Examining the Conditions that will strengthen the Success of the Iterative Stein-Rule Estimator of The Disturbance Variance in Proxy Model

D. Ünal
Department of Statistics, University of Çukurova, Adana, Turkey
dunal@cu.edu.tr

Abstract

The aim of this study is to give the conditions, in a linear regression model with proxy variables, when is the difference of variances of two estimators getting closer to each other. One of the mentioned estimators is the iterative Stein-rule estimator (ISRE) of the disturbance variance which is obtained by taking the Stein-rule estimator of the parameters in the estimator of the disturbance variance; one is the usual ordinary least squares (OLS) estimator of the disturbance variance. For that purpose the theoretical difference of variances is derived and the numerical analysis is handled to see the pattern of given theoretical difference.

Keywords: Stein-rule, Proxy variable, Iterative Stein-rule.
Subject classification codes: **62E99**

1. Introduction:

In some applied statistical analysis, some variables may not be observed by the researcher. In this situation, the researcher may ignore the variable or take approximate information for this unobservable variable(s). New variables obtained by taking the approximate information for the unobservable variables can be defined as Proxy variable.

In proxy variable case, the problem is how to treat this proxy information. There have been some studies about the performance of the Proxy variables in a linear regression model.

For instance, Wickens and McCallum showed that the use of proxy variable yields smaller bias (Wickens and McCallum 1972). Aigner showed that the use of a proxy variable does not necessarily lead to smaller mean squared error (MSE) (Aigner 1974). Namba and Ohtani derived the explicit formula of the predictive mean squared error of the Stein-rule estimator and the positive part Stein-rule estimator for the regression coefficients when the proxy variables are used (Namba and Ohtani 2006).

Ohtani considered the estimator of the disturbance variance in a linear regression when the Stein-rule estimator is used in place of the ordinary least squares (OLS) estimator and called as the iterative Stein-rule estimator (ISRE) of the disturbance variance (Ohtani 2006). Ünal and Akdeniz defined an iterative positive part Stein-rule estimator with proxy variables of the disturbance variance in a linear regression model with proxy variables. In that paper, they defined a pre-test estimator to get the MSE of the iterative positive part Stein-rule estimator (Ünal and Akdeniz 2006). Also they used incomplete beta functions and partial derivatives to analyse the performance of MSE of this estimator.

The ISRE of the disturbance variance in proxy models was also compared to OLS estimator of the disturbance variance in a linear regression model with proxy variables with respect to MSE criterion (Ünal 2010). Ünal gave the conditions to dominate the variance of the ISRE of the disturbance variance by usual OLS estimator of the disturbance variance theoretically in a linear regression model (Ünal 2007).

In this study, the variance formula for the iterative Stein-rule estimator of the disturbance variance in proxy models is given. Also the difference of variances of the ISRE and the usual OLS estimator of the disturbance variance in proxy model is demonstrated theoretically. And the conditions for which the difference of variances getting closer to each other, are taking into consideration. For these purposes:

In Section 2, the model and the estimators are constituted. In Section 3, the theoretical formula for the variance of the ISRE of the disturbance variance is obtained. In Section 4, theoretical difference between variance of the ISRE and the variance of the usual OLS estimator of the disturbance variance is given. In Section 5, numerical analysis is handled to see the pattern of given theoretical difference in section 4, and observe how it changes with changing values of parameters. And the last part concludes the paper by giving overall results of paper.

2. The Proxy Model

Let us first consider the partitioned linear regression model

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon \,, \qquad \varepsilon \sim N(0, \sigma^2 I_n) \tag{1}$$

where $y: n\times 1$ is observation vector of a dependent variable, $X_1 : n\times k_1$ and $X_2 : n\times k_2$ are matrices of observations of independent variables, $\beta_1 : k_1 \times 1$ and $\beta_2 : k_2\times 1$ are vectors of parameters and $\varepsilon : n\times 1$ is a vector of normal disturbance terms.

We assume the existence of X_2^* as the matrix of the proxy variables though X_2 is unobservable. Now, let us consider the linear regression model with proxy variable X_2^* in place of X_2:

$$y = X_1\beta_1 + X_2^*\beta_2^* + u^*, \qquad u^* \sim N(X_2\beta_2 - X_2^*\beta_2^*, \sigma^2 I_n) \tag{2}$$

where $u^* = X_2\beta_2 - X_2^*\beta_2^* + \varepsilon$ (Namba and Othani 2006). That is, choosing X_2^* instead of X_2 in the estimator of the disturbance variance collapses linear regression model to the linear regression model with proxy variable. Assume the nxk matrix $X^* = \left[X_1, X_2^*\right]$ has full column rank for $k = k_1 + k_2$ and consider the usual estimator $b^* = S^{*-1}X^{*\prime}y$ for parameter vector $\beta^* = \left[\beta_1', \beta_2'^*\right]'$, $\beta_2^* : k_2 \times 1$ based on the model (2) where $S^* = X^{*\prime}X^*$.

Also Stein-rule estimator in a linear regression model with proxy variables for the parameter vector β^* is

$$b^*_{SRP} = \left[1 - \frac{a e^{*\prime} e^*}{b^{*\prime} S^* b^*}\right] b^* \tag{3}$$

where $e^* = y - X^* b^*$ and $0 \le a \le 2(k-2)/(n-k+2)$.

The usual OLS estimator of the disturbance variance in a linear regression model with proxy variables can be given as,

$$s^{*2} = \frac{(y - X^* b^*)'(y - X^* b^*)}{n-k} \tag{4}$$

Let us take the iterative estimator of the disturbance variance in the model with proxy variables (Ünal 2010). This iterative disturbance variance is constituted by using b^*_{SRP} instead of b^* in the model with proxy variables:

$$\hat{\sigma}^{*2}_{IP} = \frac{(y - X^* b^*_{SRP})'(y - X^* b^*_{SRP})}{n-k} \tag{5}$$

where $b^*_{SRP} = \left[1 - \frac{a e^{*\prime} e^*}{b^{*\prime} S^* b^*}\right] b^*$. It is readily shown that equation (5) can be equivalently expressed as

$$\hat{\sigma}^{*2}_{IP} = \frac{1}{v}\left[e^{*\prime} e^* + a^2 \frac{(e^{*\prime} e^*)^2}{b^{*\prime} X^{*\prime} X^* b^*}\right] \tag{6}$$

providing $e^* = y - X^* b^*$, $(X^{*\prime} X^*) b^* = X^{*\prime} y$, $v = n-k$.

3. Theoretical Formula for the Variance of $\hat{\sigma}^{*2}_{IP}$

The iterative disturbance variance in equation (6) can be reduced to more feasible form by replacing $\frac{b^{*\prime} X^{*\prime} X^* b^*}{\sigma^2}$ with v_1 and $\frac{e^{*\prime} e^*}{\sigma^2}$ with v_2 as follows:

$$\hat{\sigma}^{*2}_{IP} = \frac{\sigma^2}{v}\left[v_2 + a^2 \frac{v_2^2}{v_1}\right] \tag{7}$$

Here, the quadratic forms v_1 and v_2 have independent non central chi-squared distributions i.e., $v_1 \sim \chi'^2_k(\lambda_1^*)$ and $v_2 \sim \chi'^2_v(\lambda_2^*)$ with non centrality parameters $\lambda_1^* = \frac{\beta' X' X^* S^{*-1} X^{*\prime} X \beta}{\sigma^2}$ and $\lambda_2^* = \frac{\beta' X' \left[I - X^*(X^{*\prime} X^*)^{-1} X^{*\prime}\right] X \beta}{\sigma^2}$.

Then the variance of the equation (7) is given by $V(\hat{\sigma}_{IP}^{*2}) = E\left[\left[\hat{\sigma}_{IP}^{*2} - E(\hat{\sigma}_{IP}^{*2})\right]^2\right]$. In this expression the following expectation formula (Ünal 2006)

$$E(\hat{\sigma}_{IP}^{*2}) = -\sigma^2 - \frac{\sigma^2\lambda_2^*}{v} - \frac{\sigma^2}{v}a^2 e^{-(\lambda_1^*+\lambda_2^*)}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{\lambda_1^{*i}+\lambda_2^{*j}(v+2j)(v+2j+2)}{i!\,j!(k+2i-2)}$$

should be used for further derivation. Then,

$$V(\hat{\sigma}_{IP}^{*2}) = E\left[\left[\frac{\sigma^2 v_2}{v} + \frac{\sigma^2 a^2 v_2^2}{vv_1} - \sigma^2 - \frac{\sigma^2\lambda_2^*}{v} - \frac{\sigma^2}{v}a^2 e^{-(\lambda_1^*+\lambda_2^*)}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{\lambda_1^{*i}+\lambda_2^{*j}(v+2j)(v+2j+2)}{i!\,j!(k+2i-2)}\right]^2\right] \quad (8)$$

is derived. It seems reasonable to consider $M = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}(v+2j)(v+2j+2)}{i!\,j!(k+2i-2)}$ while studying in equation (8). It can be shown that the previous equation can be equivalently expressed as

$$V(\hat{\sigma}_{IP}^{*2}) = E\begin{bmatrix} \frac{\sigma^4 v_2^2}{v^2} + \frac{\sigma^4 a^4 v_2^4}{v^2 v_1^2} + \sigma^4 + \frac{\sigma^4\lambda_2^{*2}}{v^2} + \frac{\sigma^4}{v^2}a^4 e^{-2(\lambda_1^*+\lambda_2^*)}M^2 \\ +2\frac{\sigma^4 v_2^3 a^2}{v^2 v_1} - 2\frac{\sigma^4 v_2}{v} - 2\frac{\sigma^4 v_2\lambda_2^*}{v^2} - 2\frac{\sigma^4 v_2 a^2 e^{-(\lambda_1^*+\lambda_2^*)}}{v^2}M \\ -2\frac{\sigma^4 a^2 v_2^2}{vv_1} - 2\frac{\sigma^4 a^2 v_2^2\lambda_2^*}{v^2 v_1} - 2\frac{\sigma^4 a^4 v_2^2 e^{-(\lambda_1^*+\lambda_2^*)}}{v^2 v_1}M \\ +2\frac{\sigma^4\lambda_2^*}{v} + 2\frac{\sigma^4}{v}a^2 e^{-(\lambda_1^*+\lambda_2^*)}M + 2\frac{\sigma^4 a^2\lambda_2^* e^{-(\lambda_1^*+\lambda_2^*)}}{v^2}M \end{bmatrix} \quad (9)$$

Then the iterative proxy variance will be

$$\begin{aligned} V(\hat{\sigma}_{IP}^{*2}) = {} & \frac{\sigma^4}{v^2}E\left(v_2^2\right) + \frac{\sigma^4 a^4}{v^2}E\left(\frac{v_2^4}{v_1^2}\right) + \sigma^4 + \frac{\sigma^4\lambda_2^{*2}}{v^2} + \frac{\sigma^4}{v^2}a^4 e^{-2(\lambda_1^*+\lambda_2^*)}M^2 + 2\frac{\sigma^4 a^2}{v^2}E\left(\frac{v_2^3}{v_1}\right) \\ & -2\frac{\sigma^4}{v}E\left(v_2\right) - 2\frac{\sigma^4\lambda_2^*}{v^2}E\left(v_2\right) - 2\frac{\sigma^4 a^2 e^{-(\lambda_1^*+\lambda_2^*)}}{v^2}E\left(v_2\right)M \\ & -2\frac{\sigma^4 a^2}{v}E\left(\frac{v_2^2}{v_1}\right) - 2\frac{\sigma^4 a^2\lambda_2^*}{v^2}E\left(\frac{v_2^2}{v_1}\right) - 2\frac{\sigma^4 a^4 e^{-(\lambda_1^*+\lambda_2^*)}}{v^2}E\left(\frac{v_2^2}{v_1}\right)M \\ & +2\frac{\sigma^4\lambda_2^*}{v} + 2\frac{\sigma^4}{v}a^2 e^{-(\lambda_1^*+\lambda_2^*)}M + 2\frac{\sigma^4 a^2\lambda_2^* e^{-(\lambda_1^*+\lambda_2^*)}}{v^2}M \end{aligned} \quad (10)$$

For two independently distributed non central chi-squared variables, $v_1 \sim \chi_k'^2(\lambda_1^*)$ and $v_2 \sim \chi_v'^2(\lambda_2^*)$, the following explicit formula,

$$E\left[\frac{v_1^n}{v_2^m}\right] = e^{-(\lambda_1^*+\lambda_2^*)}2^{n-m}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}\Gamma\left(\frac{k}{2}+i+n\right)\Gamma\left(\frac{v}{2}+j-m\right)}{i!\,j!\Gamma\left(\frac{k}{2}+i\right)\Gamma\left(\frac{v}{2}+j\right)} \quad (11)$$

introduced by Ünal in 2006, for the expectation of the ratio of different powers of v_1 and v_2 in proxy model can be used in equation (10) to go one step further (Ünal 2006). So, the following form is obtained for the variance of $\hat{\sigma}_{IP}^{*2}$,

$$
\begin{aligned}
V\left(\hat{\sigma}_{IP}^{*2}\right) &= \frac{4\sigma^4}{v^2} e^{-(\lambda_1^*+\lambda_2^*)} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!} \\
&+ \frac{4\sigma^4 a^4}{v^2} e^{-\left(\lambda_1^*+\lambda_2^*\right)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+3\right)\left(\frac{v}{2}+j+2\right)\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)\left(\frac{k}{2}+i-2\right)} \\
&+ \sigma^4 \quad + \frac{\sigma^4\lambda_2^{*2}}{v^2} + \frac{\sigma^4}{v^2} a^4 e^{-2\left(\lambda_1^*+\lambda_2^*\right)} M^2 \\
&+ \frac{8\sigma^4 a^2}{v^2} e^{-\left(\lambda_1^*+\lambda_2^*\right)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+2\right)\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)} \\
&- 2\frac{\sigma^4 v}{v} - 2\frac{\sigma^4\lambda_2^*}{v} - 2\frac{\sigma^4\lambda_2^* v}{v^2} - 2\frac{\sigma^4\lambda_2^{*2}}{v^2} \\
&- 2\frac{\sigma^4 a^2 e^{-\left(\lambda_1^*+\lambda_2^*\right)}}{v^2} Mv - 2\frac{\sigma^4 a^2 e^{-\left(\lambda_1^*+\lambda_2^*\right)}}{v^2} M\lambda_2^* \\
&- 4\frac{\sigma^4 a^2}{v} e^{-\left(\lambda_1^*+\lambda_2^*\right)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)} \\
&- 4\frac{\sigma^4 a^2\lambda_2^*}{v^2} e^{-\left(\lambda_1^*+\lambda_2^*\right)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)} \\
&- 4\frac{\sigma^4 a^4 e^{-2\left(\lambda_1^*+\lambda_2^*\right)}}{v^2} M \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)} \\
&+ 2\frac{\sigma^4\lambda_2^*}{v} + 2\frac{\sigma^4}{v} a^2 e^{-\left(\lambda_1^*+\lambda_2^*\right)} M + 2\frac{\sigma^4 a^2\lambda_2^* e^{-\left(\lambda_1^*+\lambda_2^*\right)}}{v^2} M
\end{aligned}
\tag{12}
$$

where $M = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \frac{\lambda_1^{*i}\lambda_2^{*j}(v+2j)(v+2j+2)}{i!\,j!(k+2i-2)}$. It is clearly seen that, the derived equation is inextricable thanks to the existence of non central chi-squared variables in ISRE of the disturbance variance and the parameters come from these non-central distributions. These features preclude the further simplification of equation (12).

4. The Difference of Variances

In this section, the difference of the variance of the iterative stein-rule estimator of the disturbance variance and the variance of the usual OLS estimator of the disturbance variance is taken in proxy model. The difference, $V\left(\hat{\sigma}_{IP}^{*2}\right)-V\left(s^{*2}\right)$, mentioned in introductory sentence is:

$$
\begin{aligned}
V\left(\hat{\sigma}_{IP}^{*2}\right)-V\left(s^{*2}\right)=&\frac{4\sigma^4 a^4}{v^2}e^{-(\lambda_1^*+\lambda_2^*)}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+3\right)\left(\frac{v}{2}+j+2\right)\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)\left(\frac{k}{2}+i-2\right)}\\
&+\frac{\sigma^4\lambda_2^{*2}}{v^2}+\frac{\sigma^4}{v^2}a^4e^{-2\left(\lambda_1^*+\lambda_2^*\right)}M^2\\
&+\frac{8\sigma^4 a^2}{v^2}e^{-\left(\lambda_1^*+\lambda_2^*\right)}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+2\right)\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)}\\
&-2\frac{\sigma^4\lambda_2^{*2}}{v^2}-2\frac{\sigma^4a^2e^{-\left(\lambda_1^*+\lambda_2^*\right)}}{v}M-2\frac{\sigma^4a^2e^{-\left(\lambda_1^*+\lambda_2^*\right)}}{v^2}M\lambda_2^*\\
&-4\frac{\sigma^4a^2}{v}e^{-\left(\lambda_1^*+\lambda_2^*\right)}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)}\\
&-4\frac{\sigma^4a^2\lambda_2^*}{v^2}e^{-\left(\lambda_1^*+\lambda_2^*\right)}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)}\\
&-4\frac{\sigma^4a^4e^{-2\left(\lambda_1^*+\lambda_2^*\right)}}{v^2}M\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}\left(\frac{v}{2}+j+1\right)\left(\frac{v}{2}+j\right)}{i!\,j!\left(\frac{k}{2}+i-1\right)}\\
&+2\frac{\sigma^4}{v}a^2e^{-\left(\lambda_1^*+\lambda_2^*\right)}M+2\frac{\sigma^4a^2\lambda_2^*e^{-\left(\lambda_1^*+\lambda_2^*\right)}}{v^2}M \qquad (13)
\end{aligned}
$$

where $M=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{\lambda_1^{*i}\lambda_2^{*j}(v+2j)(v+2j+2)}{i!\,j!(k+2i-2)}$. Beside the complexity of studying in proxy models, defining an iterative estimator brings the double difficulty because of leaving the centrality of chi-squared distributions in theoretical study. So, the equation (13) is too labyrinthine to make a comment on. And, constructing the numerical calculations is an inevitable job to overcome the impossibility of examining the meaning of this equation theoretically.

5. Numerical Analysis

In view of the complexity of the theoretical expression obtained in previous section, this section devoted to numerical evaluations and illustrations to overcome the sophisticate nature of expression (13) and to make an inference on it. That is, we can see the result by numerical evaluations to get clear insight.

For purposes of analysis, some different values which are frequently used in literature (Ohtani 1986) are considered in numerical evaluations: λ_1^*=0.0001, 0.001, 0.01, 0.1, 0.3; λ_2^*=0.0001, 0.001, 0.01, 0.1, 0.3; n =10, 20, 40, 60; k =3, 5, 6, 7, 8. It is also worth noting that a can be any value in interval $[0, a, 2(k-2)/(n-k+2)]$. The evaluations are made using Mathematica programming, which allow us to perform some complicated and tedious algebraic calculation on a computer, as well as help us to find new exact solutions and some representative results are illustrated in Table 1 and Table 2. Table 1 corresponds to the case where λ_1^*(without loss of generality λ_1^* is taken 0.1) is constant and the values of $V\left(\hat{\sigma}_{IP}^{*2}\right)-V\left(s^{*2}\right)$ is looking over for the changing values of k, n and λ_2^*. Similarly, in Table 2, the values of $V\left(\hat{\sigma}_{IP}^{*2}\right)-V\left(s^{*2}\right)$ is given for the constant λ_2^* (without loss of generality λ_2^* is taken 0.1) and the changing values of k, n and λ_1^*. It should be noticed that if the difference in equation (13) is close to zero, then the variance of $V\left(\hat{\sigma}_{IP}^{*2}\right)$ is close to $V\left(s^{*2}\right)$, i.e. the variances $V\left(\hat{\sigma}_{IP}^{*2}\right)$ and $V\left(s^{*2}\right)$ are approximately equal.

Some broad features emerges from the numerical results are;

- For fixed values of λ_1^*, λ_2^*, if n increases for fixed k then the difference of variances given in equation (13) decreases. More than this, increasing in n with decreasing k mentioned difference in (13) seems gained to 0.
- For fixed values of λ_1^*, λ_2^* and n, variances of $\hat{\sigma}_{IP}^{*2}$ and s^{*2} are getting closer by decreasing in k.
- The values of non-centrality parameters $(\lambda_1^*, \lambda_2^*)$ approaching to zero mean that the non-central chi-squared distribution approaches to central chi-squared distribution. For example, for constant λ_1^* and decreasing λ_2^* will make the difference of variances decreasing. For fixed values of n and k bring good luck to the variance differences by decreasing in λ_2^*. In other words, getting the distribution of $v_2 \sim \chi_v'^2(\lambda_2^*)$ to centrality strengthens the achievement of proposed iterative estimator (ISRE in proxy model) by means of variance criterion. But it does not assert the same luckiness by decreasing in λ_1^*. In whole study, for the case where the increases in difference of variances come across, different values of other parameters are also considered to catch the decreasing case of differences of variances. Then to overcome the unluckiness of decreasing in λ_1^* we study on it comprehensively. As a result of such an effort, we see that for the λ_1^* values

which are close to zero, the difference will decrease $\forall\ k < 5$, $\forall n$. Also if $\mathrm{k} = 5$, decreasing of difference of variances can be provided just for $n \geq 40$ in decreasing λ_1^*. But there are no situations under which $k > 5$ for decreasing λ_1^* making the difference of variances tendency to decrease even if $n = 1000000$.

- For constant λ_1^*, decreasing in k values makes the difference smaller and decreasing in λ_2^* decreases the $V\left(\hat{\sigma}_{IP}^{*2}\right) - V\left(s^{*2}\right)$
- For constant values of λ_1^* and k, having the bigger n takes us having smaller difference. This statement is also satisfied for λ_2^* and k.
- For constant n and λ_1^*, regardless of the alteration of λ_2^*, if k decreases then the difference will decrease also. In other words, for constant n and λ_1^*, decreasing in k makes difference decreasing even if λ_2^* increase and vice verse. That is, for constant n and λ_2^*, regardless of the variation of λ_1^*, if k decreases then the difference will decrease also.
- For fixed values of λ_1^*, λ_2^*, n, the decreasing in k justifies the decreasing the values of $V\left(\hat{\sigma}_{IP}^{*2}\right) - V\left(s^{*2}\right)$, and vice verse.

The main findings discovered from numerical analysis can be given by generalization to contribute to the literature in Conclusion.

Conclusion

In this study, in proxy model, given ISRE of the disturbance variance, $\hat{\sigma}_{IP}^{*2}$, and the usual OLS estimator of the disturbance variance, s^{*2} are taken into consideration. Also, their superiority of each other is examined by using variance criterion. Using the obtained theoretical formula for $V\left(\hat{\sigma}_{IP}^{*2}\right) - V\left(s^{*2}\right)$, the performance of variances owing to numerical computations of equations, by using mathematica code, are interpreted numerically. Especially, the conditions of the difference of variances getting closer to each other, are taking into consideration.

As noticed before, in numerical evaluations some various values are chosen for parameters which are frequently used in the literature and some represented results are given in the tables. This means that the tendency of the results is not changed when some other values are taken for the parameters.

If our aim of this study is to analyse the alteration of $V\left(\hat{\sigma}_{IP}^{*2}\right) - V\left(s^{*2}\right)$, then the general findings can be given as follows:

- Increase in value of n is the most effective factor in decrease of variance differences.

- Small values of λ_1^* and λ_2^*, approaching to the central chi-squared distribution, have a small effect on the decreasing difference in proxy models.
- Whatever the magnitudes of other parameters, small values of parameter k have an effect on decrease of variance differences also.
- In general, as n gets bigger, k gets smaller and non-centrality parameters get closer to 0, then variances of two estimators get closer to each other.
- One broad feature which emerges from the results is that regardless of the values of n, whenever distribution of v_1 and v_2 tends to central chi-squared distribution, the difference will decrease for $k < 5$. In other words, it is observed that, whenever $k < 5$, neither estimator strictly dominates the other, in all parts of the parameter space. These results are compatible with the result of the paper which concerned with the similarly proposed estimator of the disturbance variance in non-proxy cases and compared them by using MSE criterion (Ohtani 1987). In that paper Ohtani showed that, ISRE of the disturbance variance is dominated by the usual estimator of the disturbance variance based on the OLS estimator under the MSE criterion, if the number of regressors, k, is greater than or equal to five. Then it may convince us that ignoring biasness or facing with proxyness are still taking us similar results having it even more difficult and having it in strict sense (Ohtani 1987).

In many of situations that we have considered, the former estimator seems to have the edge on the latter for smaller k (exactly, $k < 5$) values with respect to variance criterion. Given these findings, we faced with, for the optimal choice of the parameters (as n gets bigger, k gets smaller and non-centrality parameters get closer to 0), the $V\left(\hat{\sigma}_{IP}^{*2}\right)$ needs no further discussion getting closer to $V\left(s^{*2}\right)$ in proxy models.

Acknowledgement

I am indebted to an anonymous reviewer of this paper for providing insightful comments and directions to improve this work. I also would like to thank my MS student Burcu Alpman for her contribution on numerical analysis.

References

1. Aigner, D. J. (1974). MSE-Dominance of Least Squares with Errors of Observation, Journal of Econometrics, 2, 365-372.
2. McCallum, B.T. (1972). Relative Asymptotic Bias from Errors of Omission and Measurement, Econometrica, 40, 757-758.
3. Namba, A., Ohtani, K. (2006). PMSE Performance of the Stein-Rule and Positive Part Stein-Rule Estimators in a Linear Regression Model with or without Proxy Variables, Statistics and Probability Letters, 76, 898-906.

4. Ohtani, K. (1986). A Distribution Function of the F Ratio when the Stein-Rule Estimator is used in Place of the Ordinary Least Squares Estimators, Economic Letters, 21, 257-260.

5. Ohtani, K. (1987). Inadmissibility of the Iterative Stein-Rule Estimator of the Disturbance Variance in a Linear Regression, Journal of Econometrics, 24, 51-55.

6. Ünal, D. (2006). The Doubly Noncentral F Distribution in a Regression Model with Proxy Variables, International Journal of Pure and Applied Mathematics, 30, 3, 387-392.

7. Ünal, D. (2007). More on the Inadmissibility of the Iterative Stein-rule Estimator of the Disturbance Variance in a Linear Regression, Far East Journal of Theoretical Statistics, 21, 1, 1-12.

8. Ünal D. (2010). The Effects of the Proxy Information on the Iterative Stein-Rule Estimator of the Disturbance Variance, Statistical Papers, 51, 2, 477-484.

9. Ünal, D., Akdeniz F. (2006). The Iterative Stein-Rule Estimator of the Disturbance Variance in a Linear Regression Model when the Proxy Variables are Used, Selçuk Journal of Applied Mathematics, 7, 2, 13-25

10. Wickens, M. R. A. (1972). Note on the Use of Proxy Variables, Econometrica, 40, 4, 759-761.

Appendix for Tables

Some selected results are given in the following tables.

Table 1: λ_1^* = 0.1 is constant

n	k	λ_2^*	$V(\hat{\sigma}_{IP}^{*2})-V(s^{*2})$	n	k	λ_2^*	$V(\hat{\sigma}_{IP}^{*2})-V(s^{*2})$
10	3	0.0001	0.217508	40	3	0.0001	0.0153744
		0.01	0.219523			0.01	0.0154504
		0.3	0.277283			0.3	0.0176953
	5	0.0001	49.364		5	0.0001	1.07145
		0.01	50.0525			0.01	1.07398
		0.3	72.3605			0.3	1.15433
	8	0.0001	1201.43		8	0.0001	7.45886
		0.01	1236.86			0.01	7.46046
		0.3	2372.57			0.3	7.60504
20	3	0.0001	0.0609396	60	3	0.0001	0.00682612
		0.01	0.0613391			0.01	0.00685627
		0.3	0.0732887			0.3	0.00774147
	5	0.0001	5.77401		5	0.0001	0.434919
		0.01	5.80198			0.01	0.435661
		0.3	6.67644			0.3	0.459602
	8	0.0001	49.706		8	0.0001	2.86967
		0.01	49.8659			0.01	2.86834
		0.3	55.1601			0.3	2.8689

Table 2: λ_2^* =0.1 is constant

n	k	λ_1^*	$V(\hat{\sigma}_{IP}^{*2})-V(s^{*2})$	n	k	λ_1^*	$V(\hat{\sigma}_{IP}^{*2})-V(s^{*2})$
10	3	0.0001	0.0641906	40	3	0.0001	0.00996405
		0.01	0.0817526			0.01	0.0105861
		0.3	0.563305			0.3	0.0279308
	5	0.0001	57.4714		5	0.0001	1.09414
		0.01	57.3692			0.01	1.09437
		0.3	55.1872			0.3	1.11242
	8	0.0001	1621.26		8	0.0001	7.59664
		0.01	1615.94			0.01	7.58483
		0.3	1479.28			0.3	7.31694
20	3	0.0001	0.0356785	60	3	0.0001	0.00452831
		0.01	0.0386351			0.01	0.00479017
		0.3	0.120632			0.3	0.0121036
	5	0.0001	6.07764		5	0.0001	0.440426
		0.01	6.07535			0.01	0.4406
		0.3	6.07751			0.3	0.450054
	8	0.0001	52.3525		8	0.0001	2.9
		0.01	52.2509			0.01	2.89579
		0.3	49.8389			0.3	2.80188