THE PERFORMANCE OF SERIAL CORRELATION PRELIMINARY TEST ESTIMATORS UNDER ASYMMETRY LOSS FUNCTIONS

S. N. Nakale [1]

Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa.

e-mail: *sammynakale@gmail.com*

J. Kleyn

Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa.

M. Arashi

Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa.

Department of Statistics, School of Mathematical Sciences, University of Shahrood, Shahrood, Iran.

A. Bekker

Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa.

Key words: Serial correlation, preliminary test estimation, LINEX loss, BLINEX loss, risk function.

Summary: The risk performances, under the symmetric squared error loss function, of the estimators of the regression coefficients after a preliminary test for serial correlation have been widely investigated in the literature. However, it is well known that the use of the symmetric loss functions is inappropriate in estimation problems where underestimation and overestimation have different consequences. We consider the Linear Exponential and Bounded Linear Exponential loss functions which allows for asymmetry. The risks of the estimators are derived and numerically evaluated by using simulations.

---

[1] Corresponding author.

AMS: 62M10, 62H15, 62C99

1. Introduction

The classical linear regression model assumption of no serial correlation in the error terms is often not plausible in time series regression models. Cochrane and Orcutt (1949), in particular, found that error terms in most economic models are positively correlated. Under the uncertainty of the validity of this assumption, serial correlation test procedures such as the Durbin and Watson (1950, 1951, 1971) is used to test for the significance of the serial correlation in the error terms. If the hypothesis of no serial correlation is accepted, then the regression coefficients are estimated by the ordinary least squares (OLS), otherwise a feasible generalised least squares (FGLS) estimator that corrects for serial correlation is considered. This estimation strategy is what is referred to in the literature as the serial correlation preliminary test (pretest) estimation.

The sampling performance of the serial correlation pretest estimator relative to its component estimators, OLS and FGLS, have been well investigated in the literature (see Morey, 1975; Judge and Bock, 1978; Fomby and Guilkey, 1978; King and Giles, 1984; Folmer, 1988, etc.). Comparisons of the estimators' performances in these studies are based on risks under the symmetric squared error (SE) loss function. Monte Carlo results from these studies are in agreement that there are significant risk gains by considering the pretest estimator, compared to the OLS estimator for serial correlation coefficient (ρ) approximately greater than 0.3. Moreover, the pretest estimator compares favorably to the FGLS estimator for small values of ρ, and performs just as well for medium and larger values of ρ.

However, in some estimation problems under- and overestimation estimation of the regression coefficients do not have the same consequences as implied by the use of the symmetric loss functions. In economic models, for example, underestimation and overestimation of a regression coefficient are more likely to have different consequences in terms of the appropriate policy tools to be applied and the implications thereof. Therefore, as recognised by Varian (1975), Zellner (1986) and Wen and Levy (2001), the symmetric loss functions such as the SE loss are inappropriately used in some estimation problems. In this paper we consider the asymmetry loss functions, Linear-Exponential (LINEX) as proposed by Varian (1975) and the Bounded Linear-Exponential (BLINEX) introduced by Wen and Levy (2001). We derive the risks of the OLS, GLS, FGLS and pretest estimators under the first order autoregressive error terms specification. The risk properties under asymmetry and the effect of the loss asymmetry are evaluated numerically using Monte Carlo simulations.

The next section discusses the model, estimators and the test procedure under consideration. Section 3 presents the risk derivations, the numerical evaluations of the risk function and discussions follow in Section 4. An application is considered in Section 5, and we conclude in Section 6.

2. The estimation problem

Consider the classical linear regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{1}$$

where $\mathbf{y}$ is a $T \times 1$ vector of observations on a dependent variable, $\mathbf{X}$ is a $T \times k$ nonstochastic design matrix of full column rank k, $\boldsymbol{\beta}$ is a $k \times 1$ vector of unknown regression coefficients and $\mathbf{u}$ is a $T \times 1$

vector of the error terms. We are concerned with estimating the regression coefficients when the error terms are (as frequently assumed in applied econometric modelling) generated by a stationary first order autoregressive process

$$u_t = \rho u_{t-1} + e_t, \qquad (e_1, e_2, ..., e_T)' \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I}), 0 \leq \rho < 1.$$

Under this assumption, it can be shown (see, for example, Gujarati, 2003) that

$$\mathbf{u} \sim N(\mathbf{0}, \sigma_e^2 \boldsymbol{\phi})$$

where

$$\boldsymbol{\phi} = \frac{1}{(1-\rho^2)} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{bmatrix}_{(T \times T)}.$$

It is well known that under serial correlation, the OLS estimator does not have optimal statistical properties. For example, for positive serially correlated error terms, the most likely situation with economic time series, Judge, Griffiths, Hill, Lutkepohl and Lee (1985) have shown that the variances of the OLS regression coefficients are biased downward. There are a number of FGLS estimator that have been proposed to correct for first order autoregressive error terms. The most commonly used are the Cochrane and Orcutt (1949), the Prais and Winsten (1954), Durbin's (1960) and the Maximum Likelihood (ML) estimator of Beach and Mackinnon (1978). The Cochrane-Orcutt, Prais-Winsten and Durbin estimators differ in the estimate of ρ used and the transformation method applied. Monte Carlo evidence from Rao and Griliches (1969) and Judge and Bock (1978), among others, suggest that FGLS estimators are relatively more efficient than the OLS estimator for $|\rho| \geq 0.3$.

However, in empirical work the value of ρ is unknown, that is, it is not known whether the level of serial correlation is significant or not. Therefore, it is necessary to perform a preliminary test for serial correlation by testing the hypotheses

$$H_0 : \rho = 0 \qquad H_1 : \rho > 0.$$

The Durbin-Watson (1950; 1951; 1971) is the commonly used test procedure for serial correlation. The test statistic, d, is given by

$$d = \frac{\sum_{t=2}^{T} (\widehat{u}_t - \widehat{u}_{t-1})^2}{\sum_{t=1}^{T} \widehat{u}_t^2} \qquad \widehat{u}_t = (y_t - \mathbf{X}_t \widehat{\boldsymbol{\beta}}_{OLS}) \qquad t = 1, 2, ..., T. \tag{2}$$

If d_α is the critical value for a specified level of significance, for $d > d_\alpha$ H_0 is not rejected and if $d \leq d_\alpha$ H_0 is rejected. The preliminary test estimation strategy is to use the OLS estimator if we cannot reject the null hypothesis and to use a FGLS estimator if the null hypothesis is rejected.

Hence, the serial correlation pretest estimator is defined as

$$\widehat{\boldsymbol{\beta}}_{PTE} = I_A(d)\widehat{\boldsymbol{\beta}}_{OLS} + (1 - I_A(d))\widehat{\boldsymbol{\beta}}_{FGLS} \tag{3}$$

where

$$I_A(d) = \begin{cases} 1 & \text{if } H_0 \text{ is accepted} \\ 0 & \text{otherwise.} \end{cases}$$

The Durbin and the ML estimators are the most preferable FGLS estimators (see, King and Giles, 1984; Folmer, 1988). Consequently, we consider only these two estimators for the FGLS component of the pretest estimator. Fomby and Guilkey (1978) suggested that an optimal level of significance for the pretest should be about 50%.

As noted by Simons (1988), in the context of serial correlation preliminary test estimation, the usual hypothesis of linear restrictions in the general preliminary test estimation is replaced by the hypothesis of no serial correlation which then implies that the OLS estimator is now the restricted estimator whereas the FGLS estimator is the unrestricted estimator. For a discussion on the general preliminary test estimation and other related topics see Judge and Bock (1978), Saleh and Kibria (1993, 2011), Saleh (2006), Arashi, Tabatabaey and Hassanzadeh (2009) and Arashi (2009, 2012), among others.

3. Risk derivations

In any estimation problem, it is important to choose an appropriate loss function by taking the practical implications into consideration. Even though the symmetric SE loss function is the most frequently used in the literature, this is inappropriate in applications where the consequences of overestimation may be more serious than underestimation, or vice versa. Varian (1975) proposed the asymmetric LINEX loss function which has exponential losses on one side of zero and linear losses on the other. The LINEX loss is however unbounded. Wen and Levy (2001), on the other hand, introduced the BLINEX loss function which is asymmetric and bounded, and thus more useful in practical applications where it is required to have an upper bound on the loss. In this section we present the risk derivations of the OLS, GLS, FGLS and pretest estimators under the asymmetric LINEX and BLINEX loss functions. For the BLINEX loss, however, the nature of the loss function makes it difficult to derive the exact expressions of the risks, hence we only present approximated risk functions.

3.1. Risk under LINEX loss

Suppose $\widehat{\boldsymbol{\beta}}$ is an estimator of the unknown parameter vector $\boldsymbol{\beta}$. The LINEX loss takes the form

$$\mathscr{L}_{LINEX}(\widehat{\boldsymbol{\beta}};\boldsymbol{\beta}) = c[exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})] - \mathbf{a}'(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) - 1] \tag{4}$$

where $\mathbf{a} = (a_1, ..., a_k)'$ and $a_i \neq 0$ for $i = 1, ..., k$ and $c > 0$. The properties of the LINEX loss are discussed in Varian (1975). The parameter c is a scale parameter, and generally assumed to be equal to unity. The asymmetry vector parameter $\mathbf{a}$ determines the shape of the loss function, with the signs of the a_i's reflecting the direction of asymmetry and their magnitudes the degree of asymmetry. The LINEX loss is quite asymmetric for large values of $\|\mathbf{a}\|$ and almost symmetric for small values of $\|\mathbf{a}\|$. For simplicity, we will consider asymmetry vectors $\mathbf{a}$ where the values of the a_i's are the same.

3.1.1. Risk function of OLS estimator

The risk function of the OLS estimator, $\widehat{\boldsymbol{\beta}}_{OLS}$, under the LINEX loss is

$$\begin{aligned}\Re_{LINEX}(\widehat{\boldsymbol{\beta}}_{OLS};\boldsymbol{\beta}) &= E\left\{c\left[\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-1\right]\right\}\\ &= c\left[\exp(-\mathbf{a}'\boldsymbol{\beta})M_{\widehat{\boldsymbol{\beta}}_{OLS}}(\mathbf{a})-\mathbf{a}'(\boldsymbol{\beta}-\boldsymbol{\beta})-1\right]\\ &= c\left[\exp(-\mathbf{a}'\boldsymbol{\beta})\exp(\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})-1\right]\\ &= c\left[\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})-1\right] \qquad (5)\end{aligned}$$

where use is made of the fact that

$$M_{\widehat{\boldsymbol{\beta}}_{OLS}}(\mathbf{a})=\exp\left[\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}\right] \qquad (6)$$

and $\boldsymbol{\Sigma}_{OLS}=\sigma_e^2(\mathbf{X}'\mathbf{X})^{-1}$.

3.1.2. Risk functions of GLS and FGLS estimators

For the GLS estimator, $\widetilde{\boldsymbol{\beta}}_{GLS}$, the risk function under the LINEX loss is

$$\begin{aligned}\Re_{LINEX}(\widetilde{\boldsymbol{\beta}}_{GLS};\boldsymbol{\beta}) &= E\left[c\left[\exp[\mathbf{a}'(\widetilde{\boldsymbol{\beta}}_{GLS}-\boldsymbol{\beta})]-\mathbf{a}'(\widetilde{\boldsymbol{\beta}}_{GLS}-\boldsymbol{\beta})-1\right]\right]\\ &= c\left[\exp(-\mathbf{a}'\boldsymbol{\beta})M_{\widetilde{\boldsymbol{\beta}}_{GLS}}(\mathbf{a})-\mathbf{a}'(\boldsymbol{\beta}-\boldsymbol{\beta})-1\right]\\ &= c\left[\exp(-\mathbf{a}'\boldsymbol{\beta})\exp(\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a})-1\right]\\ &= c\left[\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a})-1\right] \qquad (7)\end{aligned}$$

where $M_{\widetilde{\boldsymbol{\beta}}_{GLS}}(\mathbf{a})=\exp\left[\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a}\right]$ and $\boldsymbol{\Sigma}_{GLS}=\sigma_e^2(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}$. Now since the FGLS estimator, $\widehat{\boldsymbol{\beta}}_{FGLS}$, is a GLS-type estimator, it then follows directly from equation (7) that the risk function of the FGLS estimator under the LINEX loss function is given by

$$\Re_{LINEX}(\widehat{\boldsymbol{\beta}}_{FGLS};\boldsymbol{\beta})=c\left[\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})-1\right] \qquad (8)$$

where $\boldsymbol{\Sigma}_{FGLS}=\sigma_e^2(\mathbf{X}'\widehat{\phi}^{-1}\mathbf{X})^{-1}$.

3.1.3. Risk function of the PTE

The following property of expected values is used in deriving the risk function of the pretest estimator:

$$E[A] \equiv E[E[A|B]], \text{ where } A \equiv \mathscr{L}(\widehat{\boldsymbol{\beta}}_{PTE};\boldsymbol{\beta}) \text{ and } B \equiv \text{outcome of the hypothesis test.} \tag{9}$$

Making use of equation (9), the risk function of the pretest estimator, $\widehat{\boldsymbol{\beta}}_{PTE}$, for a specified level of significance, η, is given by

$$\begin{aligned}
\Re_{LINEX}(\widehat{\boldsymbol{\beta}}_{PTE};\boldsymbol{\beta}) &= E\left[c\left[\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta})]-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta})-1\right]\right] \\
&= E\left[c\left[\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-1\right]|H_0\right]P(d>d_\eta) \\
&\quad +E\left[c\left[\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{FGLS}-\boldsymbol{\beta})]-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{FGLS}-\boldsymbol{\beta})-1\right]|H_1\right]P(d\le d_\eta) \\
&= \Re_{LINEX}(\widehat{\boldsymbol{\beta}}_{OLS};\boldsymbol{\beta})(1-\eta)+\Re_{LINEX}(\widehat{\boldsymbol{\beta}}_{FGLS};\boldsymbol{\beta})\eta, \text{ with } \eta=P(d\le d_\eta) \\
&= c\left[\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})(1-\eta)+\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})\eta-1\right] \text{ using equations (5) and (8)}
\end{aligned}$$

3.2. Risk under BLINEX loss

If $\widehat{\boldsymbol{\beta}}$ is the estimator of the unknown parameter vector $\boldsymbol{\beta}$ then the BLINEX loss function is defined as

$$\begin{aligned}
\mathscr{L}_{BLINEX}(\widehat{\boldsymbol{\beta}};\boldsymbol{\beta}) &= \frac{\mathscr{L}_{LINEX}(\widehat{\boldsymbol{\beta}};\boldsymbol{\beta})}{1+\lambda\mathscr{L}_{LINEX}(\widehat{\boldsymbol{\beta}};\boldsymbol{\beta})} \\
&= \frac{1}{\lambda}\left[1-\frac{1}{1+b[\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}))-\mathbf{a}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-1]}\right]
\end{aligned} \tag{10}$$

where $\mathbf{a}=(a_1,...,a_k)'$ and $a_i \neq 0$ for $i=1,...,k$, $\lambda>0$ and $b=\lambda c>0$. The BLINEX loss is bounded by 0 and $\frac{1}{\lambda}$ and the signs of the a_i's reflects the direction of asymmetry, with negative a_i's penalising the negative errors more heavily and vise versa for positive values of the a_i's. On the other hand, b is an asymmetry parameter and its magnitude determines the direction of asymmetry, with smaller values representing relatively higher degree of asymmetry and larger values representing relatively lower degree of asymmetry. See Wen and Levy (2001) for a more detailed discussion of the properties of the BLINEX loss.

3.2.1. Risk function of OLS estimator

The risk function of the OLS estimator, $\widehat{\boldsymbol{\beta}}_{OLS}$, under the BLINEX loss may be derived as

$$\begin{aligned}\Re_{BLINEX}(\widehat{\boldsymbol{\beta}}_{OLS};\boldsymbol{\beta}) &= E\left[\frac{1}{\lambda}\left(1-\frac{1}{1+b[\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}))-\mathbf{a}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-1]}\right)\right]\\ &= E\left[\frac{1}{\lambda}\left(1-\frac{1}{1-b[\mathbf{a}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}))+1]}\right)\right]\end{aligned}$$

and if we define

$$Z = b[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta}))+1]$$

then the risk function can be expressed as

$$\begin{aligned}\Re_{BLINEX}(\widehat{\boldsymbol{\beta}}_{OLS};\boldsymbol{\beta}) &= E\left[\frac{1}{\lambda}\left(1-\frac{1}{1-Z}\right)\right]\\ &= E\left[\frac{1}{\lambda}[1-(1+Z+Z^2+...)]\right] \text{ by applying Maclaurin series expansion}\\ &= -\frac{1}{\lambda}\left[E(Z)+E(Z^2)+...\right].\end{aligned} \tag{11}$$

We consider the expected values in equation (11) in turn. First if we take $E(Z)$ then

$$\begin{aligned}E(Z) &= E\left[b\left[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta}))+1\right]\right]\\ &= b\left[\mathbf{a}'E\left[(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})\right]-E\left[\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta}))\right]+1\right]\\ &= b\left[\mathbf{a}'(\boldsymbol{\beta}-\boldsymbol{\beta})-M_{\widehat{\boldsymbol{\beta}}_{OLS}}(\mathbf{a})\exp(-\mathbf{a}'\boldsymbol{\beta})+1\right]\\ &= b\left[-\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})+1\right] \text{ by using equation (6).}\end{aligned} \tag{12}$$

Next we take $E(Z^2)$, then

$$\begin{aligned}E(Z^2) &= E\left\{b(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]+1)'b(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]+1)\right\}\\ &= b^2\left\{E[(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})'\mathbf{a}\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]+2E[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]+E[\exp[2\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]]\right.\\ &\quad \left.-2E[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]]-2E[\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]]+1\right\}\\ &= b^2\left\{\mathbf{a}'\sigma_e^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}+2[\mathbf{a}'(\boldsymbol{\beta}-\boldsymbol{\beta})]+\exp(-2\mathbf{a}'\boldsymbol{\beta})M_{\widehat{\boldsymbol{\beta}}_{OLS}}(2\mathbf{a})\right.\\ &\quad \left.-2E[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})]]-2\exp(-\mathbf{a}'\boldsymbol{\beta})M_{\widehat{\boldsymbol{\beta}}_{OLS}}(\mathbf{a})+1\right\}\end{aligned}$$

where use is made of the fact that $E[(\widehat{\beta}_{OLS}-\boldsymbol{\beta})(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})'] = \sigma_e^2(\mathbf{X}'\mathbf{X})^{-1}$ and equation (6).

Having defined $\boldsymbol{\Sigma}_{OLS} = \sigma_e^2(\mathbf{X}'\mathbf{X})^{-1}$ then it follows that

$$\begin{aligned} E(Z^2) &= b^2\Big\{\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a} + \exp(-2\mathbf{a}'\boldsymbol{\beta})\exp(2\mathbf{a}'\boldsymbol{\beta} + 2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) \\ &\quad -2E[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})]] - 2\exp(-\mathbf{a}'\boldsymbol{\beta})\exp(\mathbf{a}'\boldsymbol{\beta} + \frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) + 1\Big\} \\ &= b^2\Big\{\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a} + \exp(2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) - 2\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) \\ &\quad -2E[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})\exp[\mathbf{a}'(\widehat{\beta}_{OLS} - \boldsymbol{\beta})]] + 1\Big\} \end{aligned} \tag{13}$$

and for the term $2E[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})]]$, in equation (13) we take

$$\begin{aligned} &((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - \boldsymbol{\Sigma}_{OLS}\mathbf{a})'\boldsymbol{\Sigma}_{OLS}^{-1}((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - \boldsymbol{\Sigma}_{OLS}\mathbf{a}) \\ &= ((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})' - (\boldsymbol{\Sigma}_{OLS}\mathbf{a})')\boldsymbol{\Sigma}_{OLS}^{-1}((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - \boldsymbol{\Sigma}_{OLS}\mathbf{a}) \\ &= (\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - (\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}\boldsymbol{\Sigma}_{OLS}\mathbf{a} \\ &\quad - (\boldsymbol{\Sigma}_{OLS}\mathbf{a})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) + (\boldsymbol{\Sigma}_{OLS}\mathbf{a})'\boldsymbol{\Sigma}_{OLS}^{-1}\boldsymbol{\Sigma}_{OLS}\mathbf{a} \\ &= (\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - (\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\mathbf{a} - \mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) + \mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a} \\ &= (\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - 2\widehat{\boldsymbol{\beta}}'_{OLS}\mathbf{a} + 2\boldsymbol{\beta}'a + \mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}. \end{aligned}$$

Therefore

$$\begin{aligned} (\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) &= ((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - \boldsymbol{\Sigma}_{OLS}\mathbf{a})'\boldsymbol{\Sigma}_{OLS}^{-1}((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - \boldsymbol{\Sigma}_{OLS}\mathbf{a}) \\ &\quad + 2\widehat{\boldsymbol{\beta}}'_{OLS}\mathbf{a} - 2\boldsymbol{\beta}'\mathbf{a} - \mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a} \end{aligned}$$

and

$$\begin{aligned} -\tfrac{1}{2}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) + \widehat{\boldsymbol{\beta}}'_{OLS}\mathbf{a} &= -\tfrac{1}{2}((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - \boldsymbol{\Sigma}_{OLS}\mathbf{a})'\boldsymbol{\Sigma}_{OLS}^{-1}((\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}) - \boldsymbol{\Sigma}_{OLS}\mathbf{a}) \\ &\quad + \boldsymbol{\beta}'\mathbf{a} + \tfrac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a} \end{aligned} \tag{14}$$

which then implies that

$$\begin{aligned} &E[2\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}))] \\ &\quad = E[2\mathbf{a}'\widehat{\boldsymbol{\beta}}_{OLS}\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})) - 2\mathbf{a}'\boldsymbol{\beta}\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}))] \\ &\quad = 2\mathbf{a}'\exp(-\mathbf{a}'\boldsymbol{\beta})E[\widehat{\boldsymbol{\beta}}_{OLS}\exp(\mathbf{a}'\widehat{\boldsymbol{\beta}}_{OLS})] - 2\mathbf{a}'\boldsymbol{\beta}\exp(-\mathbf{a}'\boldsymbol{\beta})E[\exp(\mathbf{a}'\widehat{\boldsymbol{\beta}}_{OLS})] \\ &\quad = 2\mathbf{a}'\exp(-\mathbf{a}'\boldsymbol{\beta})\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\widehat{\boldsymbol{\beta}}_{OLS}\exp(\mathbf{a}'\widehat{\boldsymbol{\beta}}_{OLS})f_{\widehat{\boldsymbol{\beta}}_{OLS}}(\widehat{\boldsymbol{\beta}}_{OLS})d\widehat{\boldsymbol{\beta}}_{OLS} - 2\mathbf{a}'\boldsymbol{\beta}\exp(-\mathbf{a}'\boldsymbol{\beta})M_{\widehat{\boldsymbol{\beta}}_{OLS}}(\mathbf{a}) \\ &\quad = 2\mathbf{a}'\exp(-\mathbf{a}'\boldsymbol{\beta})\Big\{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\widehat{\boldsymbol{\beta}}_{OLS}\exp(\mathbf{a}'\widehat{\beta}_{OLS})(2\pi)^{\frac{-k}{2}}|\boldsymbol{\Sigma}_{OLS}|^{\frac{-1}{2}} \\ &\qquad \exp[-\frac{1}{2}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta})]d\widehat{\boldsymbol{\beta}}_{OLS}\Big\} - 2\mathbf{a}'\boldsymbol{\beta}\exp(-\mathbf{a}'\boldsymbol{\beta})\exp(\boldsymbol{\beta}'\mathbf{a} + \frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) \end{aligned}$$

$$= 2\mathbf{a}'\exp(-\mathbf{a}'\boldsymbol{\beta})\left\{\int\limits_{-\infty}^{\infty}\dots\int\limits_{-\infty}^{\infty}\widehat{\boldsymbol{\beta}}_{OLS}(2\pi)^{\frac{-k}{2}}|\boldsymbol{\Sigma}_{OLS}|^{\frac{-1}{2}}\right.$$

$$\left.\exp[-\frac{1}{2}(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})+\mathbf{a}'\widehat{\boldsymbol{\beta}}_{OLS}]d\widehat{\boldsymbol{\beta}}_{OLS}\right\} - 2\mathbf{a}'\boldsymbol{\beta}\exp(\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}-\mathbf{a}'\boldsymbol{\beta})$$

$$= 2\mathbf{a}'\exp(-\mathbf{a}'\boldsymbol{\beta})$$

$$\left\{\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty}\widehat{\boldsymbol{\beta}}_{OLS}(2\pi)^{\frac{-k}{2}}|\boldsymbol{\Sigma}_{OLS}|^{\frac{-1}{2}}\exp\Big[-\frac{1}{2}(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta}-\boldsymbol{\Sigma}_{OLS}\mathbf{a})'\boldsymbol{\Sigma}_{OLS}^{-1}(\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta}-\boldsymbol{\Sigma}_{OLS}\mathbf{a})+\right.$$

$$\left.\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}\Big]d\widehat{\boldsymbol{\beta}}_{OLS}\right\} - 2\mathbf{a}'\boldsymbol{\beta}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}), \text{ from equation (14)}$$

$$= 2\mathbf{a}'\exp(-\mathbf{a}'\boldsymbol{\beta})\left\{\int\limits_{-\infty}^{\infty}\dots\int\limits_{-\infty}^{\infty}\widehat{\boldsymbol{\beta}}_{OLS}(2\pi)^{\frac{-k}{2}}|\boldsymbol{\Sigma}_{OLS}|^{\frac{-1}{2}}\right.$$

$$\left.\exp\Big[-\frac{1}{2}((\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-\boldsymbol{\Sigma}_{OLS}\mathbf{a})'\boldsymbol{\Sigma}_{OLS}^{-1}((\widehat{\boldsymbol{\beta}}_{OLS}-\boldsymbol{\beta})-\boldsymbol{\Sigma}_{OLS}\mathbf{a})\Big]d\widehat{\boldsymbol{\beta}}_{OLS}\right\}$$

$$\exp(\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) - 2\mathbf{a}'\boldsymbol{\beta}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})$$

$$= 2\mathbf{a}'\exp(-\mathbf{a}'\boldsymbol{\beta})\left(\exp(\boldsymbol{\beta}'\mathbf{a}+\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})E(\widehat{\boldsymbol{\beta}}_{OLS})\right) - 2\mathbf{a}'\boldsymbol{\beta}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})$$

$$= 2\mathbf{a}'\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})(\boldsymbol{\beta}+\boldsymbol{\Sigma}_{OLS}\mathbf{a}) - 2\mathbf{a}'\boldsymbol{\beta}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})$$

$$= 2\mathbf{a}'\boldsymbol{\beta}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) + 2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) - 2\mathbf{a}'\boldsymbol{\beta}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})$$

$$= 2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})$$

where use is made of the fact that $E(\widehat{\boldsymbol{\beta}}_{OLS}) = \int\limits_{-\infty}^{\infty}\dots\int\limits_{-\infty}^{\infty}\widehat{\boldsymbol{\beta}}_{OLS}f_{\widehat{\beta}_{OLS}}(\widehat{\boldsymbol{\beta}}_{OLS})\ \ d\widehat{\boldsymbol{\beta}}_{OLS}$.

Therefore, it follows from above and equation (13) that

$$E(Z^2) = b^2\Big\{\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}+\exp(2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})-2\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) -2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})+1\Big\}. \quad (15)$$

Hence, from equations (12) and (15), the approximated risk function of the OLS estimator under the BLINEX risk function is

$$\begin{aligned}
&\Re_{BLINEX}(\widehat{\boldsymbol{\beta}}_{OLS};\boldsymbol{\beta})\\
&\quad= -\frac{1}{\lambda}[E(Z)+E(Z^2)+O(3)]\\
&\quad\propto -\frac{1}{\lambda}[b(-\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})+1)+b^2(\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}+\exp(2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})-2\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})\\
&\qquad\quad -2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})+1)]. \qquad (16)
\end{aligned}$$

3.2.2. Risk functions of GLS and FGLS estimators

It follows similarly to the approach used in deriving the OLS risk, that the approximated GLS risk under the BLINEX loss is given by

$$\begin{aligned}
&\Re_{BLINEX}(\widetilde{\boldsymbol{\beta}}_{GLS};\boldsymbol{\beta})\\
&\quad= -\frac{1}{\lambda}[E(Z)+E(Z^2)+O(3)]\\
&\quad\propto -\frac{1}{\lambda}[b(-\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a})+1)+b^2(\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a}+\exp(2\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a})-2\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a})\\
&\qquad\quad -2\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{GLS}\mathbf{a})+1)] \qquad (17)
\end{aligned}$$

and from equation (17) the approximated risk expression for the FGLS estimator under BLINEX is given by

$$\begin{aligned}
&\Re_{BLINEX}(\widehat{\boldsymbol{\beta}}_{FGLS};\boldsymbol{\beta})\\
&\quad\propto -\frac{1}{\lambda}[b(-\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})+1)+b^2(\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a}+\exp(2\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})-2\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})\\
&\qquad\quad -2\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})+1)]. \qquad (18)
\end{aligned}$$

3.2.3. Risk function of PTE

It then follows from equations (16), (18) and (9) that the approximated pretest estimator risk function under the BLINEX for a specified level of significance, η, is given by

$$\begin{aligned}
&\Re_{BLINEX}(\widehat{\boldsymbol{\beta}}_{PTE};\boldsymbol{\beta})\\
&\quad= E\left[\frac{1}{\lambda}\left(1-\frac{1}{1+b[\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta}))-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta})-1]}\right)\right]\\
&\quad= E\left[\frac{1}{\lambda}\left(1-\frac{1}{1+b[\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta}))-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta})-1]}\right)|H_0\right]P(d>d_\eta)\\
&\qquad+E\left[\frac{1}{\lambda}\left(1-\frac{1}{1+b[\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta}))-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_{PTE}-\boldsymbol{\beta})-1]}\right)|H_1\right]P(d\le d_\eta)
\end{aligned}$$

$$
\begin{aligned}
&= \Re_{BLINEX}(\widehat{\boldsymbol{\beta}}_{OLS};\boldsymbol{\beta})(1-\eta)+\Re_{BLINEX}(\widehat{\boldsymbol{\beta}}_{FGLS};\boldsymbol{\beta})\eta \\
&\propto \Big\{-\frac{1}{\lambda}[b(-\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})+1)+b^2(\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}+\exp(2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})-2\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}) \\
&\quad -2\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{OLS}\mathbf{a})+1)]\Big\}(1-\eta)+ \\
&\quad \Big\{-\frac{1}{\lambda}[b(-\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})+1)+b^2(\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a}+\exp(2\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})-2\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a}) \\
&\quad -2\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a}\exp(\frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}_{FGLS}\mathbf{a})+1)]\Big\}\eta.
\end{aligned}
$$

We now turn to the numerical analysis of the risk functions.

4. The Monte Carlo simulation

The data used for the Monte Carlo experiment were generated by the single explanatory variable model

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

where $X_t = \exp(0.04t) + w_t$, $w_t \sim N(0, 0.0009)$, $u_t = \rho u_{t-1} + e_t$, $e_t \sim N(0, 0.0036)$, $t = 1, 2, ..., T$ and $\beta = (\beta_1, \beta_2) = (1, 1)$. This specification of the $\mathbf{X}'s$ with a trend component was used in a study by Beach and Mackinnon (1978). The results are presented for two sample sizes, $T = 20$ and $T = 50$, and we consider ten values of ρ varying by tenths from 0.0 to 0.9. $N = 1000$ replications of the experiment were used.

The estimators compared are: the ordinary least squares (OLS), the Durbin estimator (DE), and Maximum Likelihood estimator (MLE), a pretest estimator choosing between OLS and DE, (DEPTE) and a pretest estimator choosing between OLS and MLE, (MLEPTE). Following Fomby and Guilkey (1978) we only consider $\alpha = 0.5$ for the pretest, since they found this to be the optimal level of significance. The exact critical values for the Durbin-Watson test are obtained by making use of simulations.

The risks for any of the estimator $\widehat{\boldsymbol{\beta}}$ relative to the LINEX and BLINEX loss, respectively, are calculated by

$$
\begin{aligned}
\Re_{LINEX}(\widehat{\boldsymbol{\beta}};\boldsymbol{\beta}) &= \frac{1}{N}\sum_{j=1}^{N} c\left(\exp[\mathbf{a}'(\widehat{\boldsymbol{\beta}}_j-\boldsymbol{\beta})]-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_j-\boldsymbol{\beta})-1\right) \\
\Re_{BLINEX}(\widehat{\boldsymbol{\beta}};\boldsymbol{\beta}) &= \frac{1}{N}\sum_{j=1}^{N} \frac{1}{\lambda}\left[1-\frac{1}{1+b[\exp(\mathbf{a}'(\widehat{\boldsymbol{\beta}}_j-\boldsymbol{\beta}))-\mathbf{a}'(\widehat{\boldsymbol{\beta}}_j-\boldsymbol{\beta})-1]}\right]
\end{aligned}
$$

and normalised by dividing by the GLS risks $\Re(\widetilde{\boldsymbol{\beta}}_{GLS};\boldsymbol{\beta})$ to obtain the relative risks. That is, the estimator risks are evaluated relative to the GLS risk, which is best linear unbiased estimator under serial correlation and therefore makes a good reference for other estimators performance as we expect it to strictly dominate all estimators over the entire range of ρ. Thus, for any estimator, the closer the relative risk to unity, the better the performance of that estimator relative to GLS estimator. Comparisons of the risks in terms of relative risks also makes it easier for comparability across the different loss functions. The loss parameter combinations used in the simulation are

$\mathbf{a} = (-0.5, -0.5)$ $c = 1, \mathbf{a} = (-3, -3)$ $c = 1$ and $\mathbf{a} = (-8, -8)$ $c = 1$ for the LINEX loss and $\mathbf{a} = (-0.5, -0.5)$, $b = 2$, $\lambda = 0.2$ and $\mathbf{a} = (-3, -3)$, $b = 0.0002, \lambda = 0.2$ for the BLINEX loss. All computations were done in the SAS 9.2 package.

5. Numerical evaluations and discussion

The results of our analysis are summarised in Tables 1 and 2 and Figures 1 and 2. The results suggest that some of the key results under the symmetric SE loss as reported in the literature continue to hold as the degree of loss asymmetry increases. In particular, the risks of all estimators increases with ρ and the relative ordering of the estimators performances remains the same as the degree of loss asymmetry varies. Generally, the DEPTE is preferable for small sample size *T=20*, while the MLEPTE performs the best for *T=50*. Furthermore, the risks of the FGLS estimators and their corresponding pretest estimators decreases and converge as the sample size increases, a result which is consistent with findings by Judge et al. (1985) on the asymptotic properties of the FGLS estimators. We also find that the risks of the pretest estimators are quite robust to loss asymmetry for small to medium values of ρ. For large values of ρ, however, the risks are generally decreasing as the degree of loss asymmetry increases. This is clearly illustrated in Figure 1 for the DEPTE risks under the LINEX loss and Figure 2 for the MLEPTE risks under the BLINEX loss. Hence, we note that for larger values of ρ, the risk gains of the pretest estimators over the OLS increases with higher loss asymmetry. That is, when there is a sufficient degree of asymmetry, the benefits of preliminary test estimation under asymmetry loss are even more.

Table 1: Relative risk function values under the LINEX loss.

Loss parameters	Estimator	$T = 20$				$T = 50$			
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$	$\rho = 0$	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$
a=(-0.5,-0.5) *c*=1	OLS	1.000	1.035	1.118	1.221	1.000	1.011	1.045	1.153
	DE	1.011	1.019	1.036	1.091	1.001	1.004	1.008	1.032
	MLE	1.014	1.024	1.043	1.088	1.000	1.003	1.006	1.023
	DEPTE	1.000	1.015	1.036	1.091	1.001	1.004	1.008	1.032
	MLEPTE	1.001	1.021	1.043	1.088	1.000	1.003	1.006	1.023
a=(-3,-3) *c*=1	OLS	1.000	1.034	1.118	1.220	1.000	1.011	1.044	1.148
	DE	1.010	1.018	1.034	1.086	1.001	1.004	1.008	1.028
	MLE	1.013	1.023	1.041	1.085	1.001	1.003	1.006	1.020
	DEPTE	1.000	1.014	1.035	1.086	1.001	1.004	1.008	1.028
	MLEPTE	1.001	1.020	1.041	1.084	1.000	1.003	1.006	1.020
a=(-8,-8) *c*=1	OLS	1.000	1.035	1.122	1.239	1.000	1.011	1.043	1.148
	DE	1.009	1.017	1.032	1.077	1.002	1.005	1.009	1.021
	MLE	1.013	1.022	1.038	1.082	1.001	1.004	1.007	1.013
	DEPTE	1.000	1.013	1.033	1.077	1.001	1.005	1.009	1.021
	MLEPTE	1.001	1.019	1.037	1.082	1.000	1.003	1.007	1.013

Table 2: Relative risk function values under the BLINEX loss.

Loss parameters	Estimator	$T = 20$				$T = 50$			
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$	$\rho = 0$	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$
a=(-0.5,-0.5)	OLS	1.000	1.034	1.118	1.219	1.000	1.011	1.045	1.152
b=2	DE	1.011	1.019	1.036	1.090	1.001	1.004	1.008	1.032
λ=0.2	MLE	1.014	1.024	1.043	1.088	1.000	1.003	1.006	1.023
	DEPTE	1.000	1.015	1.036	1.090	1.001	1.004	1.008	1.032
	MLEPTE	1.001	1.021	1.042	1.088	1.000	1.003	1.006	1.023
a=(-3,-3)	OLS	1.000	1.034	1.118	1.220	1.000	1.011	1.044	1.148
b=0.0002	DE	1.010	1.018	1.034	1.086	1.001	1.004	1.008	1.028
λ=0.2	MLE	1.013	1.023	1.041	1.085	1.001	1.003	1.006	1.020
	DEPTE	1.000	1.014	1.035	1.086	1.001	1.004	1.008	1.028
	MLEPTE	1.001	1.020	1.041	1.084	1.000	1.003	1.006	1.020

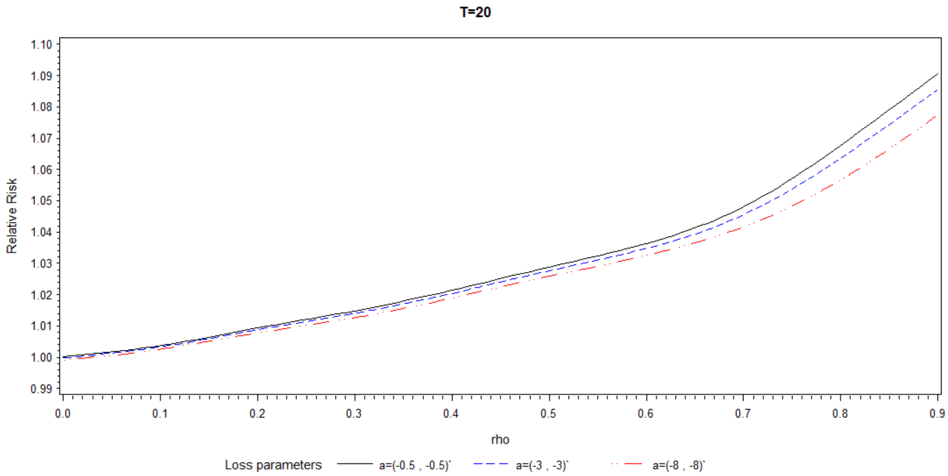

Figure 1: Relative risk functions for the Durbin pretest estimator under LINEX loss.

6. Numerical example

As an illustration, we consider the South African annual data on aggregate household consumption expenditure and aggregate household disposable income for the period 1946-1990, obtained from the South African Reserve Bank website. The aggregate household consumption function, explained in most macroeconomic introductory textbooks is estimated empirically to determine the household consumption behavior for the South African economy over the period 1946-1990, emphasising the use of the serial correlation pretest estimator and consideration for asymmetry loss in empirical work.

Keynes (1936) believed that real consumption expenditure is highly dependent on real disposable

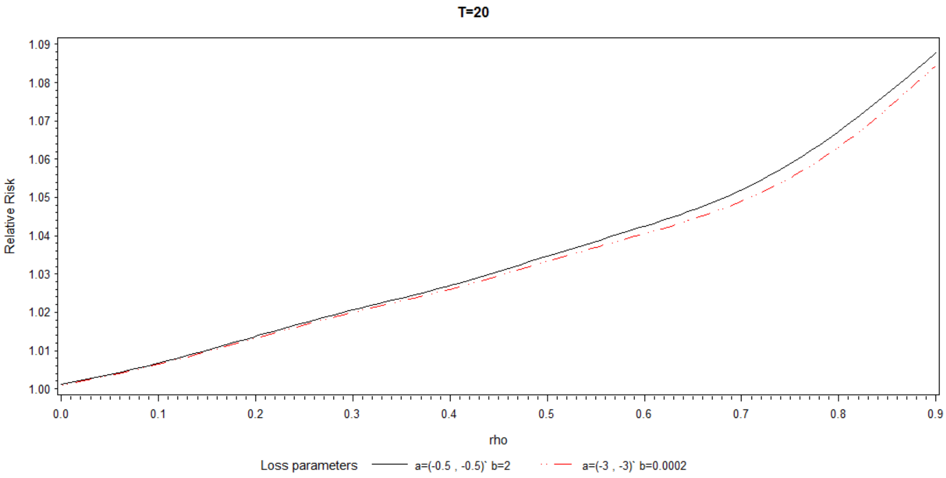

Figure 2: Relative risk functions for the Maximum likelihood pretest estimator under BLINEX loss.

income and his theory has resulted in the well known Keynesian consumption function

$$C_t = \alpha + \theta Y_t + \omega_t \tag{19}$$

where C_t is the real aggregate household consumption expenditure, Y_t is the real aggregate household disposable income and ω_t is a vector of error terms. In economics theory, α is the autonomous consumption or the minimum level of consumption and θ is the marginal propensity to consume, which indicates the rate of change in real household consumption for a unit change in the real household disposable income. According to Keynes theory of consumption, the MPC is positive but less than unity, suggesting that an increase in real household disposable income will lead to an increase in real household consumption, but this increase will be less than the increase in real household income because a proportion of real disposable income is saved.

Since some more complex specification of equation (19), with more explanatory variables, could be more realistic, we make a simplistic assumption that if there is serial correlation in the error terms it is pure serial correlation, that is serial correlation is due to the natural sequential ordering of the data and not due to the possible misspecification of the model.

Of interest here is the estimation of θ, the marginal propensity to consume, which generally plays a role in economic stabilisation through some policy multipliers. For a discussion on this, see, for example, van Zyl (1970). Simply put, for example, when a very large proportion of disposable income is consumed, then θ is close to unity, and the aggregate saving rate is very low. In such a scenario, in terms of the policy implications, on the fiscal side tax incentives can be used to curb consumer demand and encourage savings, i.e. contributions to unit trusts and on the monetary side, interest rates can be used to discourage spending by increasing the interest rates. Conversely, when households consume too little of their disposable income compared to saving, θ is very small, then fiscal policy can be used to stimulate demand by decreasing tax rates while the monetary policy tool would be to lower the interest rates. Accordingly, overestimation or underestimation of the marginal propensity to consume, could lead to wrong policy decision. Thus, when estimating θ, depending on which one has more severe consequences between overestimation and underestimation, it could be more appropriate to find an estimator that is relatively more efficient relative to a performance measure that allows for asymmetry. The direction of penalisation will then generally depend on the economic conditions.

6.1. Empirical estimation

For estimation we use the log-linear model for equation (19). Using the OLS residuals, the Durbin-Watson statistic $d = 0.511$ and the 50% exact critical value for the Durbin-Watson test $d_{0.5} = 2.05$. Therefore, since the Durbin-Watson statistic, is less than the critical value, the null hypothesis H_0 : $\rho = 0$ is rejected, suggesting that the level of serial correlation in the error terms is significant. First, we determine whether the first order autoregressive error process is the correct error generating process. This is done by making use of the tentative order selection tests based on the smallest canonical correlation (*SCAN*) and the extended sample autocorrelation function (*ESACF*) methods as presented in Table 3 where p is the order of the autoregressive, d^* is the order of integration, q is the moving average order and noting that the *ARMA* orders are identified by choosing the minimum Bayesian information criterion (*BIC*), see details in van Staden (2012). Table 3 shows that the OLS residuals are compatible with the first order autoregressive process.

Table 3: $ARMA(p+d^*,q)$ Tentative Order Selection Tests.

SCAN			*ESACF*		
$p+d^*$	q	*BIC*	$p+d^*$	q	*BIC*
1	0	−7.26506	1	0	−7.26506
0	2	−6.82147	0	2	−6.82147

Since the level of serial correlation is significant, the preliminary test estimation strategy will choose the FGLS estimator that corrects for serial correlation. It is however important to note that the sampling properties of the pretest estimator will differ from those of the FGLS estimator. Since the sampling distribution of the serial correlation pretest estimator has not been derived analytically, we make use of the non-overlapping block bootstrap procedure with block length *l=3* to estimate and compare the standard errors and confidence intervals of the Durbin pretest estimator to OLS and Durbin estimators. See Lahiri (2003) for a discussion on the block bootstrap technique.

The bootstrap estimates, standard errors and confidence interval lengths are presented in Table 4. The OLS estimate of the slope coefficient has the least standard error and a narrower confidence interval. This is to be expected because as we pointed out earlier, under serial correlation the standard errors of the OLS coefficients are underestimated and consequently the confidence intervals are inaccurate. Note that under serial correlation the sampling properties of the Durbin pretest estimator are closer to those of its FGLS component, Durbin estimator.

Table 4: Bootstrap estimates.

	$\widehat{\beta}_{OLS}$	$\widehat{\beta}_{DE}$	$\widehat{\beta}_{DEPTE\alpha=0.5}$
Slope coefficient	0.9851383	0.9840124	0.9840136
Standard error	0.0247173	0.0255237	0.0255249
95% confidence interval length	0.0968621	0.101243	0.101243

7. Conclusion

This paper extends the choice of the loss function used in comparing the estimators in models with first order autoregressive error terms to asymmetry loss functions. The results of our analysis reaffirm that there are considerable risk gains by using the serial correlation pretest estimators. For smaller to medium values of ρ, the risk performances of the pretest estimators are quite robust to loss asymmetry. For larger values of ρ, on the other hand, the relative risks of the pretest estimators decrease as the loss asymmetry increase and their risk gains over the OLS increases with higher loss asymmetry.

Acknowledgements

This work is based on the research supported by the National Research Foundation of South Africa for the grant TTK1206151317. Any opinion, finding and conclusion or recomendation expressed in this material is that of the authors and the NRF does not accept any liability in this regard. The authors would like to thank the anonymous referees for their constructive comments that improved the presentation of the paper.

References

ARASHI, M. (2009). Preliminary test estimation of the mean vector under balanced loss function. *Journal of Statistical Research*, **2** (43), 55–65.

ARASHI, M. (2012). Preliminary test and Stein estimators in simultaneous linear equations. *Linear Algebra and its Applications*, **5** (436), 1195–1211.

ARASHI, M., TABATABAEY, S. M. M., AND HASSANZADEH, M. (2009). Preliminary test methodology in ridge regression under stochastic constraints. *In 7th Seminar on Probability Stochastic Processes, Isfahan University of Technology*.

BEACH, C. M. AND MACKINNON, J. G. (1978). A maximum likelihood procedure for regression with autocorrelated errors. *Econometrica*, **46** (1), 51–58.

COCHRANE, D. AND ORCUTT, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, **44** (245), 32–61.

DURBIN, J. (1960). Estimation of parameters in time-series regression models. *Journal of the Royal Statistical Society*, **22** (1), 139–153.

DURBIN, J. AND WATSON, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika*, **37** (3/4), 409–428.

DURBIN, J. AND WATSON, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, **38** (1/2), 159–178.

DURBIN, J. AND WATSON, G. S. (1971). Testing for serial correlation in least squares regression. III. *Biometrika*, **58** (1), 1–19.

FOLMER, H. (1988). Autocorrelation pre-testing in linear models with AR(1) errors. *In* DIJKSTRA, T. K. (Editor) *On Model Uncertainty and its Statistical Implications*. Springer-Verlag, Berlin, pp. 39–55.

FOMBY, T. B. AND GUILKEY, D. K. (1978). On choosing the optimal level of significance for the Durbin-Watson test and the Bayesian alternative. *Journal of Econometrics*, **8**, 203–213.

GUJARATI, D. N. (2003). *Basic Econometrics*. McGraw-Hill, New York.

JUDGE, G. G. AND BOCK, M. E. (1978). *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics*. North-Holland, Amsterdam.

JUDGE, G. G., GRIFFITHS, W. E., HILL, R. C., LUTKEPOHL, H., AND LEE, T. (1985). *The Theory and Practice of Econometrics*. Second edition. John Wiley and Sons, New York.

KEYNES, J. M. (1936). *The General Theory of Employment Interest and Money*. Macmillan and Co, London.

KING, M. L. AND GILES, D. E. A. (1984). Autocorrelation pre-testing in the linear model estimation, testing and prediction. *Journal of Econometrics*, **25**, 35–48.

LAHIRI, S. N. (2003). *Resampling Methods for Dependent Data*. Springer Series in Statistics, Springer, Berlin.

MOREY, M. J. (1975). *The statistical consequences of autocorrelation preliminary test estimators under squared-error loss*. Masters dissertation, University of Illinois, Illinois.

PRAIS, S. J. AND WINSTEN, C. B. (1954). Trend estimators and serial correlation. Cowles Commission discussion paper 383, Chicago.

RAO, P. AND GRILICHES, Z. (1969). Small-sample properties of several two-stage regression methods in the context of autocorrelated errors. *Journal of the American Statistical Association*, **64** (325), 253–272.

SALEH, A. K. M. E. (2006). *Theory of Preliminary Test and Stein-type Estimation with Applications*. John Wiley, New York.

SALEH, A. K. M. E. AND KIBRIA, B. M. G. (1993). Performances of some new preliminary test ridge regression estimators and their properties. *Communications in Statistics-Theory and Methods*, **22**, 2747–2764.

SALEH, A. K. M. E. AND KIBRIA, B. M. G. (2011). On some ridge regression estimators: A nonparametric approach. *Journal of Nonparametric Statistics*, **23** (3), 819–851.

SIMONS, D. N. T. (1988). *The Sampling Distribution Properties of the Bayesian Estimator in the Case of Autocorrelated Error: A Monte Carlo Study*. Doctoral thesis, Simon Fraser University, Burnaby.

VAN STADEN, P. (2012). *Time Series Analysis 321 WST 321*. University of Pretoria, Pretoria.

VAN ZYL, J. C. (1970). *An Econometric Analysis of the Consumption Function in South Africa*. Doctoral thesis, University of Pretoria, Pretoria.

VARIAN, H. R. (1975). A Bayesian approach to real estate assessment. *In* FIENBERG, S. E. AND ZELLNER, A. (Editors) *Studies in Bayesian Econometrics and Statistics In Honor of Leonard J. Savage*. North-Holland, Amsterdam, pp. 195–208.

WEN, D. AND LEVY, M. S. (2001). BLINEX: A bounded asymmetric loss function with application to Bayesian estimation. *Communications in Statistics-Theory and Methods*, **30** (1), 147–153.

ZELLNER, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, **81** (394), 446–451.

Manuscript received, 2014-05-20, revised, 2014-12-29, accepted, 2015-01-22.