



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

INSTITUT FÜR STATISTIK



Shalabh & Helge Toutenburg & Christian Heumann

Stein-Rule Estimation under an Extended Balanced Loss Function

Technical Report Number 007, 2007
Department of Statistics
University of Munich

<http://www.stat.uni-muenchen.de>



Stein-Rule Estimation under an Extended Balanced Loss Function

Shalabh*, H. Toutenburg[†] and C. Heumann[‡]

December 5, 2007

Abstract

This paper extends the balanced loss function to a more general set up. The ordinary least squares and Stein-rule estimators are exposed to this general loss function with quadratic loss structure in a linear regression model. Their risks are derived when the disturbances in the linear regression model are not necessarily normally distributed. The dominance of ordinary least squares and Stein-rule estimators over each other and the effect of departure from normality assumption of disturbances on the risk property is studied.

Key words: Linear regression model, Stein-rule estimator, ordinary least squares estimator, balanced loss function, non-normal disturbances.

*Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur - 208016, India. E-mail: shalab@iitk.ac.in, shalabh1@yahoo.com (Corresponding author)

[†]Department of Statistics, Ludwig-Maximilians-University, Ludwigstr. 33, 80539 Munich, Germany. E-mail: helge.toutenburg@stat.uni-muenchen.de

[‡]Department of Statistics, Ludwig-Maximilians-University, Ludwigstr. 33, 80539 Munich, Germany. E-mail: chris@stat.uni-muenchen.de

1 Introduction

The ordinary least squares estimator is the best linear unbiased estimator of regression parameter in a linear regression model. If the criterion of linearity and unbiasedness can be dropped, then it is possible to improve upon the variability of an estimator of regression coefficient. The family of Stein-rule estimator proposed by James and Stein (1961) has smaller variability than ordinary least squares estimator under quadratic risk with a simple condition that the number of explanatory variables are more than two. A vast literature is available on the topic of Stein-rule estimation. Discussion of all the papers is out of purview of this paper. Most of the literature judges the performance of Stein-rule and ordinary least squares estimators on the basis of the concentration of estimates around the true value of the parameter. The goodness of fitted model is another criterion to judge the performance of these estimators which is generally ignored quite often.

Generally either of the criterion is employed to judge the performance. In practice, both the criterion may often be desirable, see, for instance, Shalabh (1995, 2000), Toutenburg and Shalabh (1996), and Zellner (1994) for some illustrative examples. Accordingly, Zellner (1994) has introduced the balanced loss function which is a convex combination of sum of squares of the residuals and weighted sum of squares of the estimation errors. Such loss function encompasses both the criterion, viz., concentration of estimates around true parameter and goodness of fitted of model. Further, Shalabh (1995) introduced the concept of simultaneous prediction of actual and average values of study variable in a linear regression model for within and outside sample prediction. Based on that, Shalabh (1995) has presented a predictive loss function which not only incorporate the balanced loss function of Zellner (1994) as its particular case but also measures the correlation between the goodness of fit of model and concentration of estimates around the true parameter.

The balanced loss function has received considerable attention in the literature under different set ups. For example, Rodrigues and Zellner (1994) have used the balanced loss function in the estimation of mean time to failure; the aspects of preliminary test estimation and Stein-rule estimation under the balanced loss function are discussed by Giles, Giles and Ohtani (1996), Ohtani, Giles and Giles (1997), Ohtani (1998, 1999) and Gruber (2004); estimation of normal mean is considered by Chung, Kim and Song (1997), Chung, Chanso and Dey (1999) and Sanjari and Asgharzadeh (2004); see also Chung, Kim and Song (1998), Shalabh (2001), Wan (2002), Chaturvedi and Shalabh (2004), Fikri, Wan and Akdeniz (2005) and Toutenburg and Shalabh (2005) for the application of balanced loss function in some other areas of linear models. Appreciating the popularity of balanced loss function, we have extended it further and present a general loss function called as extended balanced loss function in this paper. The predictive loss functions suggested by Zellner (1994) and Shalabh (1995) are its particular cases.

Most of the literature dealing with Stein-rule estimation of regression coefficients assume that the disturbances are normally distributed. In practice, this assumption may not hold true. How the performance of Stein-rule estimator changes under the non-normally distributed disturbances is another question. An attempt is made to judge the performance of ordinary least squares and Stein-rule estimators under the proposed extended balanced loss function when disturbances in the linear regression model are not necessarily normally distributed.

In Section 2 we describe the linear regression model and present a general loss function under quadratic loss structure. A comparison of the risk functions associated with the ordinary least squares and Stein-rule estimators is presented in Section 3, and a condition on the characterizing scalar for the superiority of Stein-rule estimators over ordinary least squares estimator is obtained. Several popular

forms of loss functions which arise as particular cases of the extended balanced loss function are also considered. The finite sample behaviour of ordinary least squares and Stein-rule estimators is studied through a Monte-Carlo simulation experiment and its findings are reported in Section 4. Some concluding remarks are then placed in Section 5.

2 Linear regression model and the loss function

Let us consider the following linear model:

$$y = X\beta + \epsilon \quad (2.1)$$

where y is a $n \times 1$ vector of n observations on the study variable, X is a $n \times p$ full column rank matrix of n observations on each of the p explanatory variables, and ϵ is a $n \times 1$ vector of disturbances. It is assumed that the elements of ϵ are independently and identically distributed following a distribution with mean 0, variance σ^2 and third moment $\gamma_1\sigma^3$ measuring skewness where γ_1 is the Pearson's coefficient of skewness of the distribution of ϵ .

Let $\tilde{\beta}$ denotes any estimator of β then the quadratic loss function which reflects the goodness of fit of the model is

$$(X\tilde{\beta} - y)'(X\tilde{\beta} - y) \quad (2.2)$$

where $\hat{y} = X\tilde{\beta}$ is the predictor for y . Similarly, the precision of estimation of $\tilde{\beta}$ is measured by the weighted loss function

$$(\tilde{\beta} - \beta)'X'X(\tilde{\beta} - \beta) . \quad (2.3)$$

Generally, either of the criterion of (2.2) or (2.3) is used to judge the performance of any estimator. Considering the goodness of fit and precision of estimation together, Zellner (1994) has considered both the criterion (2.2) and (2.3)

together and proposed the following balanced loss function:

$$BL(\tilde{\beta}) = \theta(X\tilde{\beta} - y)'(X\tilde{\beta} - y) + (1 - \theta)(\tilde{\beta} - \beta)'X'X(\tilde{\beta} - \beta) \quad (2.4)$$

where θ is a scalar lying between 0 and 1 which provides the weight assigned to the goodness of fit of model.

If we consider the set up of within sample prediction, then the predictor $X\tilde{\beta}$ is used to predict the actual value y as well as the average value $E(y)$. In certain cases, it may be desirable to consider the simultaneous prediction of y and $E(y)$, see Shalabh (1995), Toutenburg and Shalabh (1996) and Rao, Toutenburg, Shalabh and Heumann (2008) for more details and for some examples. From the prediction view of point, the loss functions (2.2) and (2.3) can be regarded as arising from the prediction of actual values y by $X\tilde{\beta}$ and the prediction of the average values $E(y) = X\beta$ by $X\tilde{\beta}$, respectively. Accordingly, Shalabh (1995) has defined a target function

$$T = \lambda y + (1 - \lambda)E(y) \quad (2.5)$$

where λ is a scalar between 0 and 1. Note that $\lambda = 0$ and $\lambda = 1$ in (2.5) provides the predictions for average and actual values of y . Any other value $0 < \lambda < 1$ provides the weight assigned to the actual value prediction and provides simultaneous prediction of actual and average values of y .

The following predictive loss function arises when we use the predictor $X\tilde{\beta}$ for simultaneous prediction of actual and average values of y through the target function (2.5):

$$\begin{aligned} PL(\tilde{\beta}) &= (X\tilde{\beta} - T)'(X\tilde{\beta} - T) \\ &= \lambda^2(X\tilde{\beta} - y)'(X\tilde{\beta} - y) + (1 - \lambda)^2(\tilde{\beta} - \beta)'X'X(\tilde{\beta} - \beta) \\ &\quad + 2\lambda(1 - \theta)(X\tilde{\beta} - T)'X(X\tilde{\beta} - T). \end{aligned} \quad (2.6)$$

Looking at the functional forms of the balanced loss function and the predictive loss function, we propose the following weighted loss function:

$$\begin{aligned} WL(\tilde{\beta}) &= \lambda_1(X\tilde{\beta} - y)'(X\tilde{\beta} - y) + \lambda_2(\tilde{\beta} - \beta)'X'X(\tilde{\beta} - \beta) \\ &\quad + (1 - \lambda_1 - \lambda_2)(X\tilde{\beta} - y)'X(\tilde{\beta} - \beta) \end{aligned} \quad (2.7)$$

where λ_1 and λ_2 are the scalars lying between 0 and 1 characterizing the loss functions. The weighted loss function in (2.7) can be called as extended balanced loss function. Clearly, the function (2.7) encompasses the loss functions (2.2), (2.3), (2.4) and (2.6) as its particular cases. Thus it is fairly general and sufficiently flexible.

3 Risk performance of ordinary least squares and Stein-rule estimators

The ordinary least squares estimator of β is given by

$$b = (X'X)^{-1}X'y \quad (3.1)$$

which is well known for its optimality in the class of linear and unbiased estimators.

If we drop the linearity and unbiasedness, there exist estimators with better performance than the ordinary least squares estimator under the risk criterion. One such interesting family of nonlinear and biased estimators of β , popularly known as Stein-rule estimators (see, James and Stein (1961)) is defined by

$$\hat{\beta} = \left[1 - \left(\frac{k}{n - p + 2} \right) \frac{y'(I - H)y}{y'Hy} \right] b \quad (3.2)$$

where

$$H = X(X'X)^{-1}X' \quad (3.3)$$

and k is a positive nonstochastic characterizing scalar; see, e.g. Judge and Bock (1978) and Saleh (2006).

A vast literature is available related to the performance of ordinary least squares and Stein-rule estimators under either of the criterion, viz., goodness of fit of model or concentration of estimates around the true parameter. We propose to judge the performance of these two estimators under the fairly general extended balanced loss function using the criterion as risk, i.e., the expected value of the weighted loss function (2.7).

From (2.1) and (2.7), we observe that

$$WL(\tilde{\beta}) = \lambda_1 \epsilon' \epsilon - \sigma(1 + \lambda_1 - \lambda_2) \epsilon' X(\tilde{\beta} - \beta) + (\tilde{\beta} - \beta)' X' X(\tilde{\beta} - \beta). \quad (3.4)$$

Setting $\tilde{\beta} = b$, we get the risk of OLSE as

$$R(b) = E[WL(b)] = \sigma^2 \lambda_1 n - \sigma^2 p(\lambda_1 - \lambda_2). \quad (3.5)$$

The exact expressions for the risk function of SRE can be derived but their nature would be sufficiently intricate. We therefore employ the large sample asymptotic approximation theory to derive the risk function. We assume that the explanatory variables are asymptotically cooperative, i.e., the limiting form of the matrix $n^{-1}X'X$ is finite and nonsingular, as n tends to infinity. Such an assumption is needed for the application of large sample asymptotic approximation theory.

Now, if we write

$$\xi = n^{\frac{1}{2}} X' \epsilon$$

and

$$\delta = n^{\frac{1}{2}} \left(\frac{\epsilon' \epsilon}{n} - 1 \right),$$

we have

$$b - \beta = \frac{1}{n^{\frac{1}{2}}} S^{-1} \xi. \quad (3.6)$$

and

$$y'(I - H)y = n + n^{\frac{1}{2}}\delta - \xi'S^{-1}\xi.$$

Next, consider the quantity

$$\begin{aligned} \frac{y'(I - H)y}{(n - p + 2)y'Hy} &= \frac{1}{n\beta'S\beta} \left(1 + \frac{\delta}{n^{\frac{1}{2}}} - \frac{1}{n}\xi'S^{-1}\xi \right) \left(1 - \frac{p-2}{n} \right)^{-1} \\ &\quad \times \left(1 + \frac{2\beta'\xi}{n^{\frac{1}{2}}\beta'S\beta} + \frac{\xi'S^{-1}\xi}{n\beta'S\beta} \right)^{-1}. \\ &= \frac{1}{n\beta'S\beta} + \frac{1}{n^{\frac{3}{2}}\beta'S\beta} \left(\delta - \frac{2\beta'\xi}{\beta'S\beta} \right) + Op(n^{-2}). \end{aligned} \quad (3.7)$$

Substituting (3.6) and (3.7) in (3.2), we find

$$\begin{aligned} (\hat{\beta} - \beta) &= \frac{1}{n^{\frac{1}{2}}}S^{-1}\xi - \frac{k}{n\beta'S\beta}\beta \\ &\quad - \frac{k}{n^{\frac{3}{2}}\beta'S\beta} \left[\delta\beta + \left(S^{-1} - \frac{2}{\beta'S\beta}\beta\beta' \right) \xi \right] + Op(n^{-2}). \end{aligned} \quad (3.8)$$

Making use of the distributional properties of ϵ , these results and neglecting terms of higher order of smallness than $O(n^{-1})$, we see from (3.8) that

$$E[\epsilon'X(\hat{\beta} - \beta)] = \sigma p - \frac{\sigma^2 k}{n\beta'S\beta} \left[\frac{\gamma_1}{n}e'X\beta + \sigma(p-2) \right] \quad (3.9)$$

$$\begin{aligned} E[(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta)] &= \sigma^2 p - \frac{\sigma^3 k}{n\beta'S\beta} \left[\frac{2\gamma_1}{n}e'X\beta + 2\sigma(p-2) - \sigma k \right]. \end{aligned} \quad (3.10)$$

Setting $\tilde{\beta} = \hat{\beta}$ in (3.4), utilizing the above results and retaining the terms to order $O(n^{-1})$, we find

$$\begin{aligned} R(\hat{\beta}) &= E[WL(\hat{\beta})] \\ &= \sigma^2 \lambda_1 n - \sigma^2 p(\lambda_1 - \lambda_2) \\ &\quad - \frac{\sigma^4 k}{n\beta'S\beta} \left[(1 - \lambda_1 + \lambda_2) \left(\frac{\gamma_1}{\sigma} \bar{X}'\beta + p - 2 \right) - k \right]. \end{aligned} \quad (3.11)$$

where $S = \frac{1}{n}X'X$ and \bar{X} is a $p \times 1$ vector of means of observations on the p explanatory variables.

It is clear from (3.5) and (3.11) that both the criteria, viz, goodness of fit and precision of estimation affect the risk performance of ordinary least squares as well as Stein-rule estimators. Using the criteria of goodness of fit and precision of estimation together has more appeal than using either of them.

Comparing (3.5) and (3.11), it is observed that the Stein-rule estimator has smaller risk to the order of our approximations, in comparison to the ordinary least squares estimator when

$$k < (1 - \lambda_1 + \lambda_2) \left(\frac{\gamma_1}{\sigma} \bar{X}'\beta + p - 2 \right) \quad (3.12)$$

provided that

$$(\lambda_1 - \lambda_2) < 1 \text{ and } \left(\frac{\gamma_1}{\sigma} \bar{X}'\beta + p - 2 \right) > 0 \quad (3.13)$$

or

$$(\lambda_1 - \lambda_2) > 1 \text{ and } \left(\frac{\gamma_1}{\sigma} \bar{X}'\beta + p - 2 \right) < 0. \quad (3.14)$$

When the distribution of disturbances is symmetric and/or \bar{X} is a null vector, i.e., the observations on the explanatory variables are taken as deviations from their corresponding means, then the condition (3.12) becomes free from unknown parameters β and is satisfied when either of the following two conditions holds true:

$$k < (1 - \lambda_1 + \lambda_2)(p - 2) \text{ and } (\lambda_1 - \lambda_2) < 1 \text{ if } p > 2 \quad (3.15)$$

$$k < (\lambda_1 - \lambda_2 - 1)(2 - p) \text{ and } (\lambda_1 - \lambda_2) > 1 \text{ if } p = 1, 2. \quad (3.16)$$

Now we examine the performance of estimators under some interesting loss functions.

First we consider the criterion of goodness of fit of model which is a particular case of (2.7) with $\lambda_1 = 1$ and $\lambda_2 = 0$. We observe from (3.5) and (3.11) that

the ordinary least squares estimator remains unbeaten by all Stein-rule estimators irrespective of the nature of observations on explanatory variables and the distribution of disturbances. This matches with the result obtained by Srivastava and Shalabh (1996, p.143) on the basis of exact risk expressions.

Considering the criterion of precision of estimation which can be obtained by setting $\lambda_1 = 0$ and $\lambda_2 = 1$ in (2.7), the Stein-rule estimators are better than the ordinary least squares estimator when

$$k < 2 \left(\frac{\gamma_1}{\sigma} \bar{X}'\beta + p - 2 \right) \quad (3.17)$$

with the rider that the quantity on the right hand side is positive; see also Vinod and Srivastava (1995).

This condition reduces to

$$k < 2(p - 2); p > 2 \quad (3.18)$$

when the distribution of disturbances is symmetric irrespective of the nature of data on the explanatory variables or \bar{X} is a null vector whether the distributions of disturbances is symmetric or asymmetric.

Similarly, the condition (3.17) is satisfied as long as (3.18) holds true provided that γ_1 and $\bar{X}'\beta$ have the same sign, i.e., $\bar{X}'\beta$ is positive for positively skewed distributions of disturbances and is negative for negatively skewed distributions of disturbances. In fact, it is possible to find Stein-rule estimators with better performance than the ordinary least squares estimator even for $p = 1$ and $p = 2$ when

$$\gamma_1 \bar{X}'\beta > 2\sigma. \quad (3.19)$$

It may be noticed (3.18) is a well-known condition for the superiority of Stein-rule estimators on the basis of exact risk under the normality of disturbances; see, e.g., Judge and Bock (1978) and Saleh (2006).

If we put $\lambda_1 = \theta$ and $\lambda_2 = (1 - \theta)$ in (2.7), we get the balanced loss function proposed by Zellner (1994). For $0 \leq \theta < 1$, a sufficient condition that the Stein-rule estimators perform better than the ordinary least squares estimator is when

$$k < 2(1 - \theta) \left(\frac{\gamma_1}{\sigma} \bar{X}'\beta + p - 2 \right). \quad (3.20)$$

When γ_1 is zero and/or \bar{X} is a null vector, the condition (3.20) assumes a simple form:

$$k < 2(1 - \theta)(p - 2); p > 2. \quad (3.21)$$

This serves as a sufficient condition for the superiority of Stein-rule estimators over ordinary least squares estimator in case of asymmetric distributions of disturbances provided that skewness coefficient γ_1 has same sign as $\bar{X}'\beta$. Further, if (3.20) holds true, one can find Stein-rule estimators that are better than ordinary least squares estimator even when there is simply one or two explanatory variables in the model.

It may be observed that the condition (3.20) has been derived by Giles, Giles and Ohtani (1996) by considering the exact risk under the normality of disturbances; see also Ohtani (1998).

Now we consider a case when $\lambda_1 = \lambda^2$ and $\lambda_2 = (1 - \lambda)^2$ in (2.7). We obtain the loss function

$$\lambda^2(X\tilde{\beta} - y)'(X\tilde{\beta} - y) + (1 - \lambda)^2(\tilde{\beta} - \beta)'X'X(\tilde{\beta} - \beta) + 2\lambda(1 - \lambda)(X\tilde{\beta} - y)'X(\tilde{\beta} - \beta).$$

It is a combination of sum of squares of the residuals, weighted sum of squares of estimation errors and weighted sum of cross products of residuals and estimation errors. This is also equal to the sum of squares of prediction errors when $X\tilde{\beta}$ is employed for the prediction of a convex combination of the actual and average values of study variable which was proposed by Shalabh (1995).

From (3.5) and (3.11), it is seen that the Stein-rule estimators have smaller risk in comparison to ordinary least squares estimator when

$$k < 2(1 - \lambda) \left(\frac{\gamma_1}{\sigma} \bar{X}'\beta + p - 2 \right) \quad (3.22)$$

which is precisely the same as (3.20) obtained from risk comparison under the balanced loss function. The condition of (3.22) with $\gamma_1 = 0$ matches with the condition of Shalabh (1995) on the basis of exact risk; see also Shalabh (1999).

4 Simulation study

The large sample asymptotic approximation theory is employed to study the risk of estimators in Section 3. The large sample theory gives an idea about the behaviour of risk of estimators in large samples. In order to study the behaviour of risk of estimators in finite samples, we conducted a Monte-Carlo simulation experiment. The main objective of the Monte-Carlo simulation experiment is to study the superiority of Stein-rule and ordinary least squares estimators over each other, effect of departure from the normal distribution and performance of estimators under the proposed loss functions in finite samples. The following set up is considered for the simulation study. The sample size was $n = 40$ and number of variables in X (number of columns of X) is $p = 6$. The matrix X is generated from a multivariate normal distribution with mean 1 for the first 3 components and mean -1 for the last 3 components. The covariance matrix of X was directly chosen as a correlation matrix with correlation 0 among the 6 columns (independence). The true (6×1) parameter vector β was set to $\beta = (-5.0, -4.9, \dots, 0.0, \dots, 4.9, 5.0)'$ (101 different values). All components of β were accordingly set to the same parameter value. The true response (without any error ϵ) was calculated according to the linear relationship $y_{\text{true}} = X\beta$ under this set up. A new response vector is generated according to the model (2.1), i.e.,

$y = y_{\text{true}} + \epsilon$ in each run. The error distribution is chosen to be an unsymmetrical beta distribution with parameters $\frac{1}{2}$ and $\frac{1}{10}$; and a symmetrical t -distribution with 5 degrees of freedom. The parameter k of the Stein-rule estimator is chosen as $k = p - 2 = 4$. The empirical risk is then calculated based on 1000 runs.

The random variables from beta and t -distributions were suitably scaled to have same mean and same variance. The 1000 runs were repeated for each of the combinations under the settings $\lambda_1 = 0.0, 0.1, \dots, 1.0$ (11 different values), $\lambda_2 = 0.0, 0.1, \dots, 1.0$ (11 different values) and $\lambda_3 = 1 - \lambda_1 - \lambda_2$.

Each setting was repeated 1000 times including a new generation of the design matrix X for each setting of β , λ_1 and λ_2 . Therefore we simulated the results from 12221 (101×11^2) different β - λ_1 - λ_2 -combinations. Each of these combinations was tested with 1000 different design matrices X and risks of regression coefficients were estimated. Within each of these 1000 simulations, another loop was embedded for the generation of new error vectors to estimate the empirical risk of ordinary least squares and Stein-rule estimators.

We have plotted the three dimension surface plots to visualize the performance of ordinary least squares and Stein-rule estimators with respect to λ_1 , λ_2 and β . Only some representative surfaces are presented here keeping in mind the length of paper.

Figure 1 and 2 show the risks of ordinary least squares and Stein-rule estimators, respectively under beta distributed errors for every β - λ_2 combination for a given specific value of λ_1 and a randomly chosen design matrix X . In figure 3, we counted how many times the Stein-rule estimator was better than the ordinary least squares estimator in all 1000 simulations with different X design matrices (since the risk of both estimators depends on the design matrix). Figures 4, 5 and 6 use another visualization showing the risk for all λ_1 - λ_2 combinations for a given specific value β . Figures 7-12 show the similar results as in figures 1-6 for

t -distributed errors.

We recall that the values of λ_1 and λ_2 indicate the weights assigned to goodness of fit and precision of estimation, respectively in the following discussion.

First we analyze the performance of ordinary least squares and Stein-rule estimators under beta distributed errors from figures 1-6. Different graphs in figure 1 give an idea about the performance of ordinary least squares estimator when λ_2 and β vary whereas λ_1 stays fixed. In such a case of given λ_1 , the perturbations in the surfaces are smaller for lower values of λ_1 and higher for larger value of λ_1 . The perturbations increase as both λ_1 and λ_2 simultaneously increase. The slopes of the surfaces remain almost unchanged with respect to β which indicates the independence of results from the values of true parameters. Similar results and trends are also observed in the surfaces of risk of Stein-rule estimator in figure 2 but the perturbations in the surfaces are higher in comparison to the perturbations in the surface of ordinary least squares estimator as in figure 1. This clearly indicates an increase in the risks of ordinary least squares and Stein-rule estimators with an increase in the value of λ_2 . This confirms the dependence of risks on the criteria of goodness of fit, precision of estimation as well as their joint effect. There is a sudden downward cusp in the surfaces of risk of Stein-rule estimator at $\beta = 0$. The depth of such cusp heavily depends on the values of λ_1 and λ_2 . It increases as λ_2 increases for a fixed λ_1 and decreases as λ_1 increases for a fixed λ_2 . The depth of cusp decreases when both λ_1 and λ_2 simultaneously increase. This suggests that an attention is needed while dealing with the Stein-rule estimator when true regression slope is expected to be around zero. Now we observe the number of cases under which the Stein-rule estimator is better than the ordinary least squares estimator from figure 3. The superiority of Stein-rule estimator over ordinary least squares estimator heavily depends on the values of λ_1 and λ_2 . When λ_1 is low, then the number of cases of superiority of Stein-rule

estimator over ordinary least squares estimator increases as λ_2 increases. When λ_1 also increases together with λ_2 , then the number of cases of superiority of Stein-rule estimator over ordinary least squares estimator also decreases. An important conclusion appears that the shapes of surfaces under the cases when $\lambda_1 = 0$ and when $\lambda_2 = 0$ are opposite in nature. This clearly indicates that the roles of goodness of fit and precision of estimation are very different and a decision about the superiority of Stein-rule and ordinary least squares estimators over each other significantly depends on the choice of decision criterion. A continuously changing performance of Stein-rule estimator over ordinary least squares estimator can be clearly observed from figure 3. We observe the dependence of risks of ordinary least squares and Stein-rule estimators on the true values of regression coefficients from the figures 4 and 5. There are 5 graphs in each of the figures 4 and 5 for different values of β . All graphs are almost same under respective estimators. This indicates that our conclusions about the effect of λ_1 and λ_2 on the risk performance of ordinary least squares and Stein-rule estimators are independent of true β . It appears that in the case of Stein-rule estimator, there are minor variations when $\beta = 0$ in comparison to the other cases when $\beta \neq 0$. The figures 4 and 5 also show the joint dependence of λ_1 and λ_2 on the risks of ordinary least squares and Stein-rule estimators. It is clearly seen that the risks of ordinary least squares and Stein-rule estimators increase as λ_1 and λ_2 simultaneously increase irrespective of the value of β . We observe the superiority of Stein-rule estimator over ordinary least squares estimator with respect to λ_1 and λ_2 jointly for a given β from figures 4 and 5. It is clear from the slopes of surfaces that Stein-rule estimator is better than ordinary least squares estimator when λ_2 is high and λ_1 is low. When λ_1 increases and λ_2 decreases simultaneously, then the number of cases of superiority of Stein-rule estimator over ordinary least squares estimator decreases. The rate of decrement does not depends on the value of β except for the case when $\beta = 0$. There is a sharp change in the number of cases of superiority when β is near to

zero.

Now we observe the figures 7-12 which shows the performance of ordinary least squares and Stein-rule estimators under t -distributed errors. The pattern of behaviour of ordinary least squares and Stein-rule estimators under t -distributed errors is nearly same as under beta distributed errors. So all the conclusions which are drawn from the figures 1-6 about the performance of ordinary least squares and Stein-rule estimators as well as the roles of λ_1 and λ_2 also follow from figures 7-12 and are valid for the case of t -distributed errors. The main difference in the respective figures is that the surfaces of plots under t -distributed errors have more perturbations than the surfaces of plots under beta distributed errors. For instance, comparing the respective plots in figures 1 and 7, we observe that the corresponding surfaces have more perturbations in figure 7 than in figure 1. Same is true for the figures 2 and 8. This clearly indicates the effect of departure from normality in terms of the departure from symmetry and peakedness of the distributions of errors on the performance of ordinary least squares and Stein-rule estimators. The degree of perturbations depends on the degree of departure from symmetry and peakedness. This also confirm the analytical findings about the effect of departure from normality.

5 Some remarks

Appreciating the simultaneous use of the two performance criteria, viz, the goodness of fitted model and the concentration of estimates around the true parameter values, for judging the efficiency of any estimation procedure for the coefficients in a linear regression model, we have presented a general loss function using the quadratic loss structure. Several popular loss functions are found to be the particular cases of it, and thus the properties of loss function are fairly general and

sufficiently flexible.

For the regression coefficient vector, we have considered the unbiased ordinary least squares and biased Stein-rule estimators. We have compared their performance according to the risk criterion under the proposed loss function and have obtained a condition on the characterizing scalar for the superiority of Stein-rule estimators over ordinary least squares estimator. It is a well known result that Stein-rule estimator is better than ordinary least squares estimator when the number of explanatory variables are more than two. We find that under such an extended balanced loss function, it is possible to find Stein-rule estimator which has a better performance than ordinary least squares estimator even when the number of explanatory variables are one or two only.

The finite sample properties from simulation results confirm the analytical findings. The performance of ordinary least squares and Stein-rule estimators very much depends on the choice of performance criterion. The effect of departure from normality is also confirmed by the simulation results.

Acknowledgement

The first author gratefully acknowledges the support from Alexander von Humboldt Foundation, Germany in the form of Humboldt Fellowship.

References

- Akdeniz, Fikri, Wan, Alan T. K. and Akdeniz, Esra (2005): ‘Generalized Liu type estimators under Zellner’s balanced loss function’, Communications in Statistics- Theory and Methods, 34, no. 8, pp. 1725-1736.
- Chaturvedi, Anoop and Shalabh (2004): ‘Risk and Pitman closeness properties of feasible generalized double k -class estimators in linear

- regression models with non-spherical disturbances under balanced loss function', *Journal of Multivariate Analysis*, 90, no. 2, 229-256.
- Chung, Younshik, Kim, Chansoo and Song, Seongho (1998): 'Linear estimators of a Poisson mean under balanced loss functions', *Statistics and Decisions*, 16, no. 3, pp. 245-257.
- Chung, Younshik, Kim, Chansoo and Dey, D. K. (1999): 'A new class of minimax estimators of multivariate normal mean vectors under balanced loss function', *Statistics and Decisions*, 17, no. 3, pp. 255-266.
- Giles, J.A., Giles, D.E.A. and Ohtani K. (1996): 'The exact risks of some pre-test and Stein-type regression estimates under balanced loss', *Communications in Statistics - Theory and Methods*, 25, pp. 2901-2924.
- Gruber, Marvin H. J. (2004): 'The efficiency of shrinkage estimators with respect to Zellner's balanced loss function', *Communications in Statistics- Theory and Methods*, 33, no. 2, pp. 235-249.
- Judge, G.G. and Bock, M.E. (1978): '**The Statistical Implication of Pre-Test and Stein-Rule Estimators in Econometrics**', North Holland.
- Ohtani, K., Giles, D. E. A. and Giles, J. A. (1997): 'The exact risk performance of a pre-test estimator in a heteroskedastic linear regression model under the balanced loss function', *Econometric Reviews*, 16, no. 1, pp. 119-130.
- Ohtani, K. (1998): 'The exact risk of a weighted average estimator of the OLS and Stein-rule estimators in regression under balanced loss', *Statistics and Decisions*, 16, pp. 35-45.
- Ohtani, K. (1999): 'Inadmissibility of the Stein-rule estimator under

- the balanced loss function', Journal of Econometrics, 88 , no. 1, pp. 193-201.
- Rao, C. R., Toutenburg, H., Shalabh and Heumann, C. (2008). *Linear Models and Generalizations: Least Squares and Alternatives*, 3rd edition. Springer.
- Rodrigues, J.; Zellner, A. (1994): 'Weighted balanced loss function and estimation of the mean time to failure', Communications in Statistics- Theory and Methods, 23, no. 12, pp. 3609-3616.
- Saleh, A.K.E. (2006): '**Theory of Preliminary Test and Stein-type Estimation with Application**', John Wiley.
- Sanjari Farsipour, N. and Asgharzadeh, A. (2004): 'Estimation of a normal mean relative to balanced loss functions', Statistical Papers, 45, no. 2, pp. 279-286.
- Shalabh (1995): 'Performance of Stein-rule procedure for simultaneous prediction of actual and average values of study variable in linear regression model', Jan Tinbergen Award Paper, Bulletin of the 50th Session of the International Statistical Institute, pp. 1375-1390.
- Shalabh (1999): 'Improving the predictions in linear regression models', Journal of Statistical Research, 33, pp. 33-39.
- Shalabh (2000): 'Prediction of values of variables in linear measurement error model', Journal of Applied Statistics, 27, pp. 475-482.
- Shalabh (2001): 'Least squares estimators in measurement error models under the balanced loss function', TEST, 10, no. 2, pp. 301-308.
- Srivastava, A.K. and Shalabh (1996): 'Efficiency properties of least squares and Stein-rule predictions in linear regression model', Journal of Applied Statistical Science, 4, pp. 141-145.

- Toutenburg, H. and Shalabh (1996): ‘Predictive performance of the methods of restricted and mixed regression estimators’, *Biometrical Journal*, 38, pp. 951-959.
- Toutenburg, H. and Shalabh (2005): ‘Estimation of regression coefficients subject to exact linear restrictions when some observations are missing and quadratic error balanced loss function is used’, *TEST*, 14, no. 2, pp. 385-396.
- Vinod, H.D. and Srivastava, V.K. (1995): ‘Large sample asymptotic properties of the double k-class estimators in linear regression models’, *Econometric Reviews*, 14, pp. 75-100.
- Wan, A. T. K. (1994): ‘Risk comparison of the inequality constrained least squares and other related estimators under balanced loss’, *Economic Letters*, 46, no. 3, pp. 203-210.
- Wan, Alan T. K. (2002): ‘On generalized ridge regression estimators under collinearity and balanced loss’, *Applied Mathematics and Computation*, 129, no. 2-3, 455-467.
- Zellner, A. (1994): ‘Bayesian and non-Bayesian estimation using balanced loss functions’ in *Statistical Decision Theory and Related Topics*, V (West Lafayette, IN, 1992), pp. 377-390, Springer, New York.

Figure 1: Risk of ordinary least squares estimator when errors are beta distributed

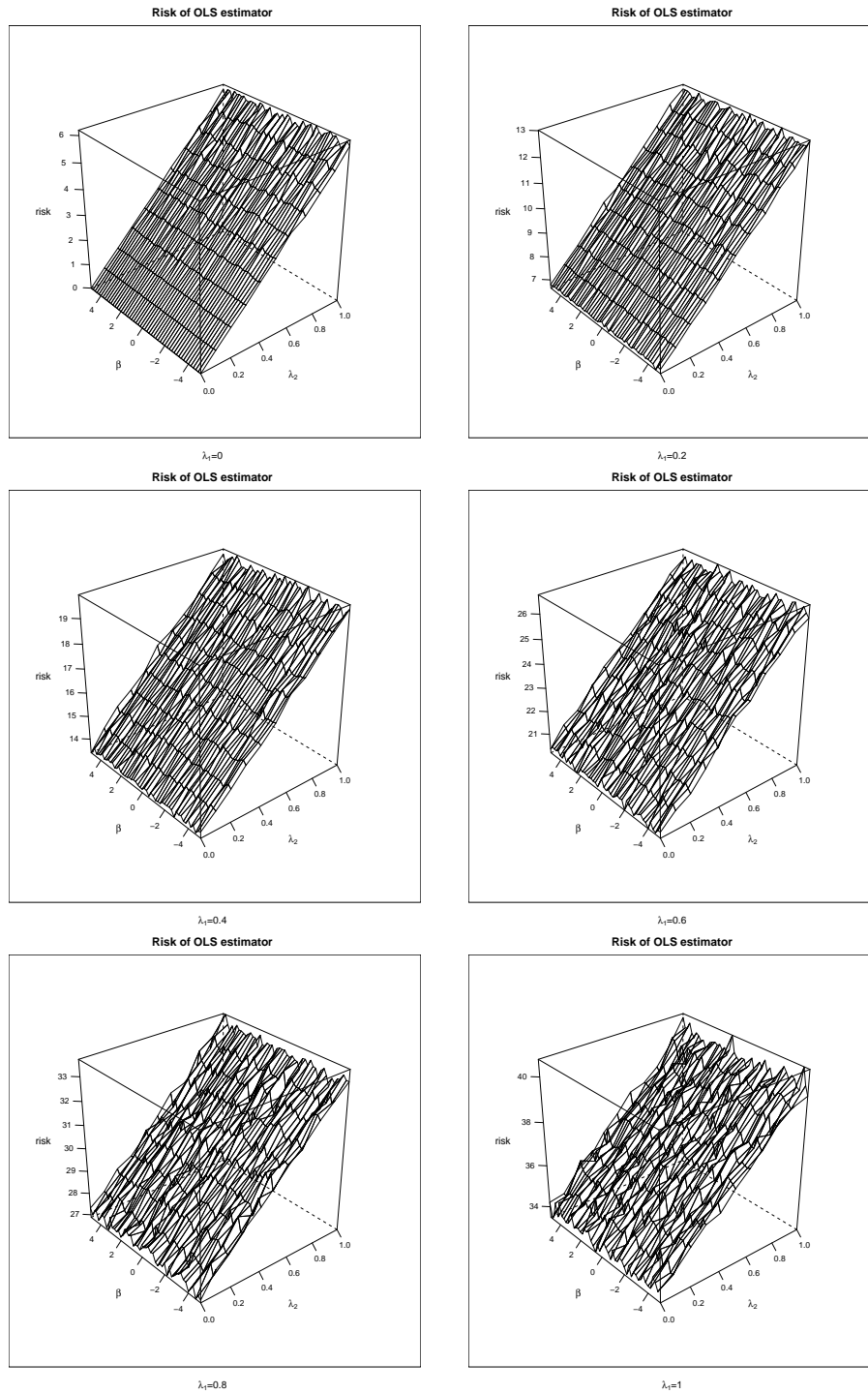


Figure 2: Risk of Stein-rule estimator when errors are beta distributed

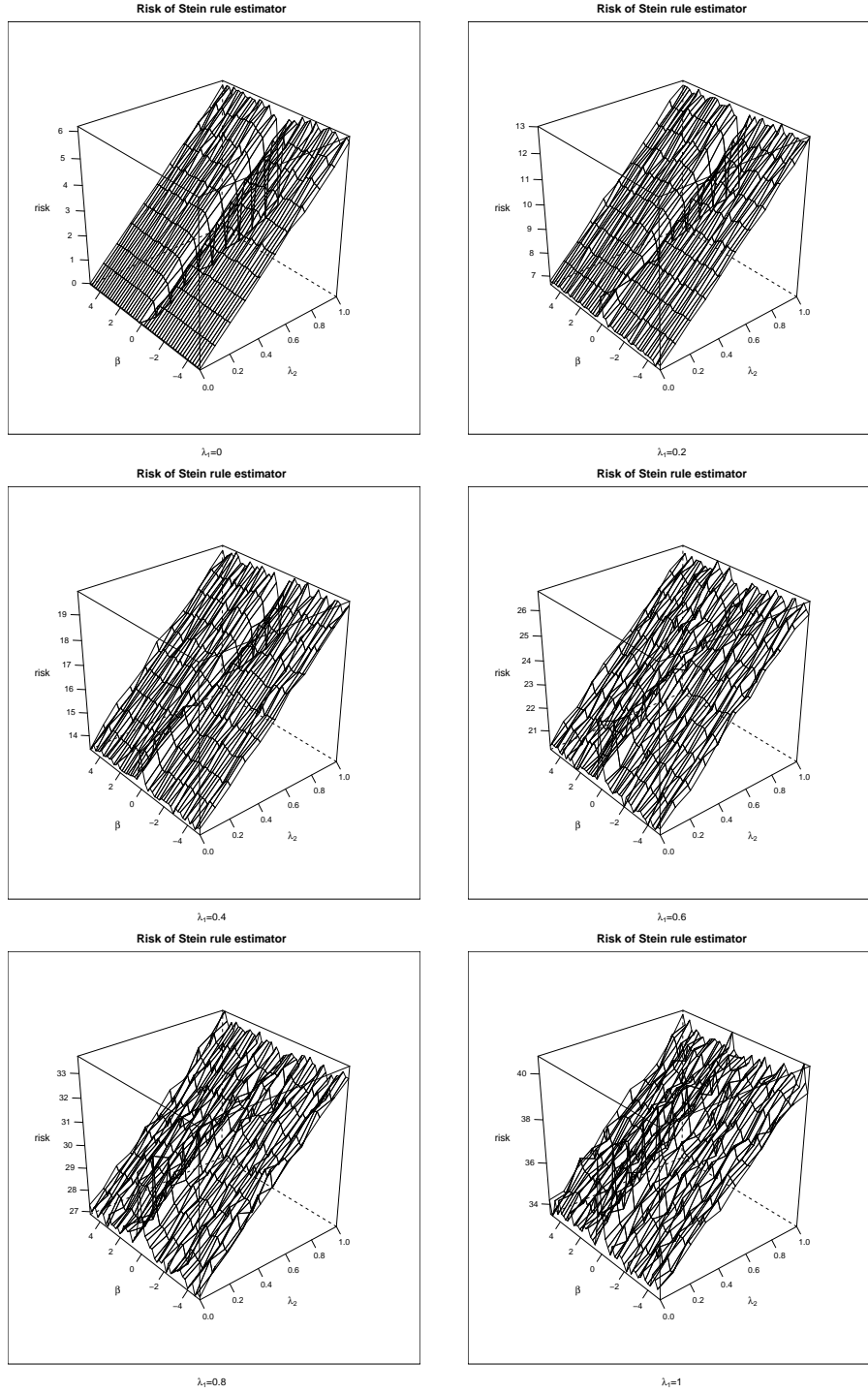


Figure 3: Number of times where Stein-rule estimator (STR) is better than ordinary least squares estimator (Beta distribution, 1000 runs)

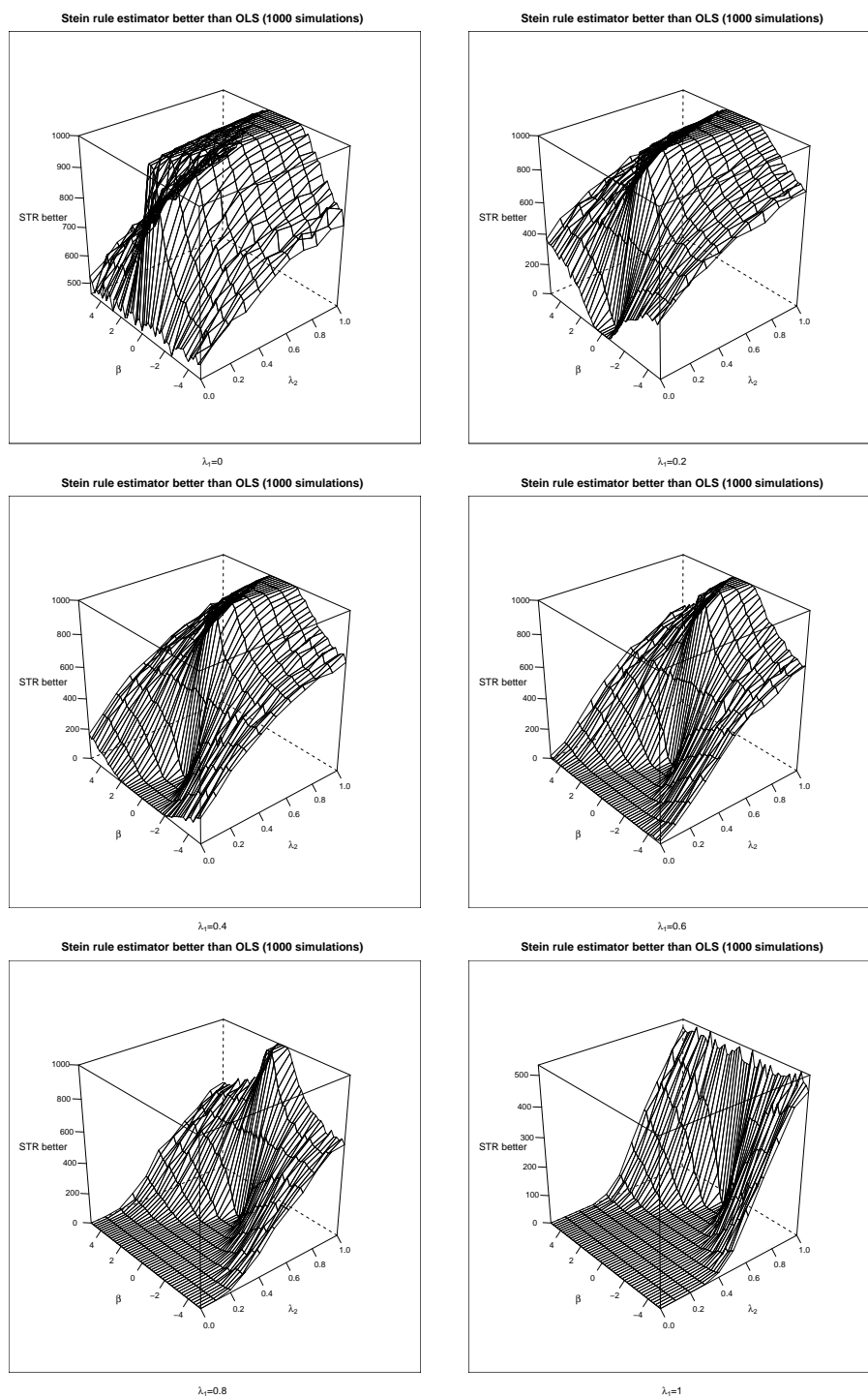


Figure 4: Risk of ordinary least squares estimator when errors are beta distributed

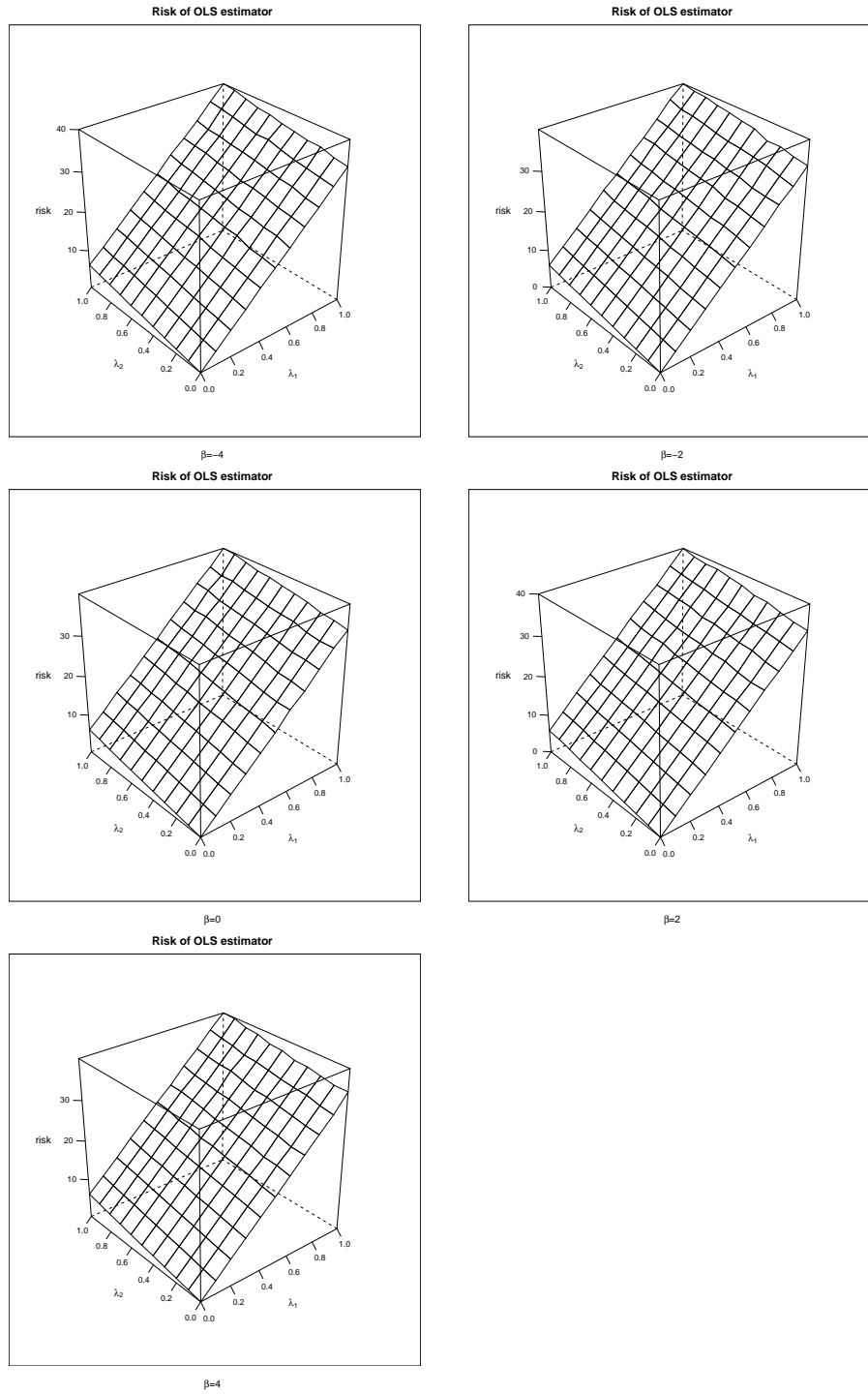


Figure 5: Risk of Stein-rule estimator when errors are beta distributed

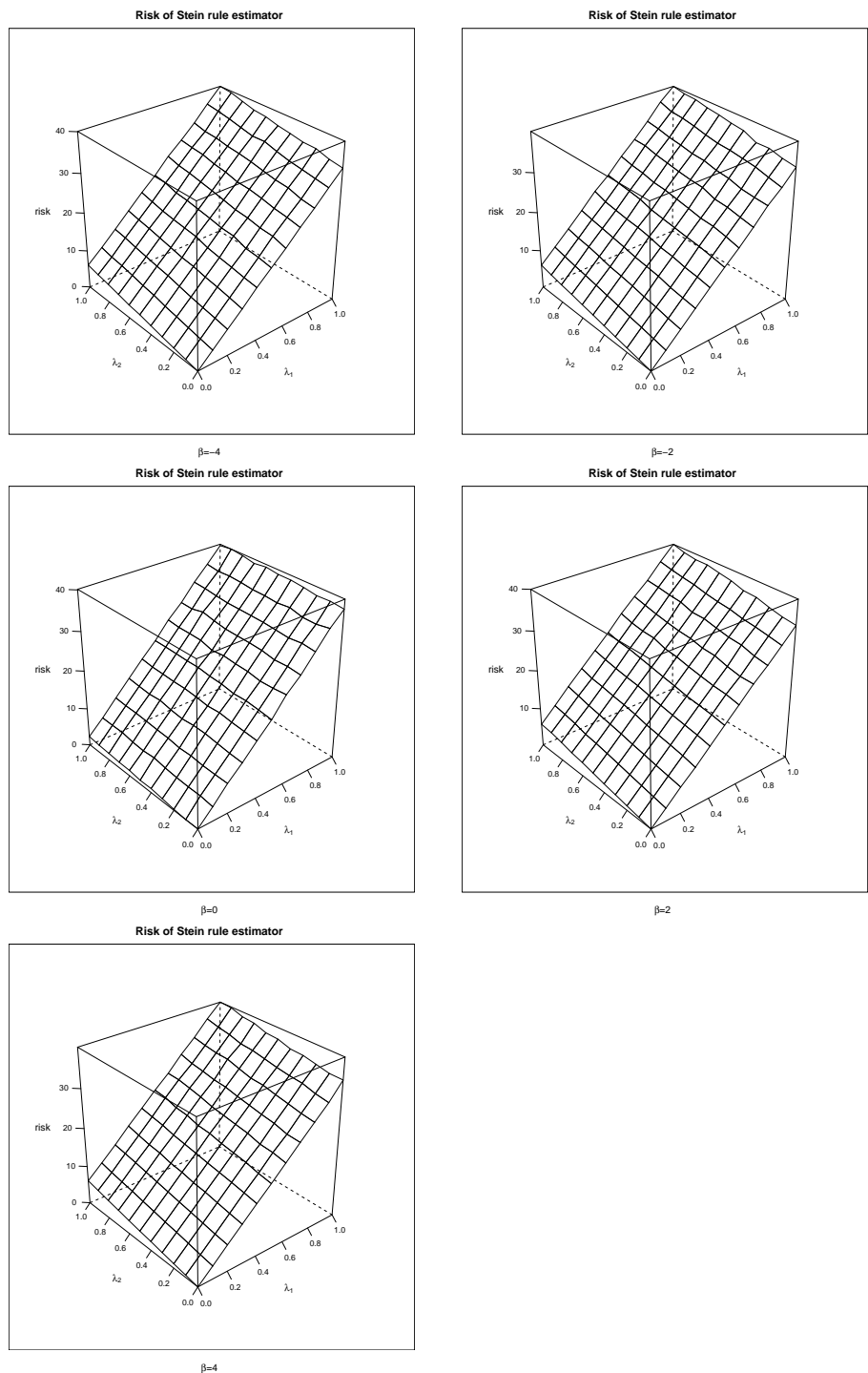


Figure 6: Number of times where Stein-rule estimator (STR) is better than ordinary least squares estimator (Beta distribution, 1000 runs)

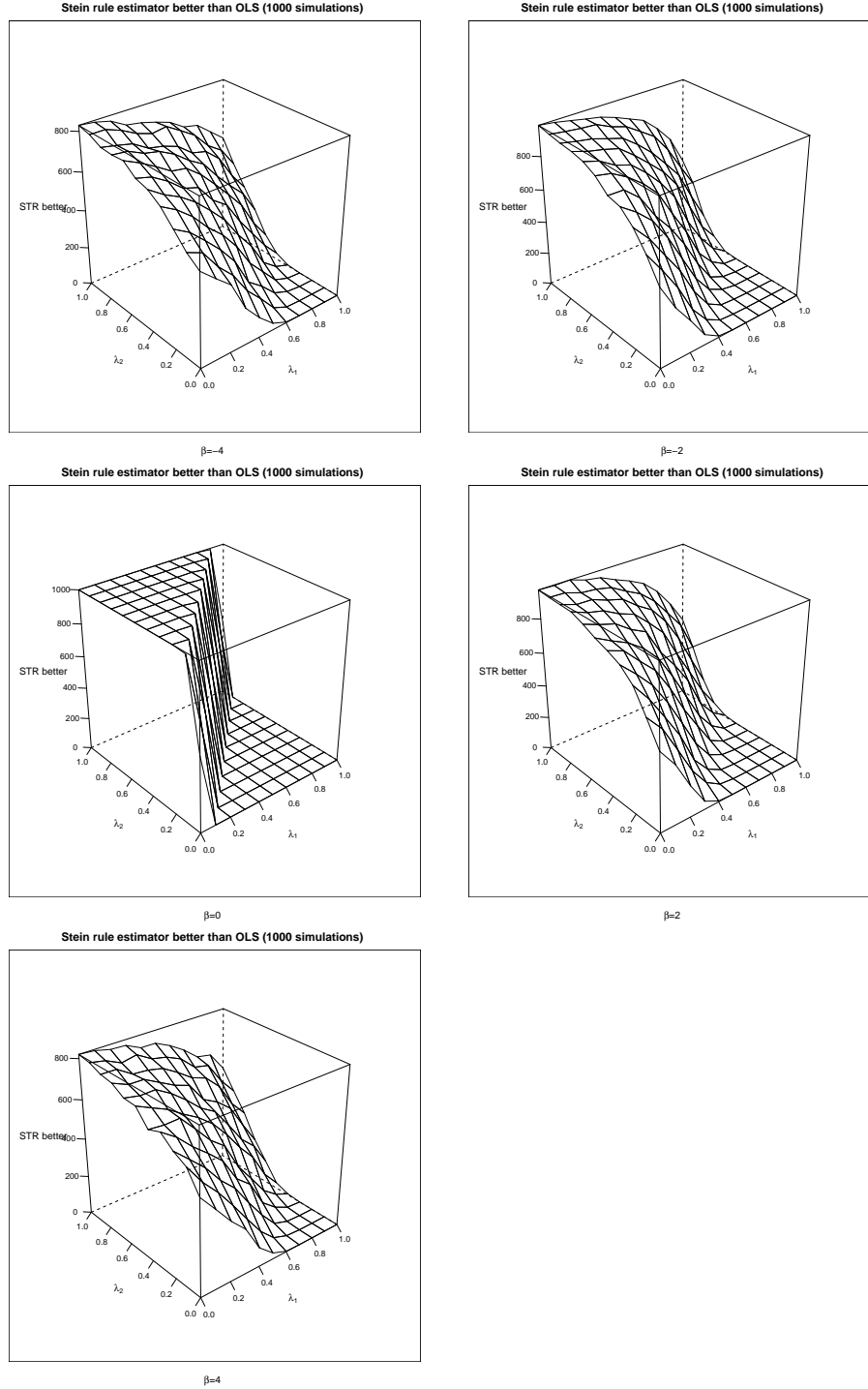


Figure 7: Risk of ordinary least squares estimator when errors are t -distributed

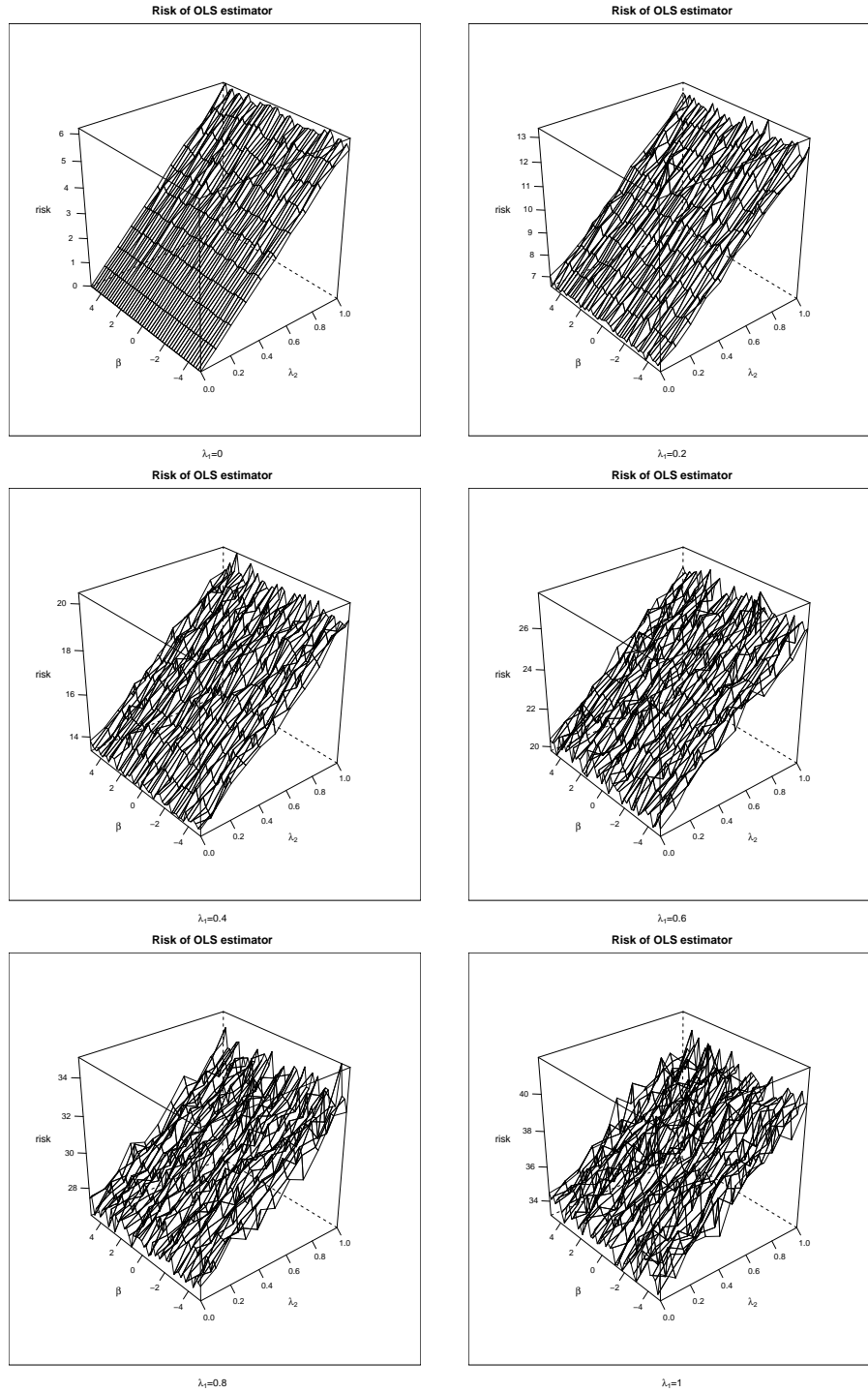


Figure 8: Risk of Stein-rule estimator when errors are t -distributed

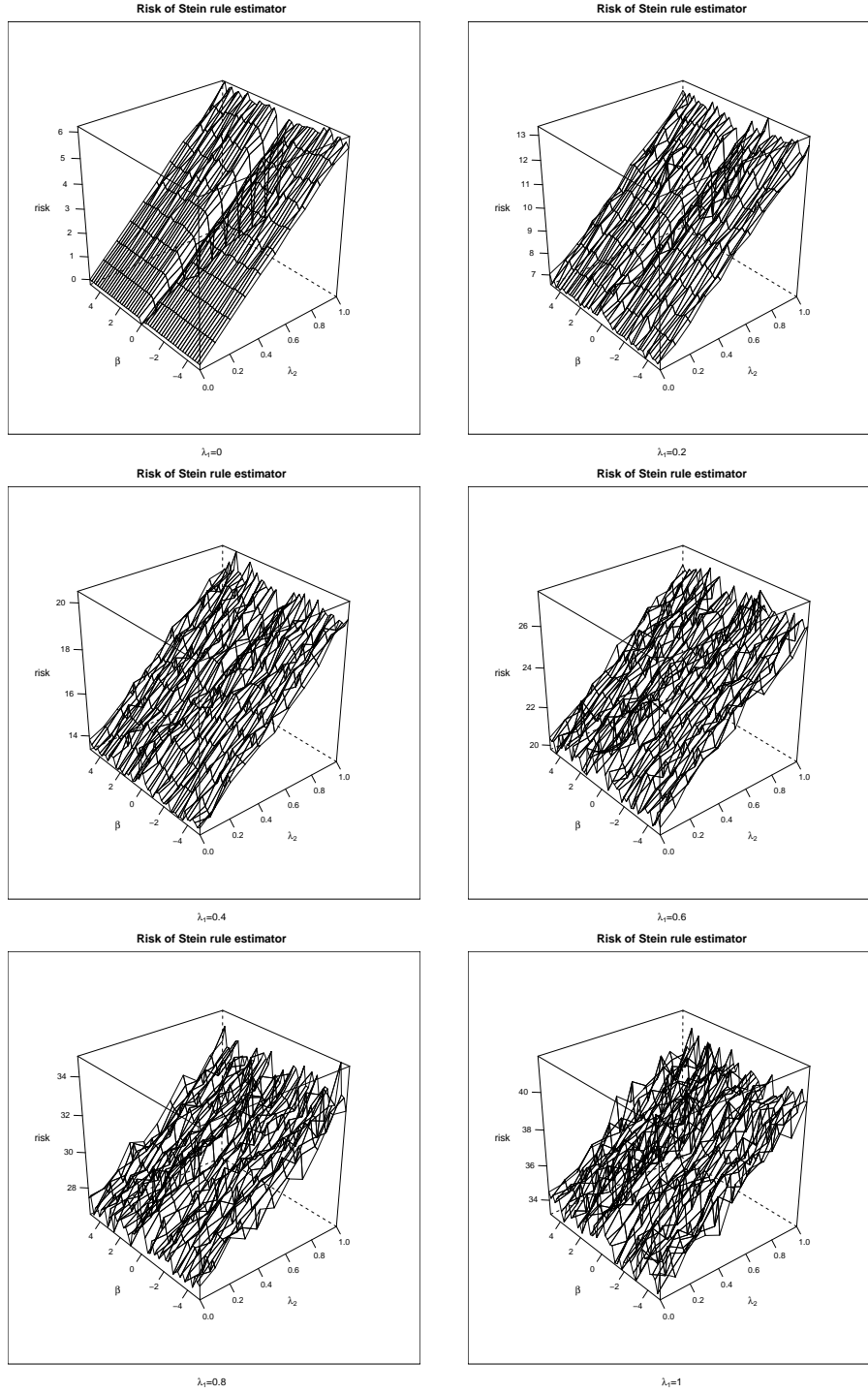


Figure 9: Number of times where Stein-rule estimator (STR) is better than ordinary least squares estimator (t -distribution, 1000 runs)

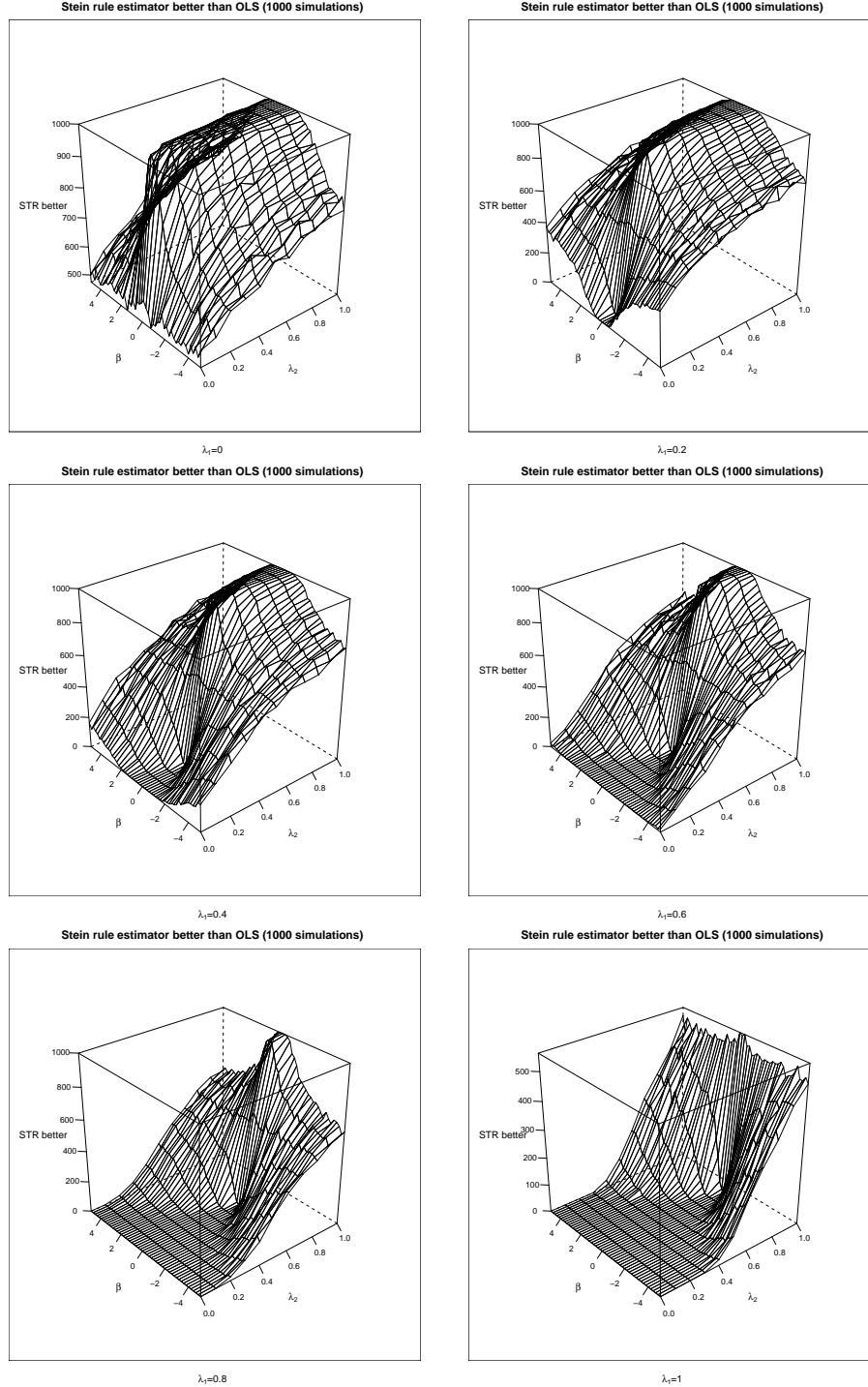


Figure 10: Risk of ordinary least squares estimator when errors are t -distributed

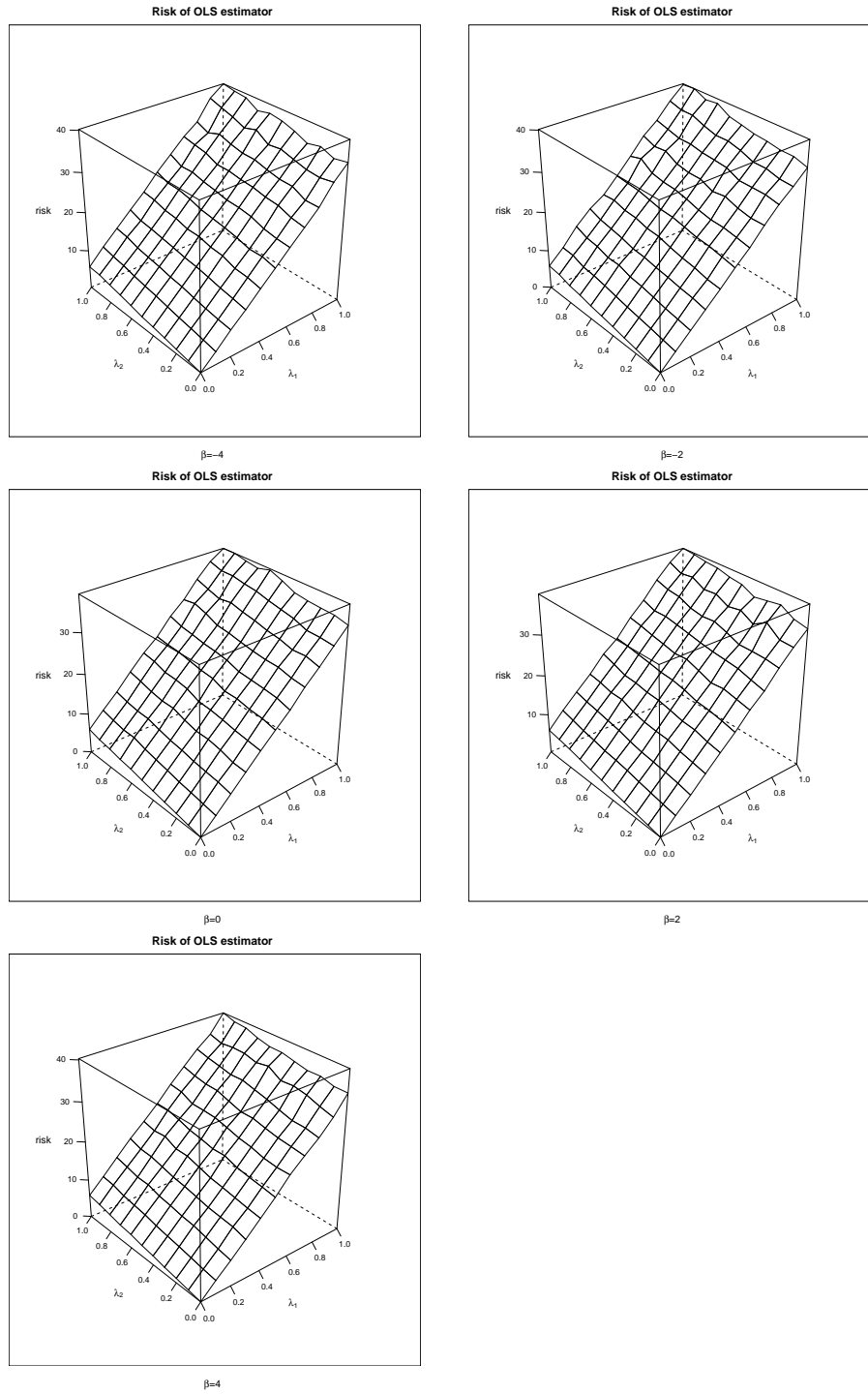


Figure 11: Risk of Stein-rule estimator when errors are t -distributed

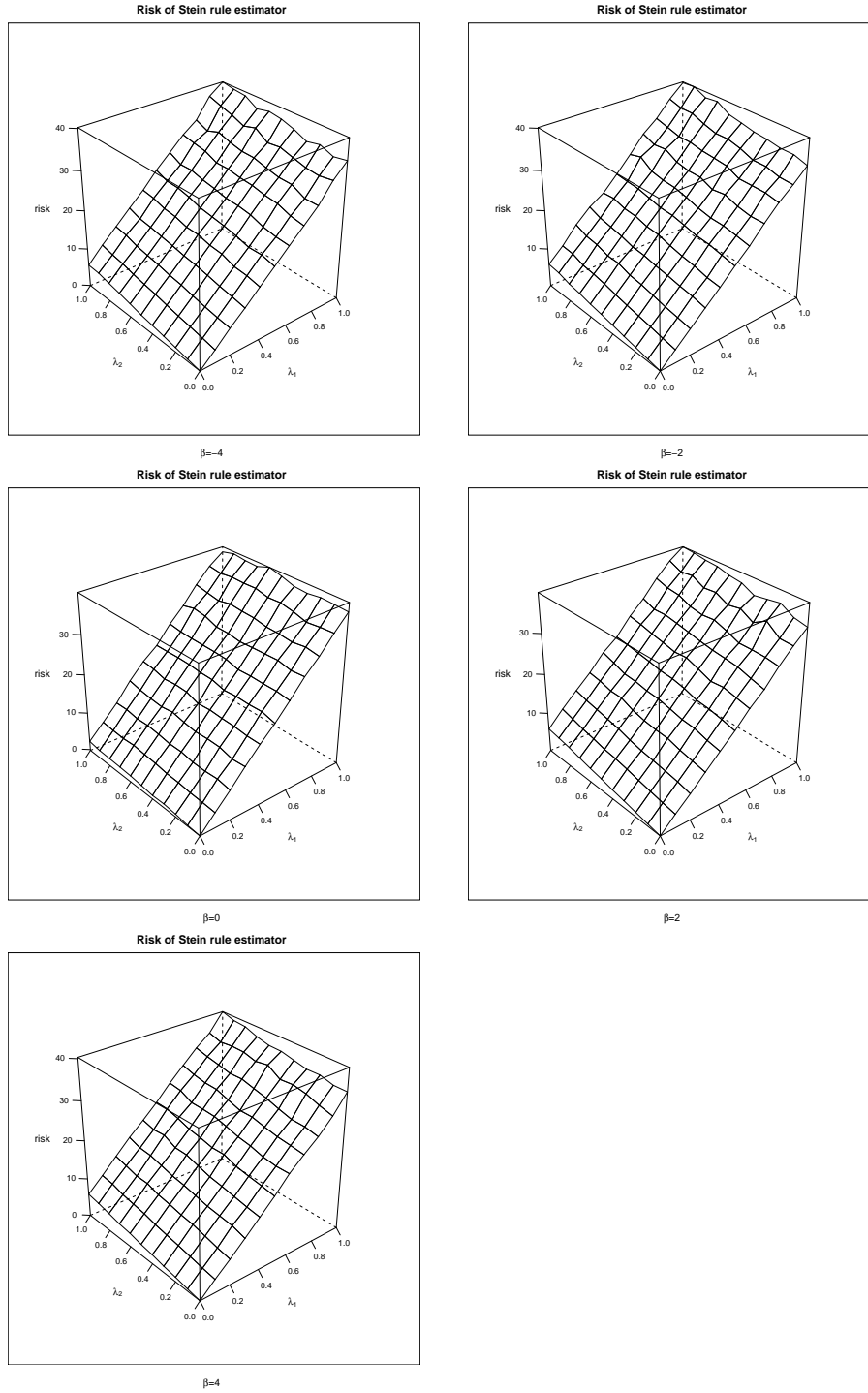


Figure 12: Number of times where Stein-rule estimator (STR) is better than ordinary least squares estimator (t -distribution, 1000 runs)

