias β_1	RMSEA β_0	RMSEA β_1
10	-2.00	0.84	1.22	0.34	0.43	-0.16	0.22	0.16	0.22
10	-1.50	0.99	1.00	0.34	0.49	-0.01	0.00	0.01	0.00
10	-1.00	1.03	0.95	0.17	0.34	0.03	-0.05	0.03	0.05
10	-0.50	1.00	0.98	0.32	0.40	0.00	-0.03	0.00	0.03
10	0.00	1.01	0.98	0.47	0.51	0.01	-0.02	0.01	0.02
10	0.50	1.00	1.01	0.41	0.45	0.00	0.01	0.00	0.01
10	1.00	0.99	1.01	0.38	0.45	-0.01	0.01	0.01	0.01
10	1.50	1.00	0.97	0.59	0.52	0.00	-0.03	0.00	0.03
10	2.00	1.00	1.01	0.50	0.52	0.00	0.01	0.00	0.01
20	-2.00	1.10	0.84	0.23	0.35	0.10	-0.16	0.10	0.16
20	-1.50	0.90	1.16	0.33	0.48	-0.10	0.16	0.10	0.16
20	-1.00	1.03	0.97	0.36	0.43	0.03	-0.03	0.03	0.03
20	-0.50	0.98	1.04	0.41	0.46	-0.02	0.04	0.02	0.04
20	0.00	0.97	1.04	0.49	0.51	-0.03	0.04	0.03	0.04
20	0.50	1.00	1.01	0.52	0.53	0.00	0.01	0.00	0.01
20	1.00	1.01	0.97	0.82	0.62	0.01	-0.03	0.01	0.03
20	1.50	1.00	1.00	0.89	0.71	0.00	-0.01	0.00	0.01
20	2.00	1.00	0.99	0.73	0.59	0.00	-0.01	0.00	0.01
30	-2.00	0.75	1.32	0.22	0.40	-0.25	0.32	0.25	0.32
30	-1.50	0.78	1.31	0.33	0.50	-0.22	0.31	0.22	0.31
30	-1.00	1.04	0.94	0.43	0.50	0.04	-0.06	0.04	0.06
30	-0.50	1.01	0.99	0.45	0.50	0.01	-0.01	0.01	0.01
30	0.00	1.00	1.01	0.71	0.65	0.00	0.01	0.00	0.01
30	0.50	1.00	1.01	0.93	0.77	0.00	0.01	0.00	0.01
30	1.00	0.99	1.02	0.93	0.80	-0.01	0.02	0.01	0.02
30	1.50	1.00	1.00	0.99	0.87	0.00	0.00	0.00	0.00
30	2.00	1.00	0.99	1.00	0.94	0.00	-0.01	0.00	0.01
50	-2.00	0.93	1.09	0.31	0.55	-0.07	0.09	0.07	0.09
50	-1.50	1.06	0.91	0.39	0.49	0.06	-0.09	0.06	0.09
50	-1.00	1.02	0.97	0.45	0.47	0.02	-0.03	0.02	0.03
50	-0.50	1.01	0.96	0.68	0.59	0.01	-0.04	0.01	0.04
50	0.00	1.01	0.98	0.93	0.82	0.01	-0.02	0.01	0.02
50	0.50	1.00	1.00	1.00	0.95	0.00	0.00	0.00	0.00
50	1.00	1.00	1.00	1.00	0.94	0.00	0.00	0.00	0.00
50	1.50	1.01	0.99	1.00	0.97	0.01	-0.01	0.01	0.01
50	2.00	1.00	1.00	1.00	0.99	0.00	0.00	0.00	0.00
100	-2.00	1.01	0.98	0.25	0.41	0.01	-0.02	0.01	0.02
100	-1.50	0.83	1.24	0.31	0.53	-0.17	0.24	0.17	0.24
100	-1.00	0.98	1.01	0.58	0.58	-0.02	0.01	0.02	0.01

n	p	β_0	β_1	Power β_0	Power β_1	Bias β_0	Bias β_1	RMSEA β_0	RMSEA β_1
100	-0.50	0.99	1.02	0.84	0.77	-0.02	0.02	0.02	0.02
100	0.00	0.99	1.01	1.00	0.95	-0.01	0.01	0.01	0.01
100	0.50	1.00	1.00	1.00	1.00	0.00	-0.01	0.00	0.01
100	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
100	1.50	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
100	2.00	1.00	0.99	1.00	1.00	0.00	-0.01	0.00	0.01
300	-2.00	1.06	0.92	0.23	0.44	0.06	-0.08	0.06	0.08
300	-1.50	1.00	1.00	0.48	0.56	0.00	0.00	0.00	0.00
300	-1.00	0.99	1.01	0.91	0.83	-0.01	0.01	0.01	0.01
300	-0.50	0.99	1.02	1.00	0.99	-0.01	0.02	0.01	0.02
300	0.00	1.00	0.99	1.00	1.00	0.00	-0.01	0.00	0.01
300	0.50	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
300	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
300	1.50	1.00	0.99	1.00	1.00	0.00	-0.01	0.00	0.01
300	2.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	-2.00	0.88	1.17	0.23	0.43	-0.12	0.17	0.12	0.17
500	-1.50	1.00	0.99	0.30	0.58	0.00	-0.01	0.00	0.01
500	-1.00	0.99	1.02	0.95	0.89	-0.01	0.02	0.01	0.02
500	-0.50	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	0.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	0.50	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	1.50	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
500	2.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00

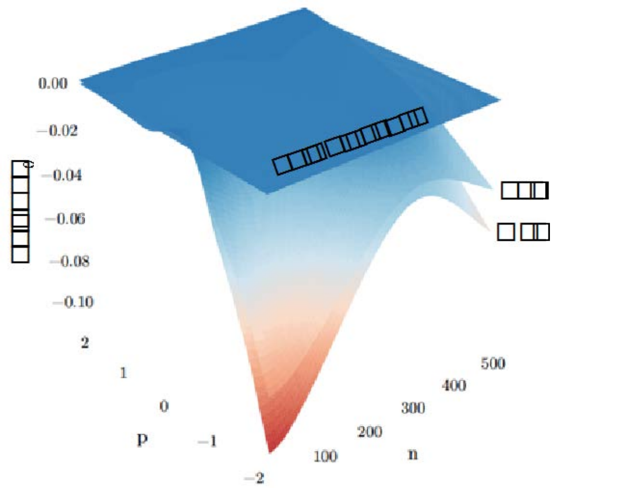
To provide a clear presentation of Table 1-2 we present the 3-D plot using smoothing loess function in Figure 1.
Figure 1 present the detail information of the parameter estimate, bias, power test, and root mean square error for different number of sample and power of heteroscedasticity



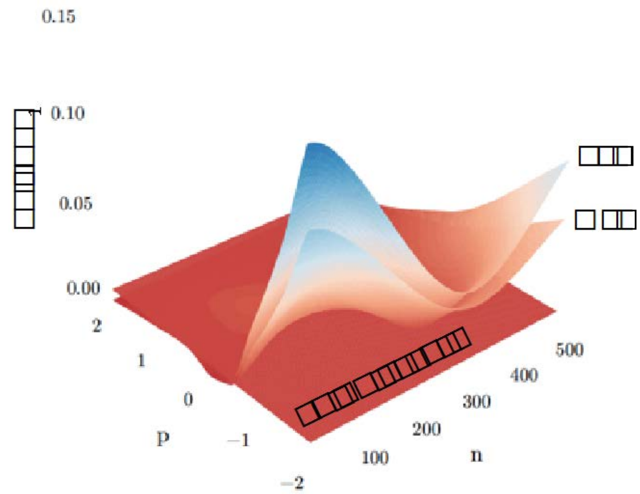
(a) Parameter estimate of β_0



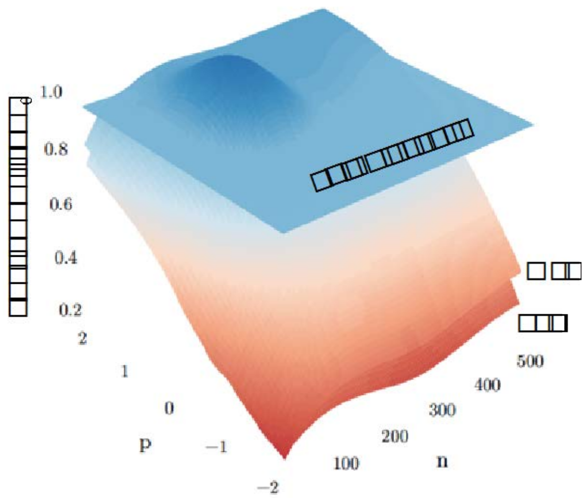
(b) Parameter estimate of β_1



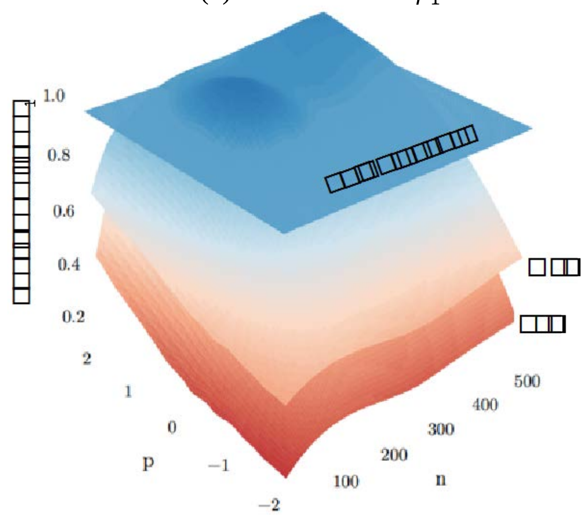
(c) Bias estimate of β_0



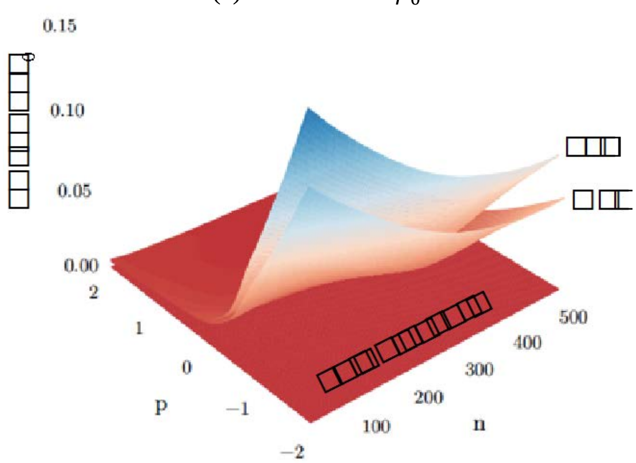
(d) Bias estimate of β_1



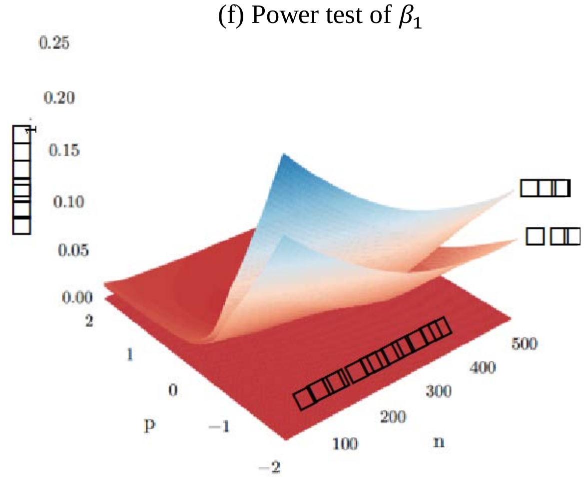
(e) Power test of β_0



(f) Power test of β_1



(g) Power test of β_0



(h) Power test of β_1

Figure 1. Simulation result: Parameter estimate, Bias, Power test and Root Mean Square Error

Figure 1 presents the detail of the comparison of ordinary least square and maximum likelihood for heteroscedasticity problem using Fisher scoring approach. We present the parameters estimate, bias, power test, and root mean square error for different number of samples and different power of heteroscedasticity function.

We can see that the estimate of parameter intercept β_0 tend to underestimate while for β_1 tend to overestimate. For negative power (p), the effect of heteroscedasticity on parameters estimate, bias, power test, and root mean square error is getting worst even though the number of sample increases.

In general, the maximum likelihood with Fisher scoring method provides a better result compare that the ordinary least square.

IV. Conclusion

The violation of the homoscedasticity assumption might lead to a misleading conclusion in regression modeling. When the variance error is a function of independent variables, the regression model may not be accurate, and the hypothesis testing might be invalid. We evaluated ordinary least square which is not dealing with heteroscedasticity problem and maximum likelihood with fisher scoring which is dealing with heteroscedasticity problem. Monte Carlo simulation result shows that the the maximum likelihood with Fisher scoring method provides a better result compare that the ordinary least square. For negative power (p), the effect of heteroscedasticity on parameters estimate, bias, power test, and root mean square error is getting worst even though the number of sample increases.

Acknowledgments

The authors thank to the Rector Universitas Padjadjaran for facilitating this research.

Reference

[1] M. Aitkin, "Modelling Variance Heterogeneity in Normal Regression Using GLIM," *Journal of the Royal Statistical Society,* vol. 36, no. 3, pp. 332-339, 1987.

[2] E. Cepeda and D. Gamerman, "Bayesian Modeling of Variance Heterogeneity in Normal Regression Models," *Brazilian Journal of Probability and Statistics,* vol. 14, no. 2, pp. 207-221, 2000.

[3] A. Sen and M. Srivastava, Regression Analysis Theory, Methods, and Applications, New York: Springer, 1990.

[4] M. Blangiardo and M. Cameletti, Saptial and Spatio-Temporal Bayesian Model with R-INLA, John Wiley: United Kingdom, 2015.

[5] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vols. PAMI-6, no. 6, pp. 721-741, 1984.

[6] E. C. Cuervo and J. A. Achcar, "Regression Models with Heteroscedasticity using Bayesian Approach," *Revista Colombiana de Estadística,* vol. 32, no. 2, pp. 267-287, 2009.

[7] W. Wiedermann, R. Artner and A. v. Eye, "Heteroscedasticity as a Basis of Direction Dependence in Reversible Linear Regression Models," *Multivariate Behavioral Research,* vol. 0, no. 0, pp. 1-20, 2017.

[8] G. Jude, Griffiths, C. Hill, H. Lutkepohl and T.-C. Lee, The Theory and Practice of Econometrics, Canada: 1985, John Wiley and Son.

[9] G. M. Cordeiro, "Corrected Maximum Likelihood Estimators in Linear Heteroskedastic Regression Models," *Brazilian Review of Econometrics,* vol. 28, no. 1, pp. 1-16, 2008.