Properties of the coefficient estimators for the linear regression model with heteroskedastic error term

Alfredas RAČKAUSKAS, Danas ZUOKAS (VU)

e-mail: alfredas.rackauskas@maf.vu.lt, danas.zuokas@mif.vu.lt

Abstract. In this paper we present estimated generalized least squares (EGLS) estimator for the coefficient vector β in the linear regression model $y = \beta X + \varepsilon$, where disturbance term can be heteroskedastic. For the heteroskedasticity of the changed segment type, using Monte-Carlo method, we investigate empirical properties of the proposed and ordinary least squares (OLS) estimators. The results show that the empirical covariance of the EGLS estimators is smaller than that of OLS estimators.

Keywords: heteroskedasticity, changed segment, Hölder norm tests.

1. Regression model coefficient estimators

Classical results in the econometric theory show that if coefficients of the linear regression model are estimated by the OLS method when error term is not homoskedastic, an estimator is obtained which, although remains unbiased, consistent and asymptotically normal, is no longer minimum variance unbiased estimator (MVUE). Thus diagnostic testing for heteroskedasticity has to be undertaken before further analysis starts. If the null hypothesis of homoskedasticity is rejected, a generalized estimator should be found.

We consider a linear regression model

$$y_j = \boldsymbol{f}^{\mathrm{T}}(j/n)\boldsymbol{\beta} + \varepsilon_j, \quad j = 1, \ldots, n,$$

where $\boldsymbol{\beta}$ is a coefficient vector of length d and $\boldsymbol{f} \colon [0, 1] \to \mathbb{R}^d$ is a given function,

$$\boldsymbol{f}(t) = \big(f_1(t), \ldots, f_d(t)\big)^{\mathrm{T}}, \quad t \in [0, 1],$$

disturbances are

$$\varepsilon_j = g(j/n)u_j,$$

with i.i.d. random variables u_j, $\mathrm{E}u_1 = 0$, $\mathrm{E}u_1^2 = 1$, and the function $g \colon [0, 1] \to \mathbb{R}$; $^{\mathrm{T}}$ denotes transposition operation. The null hypothesis H_0 is specified by $g(j/n) \equiv \sigma^2$. All other choices of g lead to heteroskedastic alternatives.

Set

$$X = \big(\boldsymbol{f}(1/n), \boldsymbol{f}(2/n), \ldots, \boldsymbol{f}(n/n)\big)^{\mathrm{T}}$$

And denote $\boldsymbol{y} = (y_1, \ldots, y_n)^{\mathrm{T}}$.

Let us assume now that the alternative hypothesis is true, i.e., $\mathrm{E}\varepsilon_j^2 = g^2(j/n) \neq \sigma^2$ for some $j = 1, \ldots, n$. It is well known (see [1], for example) that regression coefficients estimated by OLS method,

$$\widehat{\boldsymbol{\beta}} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\boldsymbol{y}, \tag{1}$$

though are unbiased, have the covariance matrix which is no longer $\sigma^2(X^{\mathrm{T}}X)^{-1}$. If we denote covariance matrix of $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$ by

$$G = \mathrm{diag}\left(g^2(1/n) \ldots, g^2(n/n)\right), \tag{2}$$

we obtain generalized least squares (GLS) estimator

$$\widehat{\boldsymbol{\beta}}_G = (X^{\mathrm{T}}G^{-1}X)^{-1}X^{\mathrm{T}}G^{-1}\boldsymbol{y}. \tag{3}$$

Even more if we replace G by some of its estimator $\widehat{G}$ we obtain estimated GLS estimator

$$\widehat{\boldsymbol{\beta}}_{EG} = (X^{\mathrm{T}}\widehat{G}^{-1}X)^{-1}X^{\mathrm{T}}\widehat{G}^{-1}\boldsymbol{y}. \tag{4}$$

It is also known that $\widehat{\boldsymbol{\beta}}_G$ and $\widehat{\boldsymbol{\beta}}_{EG}$ under regularity conditions possess at least asymptotically consistency, normality and are more efficient than $\widehat{\boldsymbol{\beta}}$.

2. Heteroskedasticity of the changed segment type

In this paper we are especially interested in the so called changed segment heteroskedasticity alternative, which can be formulated as follows:

H_1 There exist integers ℓ^* and k^* and a real $\tau \neq \sigma$ such that $g(j/n) = \tau$ if $j \in \{k^*+1, \ldots, k^*+\ell^*\}$ and $g(j/n) = \sigma$ if $j \in \{1, \ldots, n\} \setminus \{k^*+1, \ldots, k^*+\ell^*\}$.

In [2], for this type of alternative a class of test statistics is presented, which are obtained by taking a certain functional (defined in a Hölder space) of the polygonal line process constructed the partial sums of square residuals of the model. Without going much into detail assume that the residuals $\widehat{\boldsymbol{\varepsilon}} = (\widehat{\varepsilon}_1, \ldots, \widehat{\varepsilon}_n)^{\mathrm{T}}$ are defined by

$$\widehat{\varepsilon}_j = y_j - \boldsymbol{f}^{\mathrm{T}}(j/n)\widehat{\boldsymbol{\beta}}, \quad j = 1, \ldots, n,$$

with OLS estimator $\widehat{\boldsymbol{\beta}}$ defined in (1). For $0 \leqslant \alpha < 1/2$, we introduce a class of test statistics

$$T_{n,\alpha} = \max_{1 \leqslant \ell < n/2} (\ell/n)^{-\alpha} \max_{0 \leqslant k < n} \left| \sum_{j=k+1}^{k+\ell} \left(\widehat{\varepsilon}_j^2 - \frac{1}{n}\sum_{j=1}^{n} \widehat{\varepsilon}_j^2 \right) \right|.$$

Let $(W_t, t \in [0,1])$ be a standard Wiener process and let $(B_t, t \in [0,1])$ be a Brownian bridge $B_t = W_t - tW_1$, $t \in [0,1]$. The class of limiting test statistics then is

$$T_\alpha = \sup_{0 < h < 1/2} h^{-\alpha} \sup_{0 \leqslant t < 1} |B_{t+h} - B_t|, \quad 0 \leqslant \alpha < 1/2. \tag{5}$$

In this paper we will rely on theoretical results (two theorems) stated and proved in [2]. The first theorem shows that a class of test statistics $T_{n,\alpha}$ converges in distribution to T_α and the second shows that this class is consistent for the heteroskedasticity alternatives of various types. Here as a special case we will deal only with the changed segment type alternative.

Define sample variance of the squared residuals

$$\widehat{\delta}_n^2 = n^{-1} \sum_{j=1}^n \left(\widehat{\varepsilon}_j^2 - n^{-1} \sum_{j=1}^n \widehat{\varepsilon}_j^2 \right)^2. \tag{6}$$

THEOREM 1. *Let $0 < \alpha < 1/2$ and $1 \leqslant p < 1/(1-\alpha)$. Assume that H_0 holds and the function $\boldsymbol{f}$ is continuous and has finite p variation. Then*

$$n^{-1/2}\widehat{\delta}_n^{-1} T_{n,\alpha} \xrightarrow[n\to\infty]{\mathcal{D}} T_\alpha$$

if and only if

$$\lim_{t\to\infty} t^{2/(1/2-\alpha)} \mathrm{P}(|\varepsilon_1| > t) = 0.$$

The same is true for $\alpha = 0$ provided $\mathrm{E}\varepsilon_1^4 < \infty$.

THEOREM 2. *Assume that H_1 is true and $\mathrm{E}u_1^4$ is finite. Also assume that σ is fixed and $\tau = \tau_n$. If*

$$\ell^* \to \infty, \quad \ell^*/n \to \theta \in [0, 1/2), \quad (1 - \ell^*/n)|\tau^2 - \sigma^2|n^{-1/2+\alpha}(\ell^*)^{1-\alpha} \to \infty,$$

then for $0 \leqslant \alpha < 1/2$

$$n^{-1/2}\widehat{\delta}_n^{-1} T_{n,\alpha} \xrightarrow[n\to\infty]{\mathrm{P}} \infty.$$

Under H_1, if ℓ^*, k^*, τ and σ are known then G, as in (2), is a diagonal matrix with elements equal τ^2 when $j \in \{k^*+1, \ldots, k^*+\ell^*\}$, and the rest of elements equal σ^2. It is natural approach to estimate the unknown G by taking estimates $\widehat{\ell^*}$, $\widehat{k^*}$, $\widehat{\tau}$ and $\widehat{\sigma}$ instead of the corresponding parameters. This approach leads to the estimated GLS estimator of $\boldsymbol{\beta}$ defined in (4), if the null hypothesis of the homoskedasticity is rejected.

First define

$$T_n'(k, \ell) = \left| \sum_{j=k+1}^{k+\ell} \left(\widehat{\varepsilon}_j^2 - \frac{1}{n} \sum_{j=1}^n \widehat{\varepsilon}_j^2 \right) \right|,$$

$$T_{n,\alpha}'(\ell) = (\ell/n)^{-\alpha} \max_{0 \leqslant k < n} T_n'(k, \ell).$$

Then as an estimators of ℓ^* and k^* consider

$$\widehat{\ell^*} = \min\left\{ \ell: T_{n,\alpha}'(\ell) = \max_{0 < i < n/2} T_{n,\alpha}'(i) \right\},$$

$$\widehat{k^*} = \min\left\{ k: T_n'(k, \widehat{\ell^*}) = \max_{0 \leqslant i < n} T_n'(i, \widehat{\ell^*}) \right\}.$$

Next, we estimate τ as an empirical standard deviation of $\widehat{\varepsilon}$ over the index set $\{\widehat{k}^* + 1, \ldots, \widehat{k}^* + \widehat{\ell}^*\}$, and the estimate of σ is an empirical standard deviation of residuals with the indexes from $\{1, \ldots, n\} \setminus \{\widehat{k}^* + 1, \ldots, \widehat{k}^* + \widehat{\ell}^*\}$. Using these estimates we obtain $\widehat{G}$.

Unfortunately we can not yet provide the reader with the results showing that these estimators converge to their true values, nevertheless we can give empirical results to justify our reasoning.

3. Monte Carlo experiment

Set $d = 2$ and $\boldsymbol{f}(t) = 1 + t$, i.e., $f_1(i/n) \equiv 1$ and $f_2(i/n) = i/n$, $i = 1, \ldots, n$ and analyze a case with $n = 128$. We take $u_1 \sim N(0,1)$. Under the null hypothesis we fix $\sigma = 1.0$ and for the changed segment type of alternative we generate a changed segment of length $\ell^* = 16$ which begins at $k^* = 56$ and is of size $\tau = 2.0$.

In this experiment we are interested in the properties of OLS estimator $\widehat{\boldsymbol{\beta}}$, GLS estimator $\widehat{\boldsymbol{\beta}}_G$ and EGLS estimator $\widehat{\boldsymbol{\beta}}_{EG}$ defined in (1)–(4). Knowing the configuration of the experiment we can compute true values of the parameters of interest. For example, $E\widehat{\boldsymbol{\beta}} = E\widehat{\boldsymbol{\beta}}_G = E\widehat{\boldsymbol{\beta}}_{EG} = (1, 1)^{\mathrm{T}}$ both under the null hypothesis and under the alternative. Let us denote covariance matrices of the estimators by $V(\widehat{\boldsymbol{\beta}}) = E(\widehat{\boldsymbol{\beta}} - \mathbf{1})(\widehat{\boldsymbol{\beta}} - \mathbf{1})^{\mathrm{T}}$, $V(\widehat{\boldsymbol{\beta}}_G)$ and $V(\widehat{\boldsymbol{\beta}}_{EG})$. Then under H_0,

$$V(\widehat{\boldsymbol{\beta}}) = V(\widehat{\boldsymbol{\beta}}_G) = \begin{pmatrix} 0.0316 & -0.0472 \\ -0.0472 & 0.0938 \end{pmatrix},$$

and under the alternative H_1 it is easy to calculate that $V(\widehat{\boldsymbol{\beta}}) = \sigma^{*2}(X^{\mathrm{T}}X)^{-1}$, where $\sigma^{*2} = (\ell^*/n)\tau^2 + (1 - \ell^*/n)\sigma^2$. Hence,

$$V(\widehat{\boldsymbol{\beta}}) = \begin{pmatrix} 0.0435 & -0.0650 \\ -0.0650 & 0.1289 \end{pmatrix} \quad \text{and} \quad V(\widehat{\boldsymbol{\beta}}_G) = \begin{pmatrix} 0.0325 & -0.0473 \\ -0.0473 & 0.0939 \end{pmatrix}.$$

Thus, if we know the values of the change parameters, the GLS estimator reduces the variance. Next we will show that the EGLS estimator also does.

For the set $\{b_1, \ldots, b_R\}$ of realizations of any statistic b, let us denote empirical mean $Mb = R^{-1} \sum_{i=1}^{R} b_i$.

First for the model without the change we have generated $R = 1000$ replications of observations. For every replication we have computed $\widehat{\boldsymbol{\beta}}$ and $V(\widehat{\boldsymbol{\beta}})$ and obtained $MV(\widehat{\boldsymbol{\beta}}) = (0.0315, -0.0471, -0.0471, 0.0934)$ (writing the elements of a matrix in a row). The results obtained are in agreement with the theoretical findings. Next with the changed segment in variance of disturbances we again ran $R = 1000$ replications. If despite of the presence of heteroskedasticity we use OLS estimator we obtain $MV(\widehat{\boldsymbol{\beta}}) = (0.0436, -0.0651, -0.0651, 0.1292)$, just according to the theory. Then, as was proposed, we test for a changed segment and, if H_0 is rejected, find $\widehat{\boldsymbol{\beta}}_{EG}$ and $V(\widehat{\boldsymbol{\beta}}_{EG})$. We deffer the results one paragraph ahead and first add one remark.

Table 1. Empirical power of $T_{n,\alpha}$ and mean of $V(\widehat{\boldsymbol{\beta}}_{EG})$

α	0	1/16	1/8	3/16	1/4	5/16	3/8	7/16
pw	589	672	734	796	809	758	684	568
β_1	0.0295	0.0299	0.0305	0.0310	0.0317	0.0322	0.0330	0.0340
$\beta_1\beta_2$	−0.0425	−0.0434	−0.0443	−0.0453	−0.0464	−0.0474	−0.0488	−0.0506
β_2	0.0844	0.0861	0.0880	0.0899	0.0921	0.0940	0.0969	0.1004

Brownian motion in limiting statistics T_α (5) is approximated by the partial sum process $\zeta_m(0) = 0$,

$$\zeta_m(t) = \sum_{i=1}^{[mt]} Z_i + \left(mt - [mt]\right) Z_{[mt]+1} - t \sum_{i=1}^{m} Z_i, \quad t \in [0,1],$$

with normalization $m^{-1/2}$ and $Z_i \sim N(0,1)$, $i = 1, \ldots, m$. Taking a large value of m for small sample sizes we obtain significantly biased approximation for the null distribution of the test statistics (especially for α close to $1/2$). As a consequence wrong critical regions for the test statistics $T_{n,\alpha}$ are obtained. In [2] the procedure is proposed to fix this problem. First we propose to take $m = n$. Second we suggest taking Z_i^2 (and change normalization to $(2m)^{-1/2}$) instead of Z_i as for small sample sizes we can not expect CLT to work well. We use this procedure with $\alpha = i/16$, $i = 0, \ldots, 7$ and the significance level $\epsilon = 0.05$. As a result we obtain the test p-values ranging from 0.046 to 0.053, i.e., close to correct.

In Table 1 the results of experiment under H_1 are presented.

For the significance level $\epsilon = 0.05$, according to the critical regions obtained after the proposed procedure, for every α we compute the number when H_0 was rejected out of total number of replications $R = 1000$. In other words, we find empirical power (denoted pw in Table 3). The row indicated by β_i represents sample mean of estimators for the variance of β_i, $i = 1, 2$. And $\beta_1\beta_2$ denotes sample mean of estimators for the covariance between β_1 and β_2. Note however, that we have computed the sample means only from the realizations of estimators, when H_0 was rejected. Also note that all these estimators depend on α. We see that estimating $\boldsymbol{\beta}$ by $\widehat{\boldsymbol{\beta}}_{EG}$ gives the most accurate results when $\alpha = 5/16$. This confirms the authors' findings in other papers that detection of changed segments of various lengths crucially depends on the value of α.

References

1. R.C. Mittelhammer, G.G. Judge, D.J. Miller, *Econometric Foundations*, Cambridge University Press, Cambridge (2000).
2. A. Račkauskas, D. Zuokas, *New Tests of Heteroskedasticity in Linear Regression Models*, Vilnius University Faculty of Mathematics and Informatics, Preprint 2006-16 (2006).

REZIUMĖ

A. Račkauskas, D. Zuokas. Tiesinio regresinio modelio su heteroskedastiškomis paklaidomis koeficientų įverčių savybės

Šiame darbe nagrinėjamas tiesinis regresinis modelis $y = \beta X + \varepsilon$, kurio liekanos gali būti heteroskedastiškos. Yra žinoma, kad, esant heteroskedastiškumui, MK metodu rastas β įvertis $\widehat{\beta}$ nors ir išlieka nepaslinktas, tačiau turi didesnę dispersiją. Specialaus tipo heteroskedastiškumo – pasikeitusio segmento – alternatyvai siūlomas apibendrintas įvertintas MK įvertis $\widehat{\beta}_{EG}$. Monte-Carlo metodu palyginamos $\widehat{\beta}$ ir $\widehat{\beta}_{EG}$ dispersijos ir parodoma, kad jos atitinka paskaičiuotas tikrąsias, o pasiūlyto įverčio dispersija yra mažesnė.