The risk properties of a pre-test estimator for Zellner's seemingly unrelated regression model
Özcam, A.; Judge, G.; Bera, A.K.; Yancey, T.

Document version:
Publisher's PDF, also known as Version of record

Publication date:
1994

Link to publication

Citation for published version (APA):
The Risk Properties of a Pre-test Estimator for Zellner's Seemingly Unrelated Regression Model

by Ahmet Özcam, George Judge, Anil Bera and Thomas Yancey

CENTER FOR ECONOMIC RESEARCH

Board
Harry Barkema
Helmut Bester
Eric van Damme, Chairman
Frank van der Duyn Schouten
Jeffrey James

Management
Jeffrey James (Director of Graduate Studies)
Arie Kapteyn (Scientific Director)
Marie-Louise Kemperman (Managing Director)

Scientific Council
Anton Barten
Eduard Bomhoff
Willem Buiter
Jacques Drèze
Jack Kleijn
Theo van de Klundert
Jean-Jacques Laffont
Merton Miller
Piet Moerland
Philippe Naert
Pieter Ruys

Université Catholique de Louvain
Erasmus University Rotterdam
Yale University
Université Catholique de Louvain
Tilburg University
Tilburg University
Université des Sciences Sociales de Toulouse
University of Chicago
Tilburg University
Nijenrode University
Tilburg University

Residential Fellows
Hans Bloemen
Lans Bovenberg
Hans Carlsson
Jay Pil Choi
Jan Magnus
Andrew Mountford
Bezalel Peleg
Mark Steel
Frank Verboven
Oscar Volij
Karl-Erik Wärneryd

CentER
CentER, Erasmus University Rotterdam
Gothenburg University and Lund University
Columbia University
CentER, LSE
CentER
Hebrew University of Jerusalem
CentER/Department of Econometrics, Tilburg University
CentER
Hebrew University of Jerusalem
Stockholm School of Economics

Address : P.O. Box 90153, 5000 LE Tilburg, The Netherlands
Phone : +31 13 663102
Telex : 52426 kub nl
Telefax : +31 13 663066
E-mail : center@kub.nl

ISSN 0924-7874

1994
The Risk Properties of a Pre-test Estimator for Zellner’s Seemingly Unrelated Regression Model

by Ahmet Özcaml, George Judge, Anil Bera and Thomas Yancey


Reprint Series no. 143
THE RISK PROPERTIES OF A PRE-TEST ESTIMATOR FOR ZELLNER'S SEEMINGLY UNRELATED REGRESSION MODEL

AHMET ÖZCAM, GEORGE JUDGE, ANIL BERA, and THOMAS YANCEY

Research Division of Turkish Parliament; University of California, Berkeley; University of Illinois at Urbana-Champaign and Tilburg University; and University of Illinois at Urbana-Champaign

In the case of Zellner's seemingly unrelated statistical model it is well known that the efficiency of the generalized least squares estimator (GLSE) relative to that of the least squares estimator (LSE) is conditional on the magnitude of the correlation between the equation errors. Using a relevant test statistic, we analytically evaluate the risk characteristics of a seemingly unrelated regressions pre-test estimator (SURPE) that is the GLSE if a preliminary test, based on the data at hand, indicates that the correlation between equation errors is significantly different from zero, and the LSE if we accept the null hypothesis of no correlation. The small sample distribution of the test statistic, used in defining SURPE is also derived. λ(JEL C39)

1. INTRODUCTION

Since Zellner (1962) proposed the use of Aitken's generalized least squares estimator (GLSE) for a set of disturbance related regression equations, the efficiency of this estimator relative to that of the least squares estimator (LSE) has received much attention. For the uncorrelated regressors case, Zellner (1963) derived the small sample properties of the seemingly unrelated regression estimator (SURE) and noted that the distribution of the estimator converges rapidly toward a normal density. Mehta and Swamy (1976) derived the

*We are grateful to an anonymous referee for detailed comments and many helpful suggestions. We also wish to express our appreciation to David Giles and Helga Hessenius. This work was partially supported by National Science Foundation grant, SES-86-96152.
exact second moment matrix of Zellner's estimator conditional on an estimate of the variance-covariance matrix of the error terms and found that (i) the LSE is more efficient than Zellner's estimator if the correlation in the errors of the two equations is zero, or small and (ii) Zellner's estimator is better if the contemporaneous correlation is high (also see Kunitomo (1977)). They also indicate that the gain in efficiency in using Zellner's estimator is especially high when the equation error correlation coefficient is close to one, and the loss is small when the errors are mildly correlated and the degrees of freedom is greater than 12.

In this paper, we examine under a squared error loss measure the risk of the seemingly unrelated regression pre-test estimator (SURPE), which is the GLSE if a preliminary test indicates that the correlation coefficient is significantly different from zero, and the LSE if we accept the null hypothesis of no correlation. The motivation for this research comes from Zellner's suggestion that it is possible to develop a decision procedure for deciding whether to use the LSE, or the GLSE.

In section 2, we present the statistical model and the various estimators. Our main interest is to derive the risk function of the SURPE with respect to the joint distribution of the test statistic \( r = s_{12}/\sqrt{s_{11}s_{22}} \) and \( v = s_{12}/s_{22} \), where the \( s_{ij} \) \((i, j = 1, 2)\), which are defined later, are consistent estimators of the variances and the covariances of the errors. The small sample distribution of \( r \) as a function of the population correlation coefficient \( \phi \) is given in section 3. The marginal distribution of \( r \) is obtained from the joint distribution of \( r \) and \( v \). In section 4, we derive the risk function of the SURPE and compare it with those of LSE and GLSE. Section 5 summarizes the discussion of the implications of the paper.

### 2. STATISTICAL MODEL AND ESTIMATORS

Consider the following two sample regression model

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\end{bmatrix} = \begin{bmatrix} X_1 & 0 \\
0 & X_2 \\
\end{bmatrix} \begin{bmatrix} a_1 \\
a_2 \\
\end{bmatrix} + \begin{bmatrix} e_1 \\
e_2 \\
\end{bmatrix}, \quad \text{or } y = Xa + e
\]  

where \( y_i \) is a \((n \times 1)\) vector of observations, \( X_i \) is a \((n \times p)\) matrix of fixed regressors of rank \( p \), \( a_i \) is a \((p \times 1)\) unknown location vector, and \( e_i \) is an \((n \times 1)\) random error vector for \( i = 1, 2 \). We make a simplifying assumption that \( X_1'X_2 = X_2'X_1 = O_p \). Let us further assume that the equation errors are distributed as multivariate normal random variables with zero means and covariance matrix

\[
\Sigma = E \begin{bmatrix} e_1' \\
e_2' \\
\end{bmatrix} [e_1'e_2'] = E[ee'] = \begin{bmatrix} \sigma_{11} I_n & \sigma_{12} I_n \\
\sigma_{12} I_n & \sigma_{22} I_n \\
\end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\
\sigma_{21} & \sigma_{22} \\
\end{bmatrix} \otimes I_n
\]  

where \( I_n \) is an identity matrix of dimension \( n \). The LSE for this model is

\[
a^*(1) = \begin{bmatrix} (X_1'X_2)^{-1} & X_1' \\
(X_2'X_2)^{-1} & X_2' \\
\end{bmatrix} \begin{bmatrix} y_1 \\
y_2 \\
\end{bmatrix}
\]  

The Zellner SUR estimator

\[
a^*(2) = (X'\Sigma^{-1}X)^{-1} X'\Sigma^{-1}y
\]
is obtained by applying Aitken's GLSE to the whole system (2.1). The estimator in (2.4) is not feasible since it depends on unknown parameters of the Σ matrix. Replacing Σ by a consistent estimator $S$ produces Zellner's feasible GLSE, $a^*(4)$. One choice for the elements of $S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}$ is $s_{ij} = \frac{1}{n} (y_i - x_i a_i^*(1))' (y_j - X_j a_j^*(1)), i, j = 1, 2$. Now the feasible GLSE is given by

$$a^*(4) = \begin{bmatrix} X_1' & 0 \\ 0 & X_2' \end{bmatrix} \begin{bmatrix} s_{11} I_n & s_{12} I_n \\ s_{21} I_n & s_{22} I_n \end{bmatrix}^{-1} \begin{bmatrix} X_1' \\ 0 \end{bmatrix} \begin{bmatrix} s_{11} I_n & s_{12} I_n \\ s_{21} I_n & s_{22} I_n \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} (X_1'X_1)^{-1} X_1' y_1 + (s_{11}^2/s_{11}) (X_1'X_1)^{-1} X_1' y_2 \\ (X_2'X_2)^{-1} X_2' y_2 + (s_{22}^2/s_{22}) (X_2'X_2)^{-1} X_2' y_1 \end{bmatrix}$$

(2.5)

where we have used the assumption $X_1' X_2 = X_2' X_1 = O_p$ and the $s_{ij}$ are the elements of $S^{-1} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \otimes I_n$. The estimates of the variances and the covariances are obtained from the restricted residuals, that are obtained from regressing $y_i$ on $X_i (i = 1, 2)$, i.e., implicitly assuming $\phi = 0$.

The SUR pre-test estimator (SURPE) is based on the test statistic $r = s_{12}/\sqrt{s_{11} s_{22}}$ that is used to test the null hypothesis $H_0: \phi = 0$ that the population correlation coefficient $\phi$ is zero, versus a one-sided alternative $H_1: \phi > 0$. We reject the null hypothesis if $r > c$, where $c$ is the critical value chosen for the test. If we suspect a negative correlation then we reject the $H_0$, if $r < -c$. A two-sided alternative can also be set up and this would of course have implications for the properties of the implied pretest estimator. This test statistic is similar to the locally best invariant test statistic given by Kariya (1981) and the Lagrange multiplier statistic of Breusch and Pagan (1980) and Shiba and Tsurumi (1988). The pretest estimator (Judge and Bock (1978)) is defined as follows: if we accept $H_0$, the SURPE is the LSE, and otherwise it is the GLSE. This means the SURPE is

$$a^*(3) = I_{[-1,c]}(r) a^*(1) + I_{[c,1]}(r) a^*(4)$$

(2.6)

where $I_{(-)} (\cdot)$ is a zero-one indicator function.

### 3. THE SMALL SAMPLE DISTRIBUTION OF R

The distribution of SURPE $a^*(3)$ and hence its risk depends on the distribution of $r$. Therefore, in this section we derive the small sample distribution of $r$. First, we find the joint distribution of the test statistic $r$ and $v$. It is well known that $n s_{11} = x$, $n s_{22} = y$ and $n s_{12} = z$ are distributed according to the Wishart distribution with covariance matrix $\Sigma$, an degrees of freedom $t = n - 2p$. The joint density of $x$, $y$ and $z$ is given by

$$W(\Sigma, t) = k (xy - z^2)^{(t-3)/2} \exp \left[-(x/\sigma_{11} - 2z/\sqrt{\sigma_{11} \sigma_{22}} + y/\sigma_{22})/2 \right] (1 - \phi^2)$$

(3.1)

where $k = 1/[2^t \Sigma t^{t/2} \pi \Gamma(t/2) \Gamma((t-1)/2)]$. In the evaluation we made a transformation from the variables $x$, $y$ and $z$ to $r = z/\sqrt{xy}$, $v = z/y$ and $w = z$. The density, in these new variables with Jacobian $= 2w^2/vr^2$, is
\[ f(r,v,w) = k(2w^2/\nu)^{v-1/2} \exp \left\{ - \frac{w}{\sigma_{11} r^2 - 2\phi/\sqrt{\sigma_{11} \sigma_{22} + 1/\sigma_{22} v}} \right\} \] (3.2)

when \( w, v \in R, \) and \(-1 \leq r \leq +1.

Due to the nature of the transformation, the density in (3.2) is defined only when \( r,v, w \) are either all positive or all negative. As we see later, for our purpose, it is sufficient to consider only positive values of \( r. \) Therefore, from now on, we consider \( f(r,v,w) \) only when \( r, v, w \) are all positive and this means we assume a positive critical value.

Integrating (3.2) with respect to \( w, \) we have the following joint density of \( r \) and \( v 
\[ f(r,v) = 2k \frac{(1-r^2)^{v-3/2} \Gamma(v)/(\Gamma(r^2 \sigma_{11} - 2\phi/\sqrt{\sigma_{11} \sigma_{12} + 1/\sigma_{22} v})/2(1-\phi^2))}{vr' \rho} \] (3.3)

To obtain the marginal density of \( r \) from (3.3), we define

\[ g = \frac{1}{2} (1-\phi^2) \sigma_{11} \]
\[ h = \frac{-\phi}{(1-\phi^2) \sqrt{\sigma_{11} \sigma_{22}}} \]
\[ q = \frac{1}{2} (1-\phi^2) \sigma_{22} \]
\[ m = ((q/g) - hr^2/4g^3)^{1/2} \]
\[ s = v + hr^2/2g \]
\[ s = rm \tan \theta \]

\[ I_r = \int_{\theta^*}^{\pi/2} (\sin \theta)^{j-1} (\cos \theta)^{n-1} d\theta \]

\[ = \sum_{i=1}^{j/2} \binom{j-1}{i} \binom{(-1)}{(j-2i+1)} \binom{n/2}{(j-1)!/(a-j-1+2i)} \sin (\theta^*)^{j+1-2i} \cos (\theta^*)^{a-j-1+2i} \]

\[ + (j-1)!/(a-j-1+2i)! \int_{\theta^*}^{\pi/2} (\cos \theta)^a d\theta \]

and

\[ I_o = \int_{\theta^*}^{\pi/2} (\sin \theta)^{j-1} (\cos \theta)^{n-1} d\theta \]

\[ = \sum_{i=1}^{j+1/2} \binom{(-1)}{(j-1)!/(j-2i+1)!} \binom{n/2}{(j-1)!/(a-j-1+2i)!} \sin (\theta^*)^{j+1-2i} \cos (\theta^*)^{a-j-1+2i} \]

where \( \theta^* = \arctg hr/2gm, \) !! means double factorial and \( a = 2t-2. \) Then the probability density function of \( r \) is given by


\[ f(r) = 2(1-r^2)^{(t-3)/2} \frac{\Gamma(t) (1-\phi)^{t/2}}{\sqrt{\pi} \Gamma(t/2) \Gamma((t-1)/2)} \]

where \((I_e, I_o, j)\) means that we pick either \(I_e\) or \(I_o\) depending on whether \(j\) is even or odd.

In Figures 1 and 2, this distribution is plotted as a function of \(t = n-2p\) and \(\phi\). In Figure 1 where \(\phi = 0\), the distribution is symmetric for \(t = 10, 15\). The distribution for the larger \(t\) has more probability mass around zero, but goes to zero faster on either side as \(r\) differs from zero. In Figure 2, we show for \(t = 15\), the same distribution with \(\phi = .2\) and \(\phi = .4\). Under this scenario, as \(\phi\) gets larger there is more probability to the right. For example, \(P(r > 0|\phi = .2) = .72\), whereas \(P(r > 0|\phi = .4) = .88\).

![Fig. 1 The small sample distribution of r (t=10, 15 : \(\phi = 0\))](image-url)
4. THE RISK OF THE PRE-TEST ESTIMATOR

Since the derivation is symmetric and the calculations for the second sample are exactly similar, we can reduce the dimensionality of the coefficient vectors by two without affecting the results. Therefore, henceforth \( \alpha^*(1), \alpha^*(3) \) and \( \alpha^*(4) \) are \((p \times 1)\) vectors of estimators of the coefficients of the first sample only. Under squared error loss the risk of the SUPRE is given by

\[
\rho(\alpha^*(3), a_1) = trE \| I_{[-1,c]}(r) \alpha^*(1) + I_{[c+1,1]}(r) \alpha^*(4) - a_1 \|^2 \\
= trE [I_{[-1,c]}(r) (X_1'X_1)^{-1} X_1'y_1 - I_{[-1,c]}(r) a_1] \\
+ [I_{[c+1,1]}(r) (X_1'X_1)^{-1} X_1'y_1 - v(X_1'X_1)^{-1} X_1'y_2) \\
- I_{[c+1,1]}(r) a_1] \|^2 \\
\]

(4.1)
Using \((X'_1 X_1)^{-1} X'_1 y_1 = a_1 + (X'_1 X_1)^{-1} X'_1 e_1\) and 
\((X'_2 X_2)^{-1} X'_2 y_2 = (X'_2 X_2)^{-1} X'_2 e_2\) we have

\[
\rho(a^*(3), a_1) = trE [I_{[-1,1]}(r) (X'_1 X_1)^{-1} X'_1 e_1 
+ I_{(c,+1]}(r) (X'_1 X_1)^{-1} X'_1 e_1 
- I_{(c,+1]}(r) v(X'_1 X_1)^{-1} X'_1 e_1 \parallel 2 
- trE \| (X'_1 X_1)^{-1} X'_1 e_1 - I_{(c,+1]}(r) v(X'_1 X_1)^{-1} X'_1 e_1 \parallel 2
\]

(4.2)

where we can use the fact that \(I_{[-1,1]}(r) + I_{(c,+1]}(r) = 1\), since \(r \in [-1,1]\). Also, because the domains of the indicator functions are disjoint, this means that \(I_{[-1,c]}(r) I_{(c,+1]}(r) = 0\) and we obtain

\[
\rho(a^*(3), a_1) = \sigma_{11} tr(X'_1 X_1)^{-1} 
- 2 trE I_{(c,+1]}(r) v(X'_1 X_1)^{-1} X'_1 e_1 e_2' X_1 (X'_1 X_1)^{-1} 
+ trE I_{(c,+1]}(r) v^2 (X'_1 X_1)^{-1} X'_1 e_2 e_2' X_1 X'_1 (X'_1 X_1)^{-1}
\]

(4.3)

Using the independence of the following vectors, \((a^*(1), (X'_1 X_1)^{-1} X'_1 y_2, (X'_2 X'_2)^{-1} X'_2 y_1)\) and scale parameter estimates \((s_{11}, s_{22}, s_{12})\), yields

\[
\rho(a^*(3), a_1) = \sigma_{11} tr(X'_1 X_1)^{-1} 
- 2E I_{(c,+1]}(r) v trE ((X'_1 X_1)^{-1} X'_1 e_1 e_2' X_1 (X'_1 X_1)^{-1}) 
+ E(I_{(c,+1]}(r) v^2) trE ((X'_1 X_1)^{-1} X'_1 e_2 e_2' X_1 (X'_1 X_1)^{-1})
\]

(4.4)

In order to compare the risks of SURPE, Zellner’s GLSE and LSE, all risk evaluations are made with respect to the LSE risk, \(\sigma_{11} tr(X'_1 X_1)^{-1}\). Therefore, the relative risk is

\[
\frac{\rho(a^*(3), a_1)}{\rho(a^*(1), a_1)} = 1 - 2E I_{(c,+1]}(r) v(\sigma_{12}/\sigma_{11}) + E I_{(c,+1]}(r) v^2 (\sigma_{22}/\sigma_{11})
\]

(4.5)

Here we should note that \(r\) in the argument of the indicator function in (4.5) is positive unless we choose a negative value of \(c\). That is why, in section 2 the joint distribution \(f(r, v, w)\) is considered only for the positive values of \(r, v\) and \(w\) [see equation (3.2)].

The relative risk values of the SURPE with respect to that of LSE are given as a function of the population correlation coefficient \(\phi\) and the critical value of the test \(c\), in Table 1, for \(t = 10, 15, \text{ and } 20\) respectively, when \(\sigma_{11} = \sigma_{22} = 1\). These values are obtained by calculating the expectations in (4.5) with respect to the joint distribution of \(r\) and \(v\) given in equation (3.5). These expectations were solved numerically since analytical approach involved intractable algebraic computations.

From the tabulated values of the relative risk of SURPE, that is a function of \(\phi\) and the critical value \(c\) used in the preliminary testing, we notice that over the range of the \((\phi, c)\) parameter space, the relative risks of the pretest estimators cross. As larger and larger critical values are used, the LSE is used more frequently and this causes the relative risk
### Table 1

Relative risk values of SURPE as a function of the population correlation coefficient $\phi$ and the critical value $c$

<table>
<thead>
<tr>
<th>$c$</th>
<th>$0.1$</th>
<th>$0.3$</th>
<th>$0.5$</th>
<th>$0.7$</th>
<th>$0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.9$</td>
<td>$1.0004$</td>
<td>$1.0009$</td>
<td>$1.0002$</td>
<td>$0.9775$</td>
<td>$0.5551$</td>
</tr>
<tr>
<td>$0.8$</td>
<td>$1.0040$</td>
<td>$1.0072$</td>
<td>$0.9967$</td>
<td>$0.8753$</td>
<td>$0.3030$</td>
</tr>
<tr>
<td>$0.7$</td>
<td>$1.0133$</td>
<td>$1.0180$</td>
<td>$0.9803$</td>
<td>$0.7652$</td>
<td>$0.2413$</td>
</tr>
<tr>
<td>$0.6$</td>
<td>$1.0273$</td>
<td>$1.0273$</td>
<td>$0.9517$</td>
<td>$0.6837$</td>
<td>$0.2247$</td>
</tr>
<tr>
<td>$0.5$</td>
<td>$1.0425$</td>
<td>$1.0303$</td>
<td>$0.9187$</td>
<td>$0.6332$</td>
<td>$0.2196$</td>
</tr>
<tr>
<td>$0.4$</td>
<td>$1.0552$</td>
<td>$1.0263$</td>
<td>$0.8887$</td>
<td>$0.6050$</td>
<td>$0.2179$</td>
</tr>
<tr>
<td>$0.3$</td>
<td>$1.0630$</td>
<td>$1.0178$</td>
<td>$0.8660$</td>
<td>$0.5907$</td>
<td>$0.2174$</td>
</tr>
<tr>
<td>$0.2$</td>
<td>$1.0648$</td>
<td>$0.9997$</td>
<td>$0.8426$</td>
<td>$0.5815$</td>
<td>$0.2172$</td>
</tr>
<tr>
<td>$t = 15$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.9$</td>
<td>$1.0000$</td>
<td>$1.0000$</td>
<td>$1.0000$</td>
<td>$0.9924$</td>
<td>$0.5623$</td>
</tr>
<tr>
<td>$0.8$</td>
<td>$1.0001$</td>
<td>$1.0005$</td>
<td>$0.9870$</td>
<td>$0.8163$</td>
<td>$0.2563$</td>
</tr>
<tr>
<td>$0.7$</td>
<td>$1.0017$</td>
<td>$1.0041$</td>
<td>$0.9807$</td>
<td>$0.7554$</td>
<td>$0.2129$</td>
</tr>
<tr>
<td>$0.6$</td>
<td>$1.0064$</td>
<td>$1.0085$</td>
<td>$0.9436$</td>
<td>$0.6459$</td>
<td>$0.2128$</td>
</tr>
<tr>
<td>$0.5$</td>
<td>$1.0146$</td>
<td>$1.0085$</td>
<td>$0.8967$</td>
<td>$0.5880$</td>
<td>$0.2048$</td>
</tr>
<tr>
<td>$0.4$</td>
<td>$1.0240$</td>
<td>$1.0011$</td>
<td>$0.8553$</td>
<td>$0.5626$</td>
<td>$0.2047$</td>
</tr>
<tr>
<td>$0.3$</td>
<td>$1.0310$</td>
<td>$0.9885$</td>
<td>$0.8271$</td>
<td>$0.5530$</td>
<td>$0.2046$</td>
</tr>
<tr>
<td>$0.2$</td>
<td>$1.0307$</td>
<td>$0.9651$</td>
<td>$0.8049$</td>
<td>$0.5491$</td>
<td>$0.2046$</td>
</tr>
<tr>
<td>$t = 20$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.9$</td>
<td>$1.0000$</td>
<td>$1.0000$</td>
<td>$1.0000$</td>
<td>$0.9972$</td>
<td>$0.5665$</td>
</tr>
<tr>
<td>$0.8$</td>
<td>$1.0000$</td>
<td>$1.0002$</td>
<td>$0.9987$</td>
<td>$0.9192$</td>
<td>$0.2348$</td>
</tr>
<tr>
<td>$0.7$</td>
<td>$1.0004$</td>
<td>$1.0015$</td>
<td>$0.9848$</td>
<td>$0.7528$</td>
<td>$0.2200$</td>
</tr>
<tr>
<td>$0.6$</td>
<td>$1.0022$</td>
<td>$1.0040$</td>
<td>$0.9450$</td>
<td>$0.6266$</td>
<td>$0.2195$</td>
</tr>
<tr>
<td>$0.5$</td>
<td>$1.0070$</td>
<td>$1.0031$</td>
<td>$0.8979$</td>
<td>$0.5675$</td>
<td>$0.2135$</td>
</tr>
<tr>
<td>$0.4$</td>
<td>$1.0143$</td>
<td>$0.9942$</td>
<td>$0.8413$</td>
<td>$0.5465$</td>
<td>$0.2090$</td>
</tr>
<tr>
<td>$0.3$</td>
<td>$1.0207$</td>
<td>$0.9790$</td>
<td>$0.8107$</td>
<td>$0.5402$</td>
<td>$0.2088$</td>
</tr>
<tr>
<td>$0.2$</td>
<td>$1.0212$</td>
<td>$0.9524$</td>
<td>$0.7907$</td>
<td>$0.5376$</td>
<td>$0.2086$</td>
</tr>
</tbody>
</table>

The effect of degrees of freedom on these results is minimal.

The critical values of the SURPE for significance levels .05 and .10 are respectively .60 and .45. The relative risks of LSE and Zellner's GLSE for $t = 10$ are presented in Figure 3. The risk values of Zellner's estimator are taken from Zellner (1963, p. 983). It should be noted that Zellner's (1963) considers unrestricted residuals whereas in this paper we use restricted residuals. Revankar (1976) finds that in many practical situations there is little to choose between the feasible GLSE using the two definitions of the residuals.
on efficiency grounds. Therefore, our use of Zellner's results could be partially justified. Many earlier papers discussed properties of feasible GLSE and those are not repeated here. From Figure 3 we observe that the relative risk of the SURPE with \( c = .60 \), starts below that of \( c = .45 \), crosses the latter around \( \phi = .3 \), and remains above for all \( \phi > .3 \). This means that throughout the \((c, \phi)\) parameters space, no one SURPE is risk superior to the other. The SURPE with \( c = .6 \) is risk superior to SURPE with \( c = .45 \), for \( \phi \) close to zero. In turn it is risk inferior once \( \phi \) exceeds .3. This relationship between the SURPE's with different critical values holds true throughout. In general, as can be observed from Table 1, the SURPE with a larger critical value has a small sampling variability when \( \phi \) is small, but then performs worse after its risk crosses that of the SURPE with a smaller critical value.

The relative risk function of Zellner's GLSE is also presented in Figure 3. Its risk is highest for small \( \phi \), and then crosses the risks of LSE, SURPE \((c = .6)\) and finally SURPE \((c = .45)\) as \( \phi \) gets larger. Therefore, under squared error loss, none of the estimators in
Figure 3 dominates. However, it is interesting to note that there is a range of \( \phi \) where SURPE is better than both LSE and GLSE. This is not the case in the regression coefficient pretesting. However, this result is observed in other pre-test situations, for example, see Toyoda and Wallace (1975), Ohtani and Toyoda (1978, 1980) and Ohtani (1988). A possible reason for this might be the fact that \( 0 \leq \phi \leq 1 \) prevents the pretest from making any disastrous type I and type II errors. The SURPE with \( 0 < c < 1 \) at \( \phi = 0 \) starts with a risk in between that of the LSE and the GLSE. It ends with a risk in between these two estimators when \( \phi = 1 \). One can also see that the SURPE has a substantial risk gain over the LSE for large \( \phi \), and the risk loss is modest when \( \phi \) is close to zero. When the critical value \( c \) takes on extreme values, the risk of SURPE approaches the risk of the LSE or the risk of the GLSE depending whether \( c \) tends to 1 or to -1. Similar comparisons can be made for the same estimators in Figure 4 with \( t = 10 \) where
the critical values .5 and .35 correspond to significance levels .05 and .1 respectively. As \( t \) increases, Zellner's GLSE becomes more efficient, and in fact approaches asymptotic efficiency levels.

5. SUMMARY AND LIMITATIONS

We have made risk comparisons between the SURPE, LSE and Zellner's GLSE in the two sample seemingly unrelated regression model and found that no one estimator is uniformly superior. However, we can now determine the risk gains that accrue when the pre-test estimator is used to take advantage of the risk superiority of LSE, when \( \phi \) is close to zero, and the GLSE is used when \( \phi \) is close to 1. Alternatively, we can determine the risk consequences of always using the pre-test rule. Our results suggest searching for an optimal critical value for the pre-test according to some optimality criterion. This is a major issue, and is enough for another paper in its own right. There are a number of studies which investigate this problem of finding optimal critical values for other pre-test problems, for example, Toyoda and Wallace (1975, 1976) and Ohtani and Toyoda (1980) derived optimal critical points using a minimum average relative risk criterion while Ohtani and Toyoda (1978) used a minimax regret criterion. Until an optimal critical value has been developed for SURPE, our results suggest that for sample sizes and critical values normally used in practice, if the applied researcher uses SURPE then (1) the risk consequences relative to GLSE will be minimal and (2) significant risk gain over LSE will accrue over much of the \( \phi \) parameter space. Thus contrary to many other pre-testing situations, our risk results point to the normative content of SURPE in applied risk. We should also mention that our results have been obtained under some restrictive assumptions such as the regressors are orthogonal and the two regression equations have the same variance and the same number of regressors. It is not clear whether our results will be still valid when these restrictions are relaxed. We leave these important issues for future research.

REFERENCES


Reprint Series, CentER, Tilburg University, The Netherlands:


No. 5  Th. ten Raa and F. van der Ploeg, A statistical approach to the problem of negatives in input-output analysis, *Economic Modelling*, vol. 6, no. 1, 1989, pp. 2 - 19.


No. 8  Th. van de Klundert and F. van der Ploeg, Wage rigidity and capital mobility in an optimizing model of a small open economy, *De Economist*, vol. 137, nr. 1, 1989, pp. 47 - 75.


No. 68  F. van der Ploeg, Macroeconomic policy coordination issues during the various phases of economic and monetary integration in Europe, European Economy - The Economics of EMU, Commission of the European Communities, special edition no. 1, 1991, pp. 136 - 164.


<table>
<thead>
<tr>
<th>No.</th>
<th>Reference</th>
</tr>
</thead>
</table>


<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Authors</th>
<th>Source</th>
</tr>
</thead>
</table>