BEBR
FACULTY WORKING
PAPER NO. 90-1668

The Risk Properties of a Pre-Test Estimator for Zellner's Seemingly Unrelated Regression Model

Ahmet Ozcam
George Judge
Anil Bera
Thomas Yancey

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 90-1668

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

July 1990

The Risk Properties of a Pre-Test Estimator for Zellner's Seemingly Unrelated Regression Model

Ahmet Özcam
Research Division of Turkish Parliment
Ankara, Turkey

George Judge
University of California, Berkeley

Anil Bera & Thomas Yancey
Department of Economics
University of Illinois at Urbana-Champaign

Abstract

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.

Key Words: Risk, Pre-Test estimator, Least squares estimator, Generalized least squares estimator, Seemingly unrelated regression model, Test statistic.

THE RISK PROPERTIES OF A PRE-TEST ESTIMATOR FOR ZELLNER'S SEEMINGLY UNRELATED REGRESSION MODEL[1]

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 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),

---

[1]The authors acknowledge the helpful comments of David Giles and Helga Hessenius. This work was partially supported by National Science Foundation grant, SES-86-96152.

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 ϕ 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 and discusses the implications of the paper.

2. Statistical Model and Estimators

Consider the following two sample regression model

$$\begin{bmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \underset{\sim}{\alpha}_1 \\ \underset{\sim}{\alpha}_2 \end{bmatrix} + \begin{bmatrix} \underset{\sim}{e}_1 \\ \underset{\sim}{e}_2 \end{bmatrix}, \text{ or } \underset{\sim}{y} = X\underset{\sim}{\alpha} + \underset{\sim}{e} \qquad (2.1)$$

where y_i is a (nxl) vector of observations, X_i is a (nxp) matrix of fixed regressors of rank p, $\underset{\sim}{\alpha}_i$ is a (pxl) unknown location vector, and $\underset{\sim}{e}_i$ is an (nxl) random error vector for i = 1,2. For expository purposes we assume that $X_1'X_2 = X_2'X_1 = 0_p$. We further assume that the equation errors are distributed as multivariate normal random variables with zero means and covariance matrix

$$\Sigma = E\begin{bmatrix} \underset{\sim}{e}_1 \\ \underset{\sim}{e}_2 \end{bmatrix} [\underset{\sim}{e}_1' \ \underset{\sim}{e}_2'] = E[\underset{\sim}{e}\underset{\sim}{e}'] = \begin{bmatrix} \sigma_{11}I_n & \sigma_{12}I_n \\ \sigma_{21}I_n & \sigma_{22}I_n \end{bmatrix} \tag{2.2}$$

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

$$\underset{\sim}{\alpha}^*(1) = \begin{bmatrix} (X_1'X_1)^{-1}X_1'y_1 \\ (X_2'X_2)^{-1}X_2'y_2 \end{bmatrix} \tag{2.3}$$

The Zellner SUR estimator

$$\underset{\sim}{\alpha}^*(2) = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\underset{\sim}{y} \tag{2.4}$$

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, α*(4). One choice for the elements of $S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}$ is $s_{ij} = \frac{1}{n}(\underset{\sim}{y}_i - X_i\underset{\sim}{\alpha}_i^*(1))'(\underset{\sim}{y}_j - X_j\underset{\sim}{\alpha}_j^*(1))$, $i,j = 1,2$.

Now the feasible GLSE is given by

$$\alpha^*(4) = \left[\begin{bmatrix} X_1' & 0 \\ 0 & X_2' \end{bmatrix}\begin{bmatrix} s^{11}I_n & s^{12}I_n \\ s^{21}I_n & s^{22}I_n \end{bmatrix}\begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}\right]^{-1}\begin{bmatrix} X_1' & 0 \\ 0 & X_2' \end{bmatrix}\begin{bmatrix} s^{11}I_n & s^{12}I_n \\ s^{21}I_n & s^{22}I_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} (X_1'X_1)^{-1} X_1'y_1 - (s^{12}/s^{11})(X_1'X_1)^{-1}X_1'y_2 \\ (X_2'X_2)^{-1} X_2'y_2 - (s^{12}/s^{22})(X_2'X_2)^{-1}X_2'y_1 \end{bmatrix} \qquad (2.5)$$

where we have used the assumption $X_1'X_2 = X_2'X_1 = 0_p$ and the s^{ij} are the elements of $S^{-1} = \begin{bmatrix} s^{11} & s^{12} \\ s^{21} & s^{22} \end{bmatrix}$. The estimates of the variances and the covariances are obtained from the restricted residuals, that are obtained from regressing $\underset{\sim}{y}_i$ 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_o: $\phi = 0$ that the population correlation coefficient ϕ is zero, versus a one-sided alternative H_a: $\phi > 0$. 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_o, 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_o, the SURPE is the LSE, and otherwise it is the GLSE. This means the SURPE is

$$\underset{\sim}{\alpha}^*(3) = I_{[-1,c]}(r)\underset{\sim}{\alpha}^*(1) + I_{(c,+1]}(r)\underset{\sim}{\alpha}^*(4) \tag{2.6}$$

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

3. The Small Sample Distribution of r

The distribution of SURPE $\underset{\sim}{\alpha}^*(3)$ 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 $ns_{11}=x$, $ns_{22}=y$ and $ns_{12}=z$ are distributed according to the Wishart distribution with covariance matrix Σ, and degrees of freedom $t = n-2p$. The joint density of x, y and z is given by

$$W(\Sigma,t) = k(xy-z^2)^{(t-3)/2}\exp[-(x/\sigma_{11} - 2\phi z/\sqrt{\sigma_{11}\sigma_{22}} + y/\sigma_{22})/2(1-\phi^2)] \tag{3.1}$$

where $k = 1/[2^t|\Sigma|^{t/2}\sqrt{\pi}\,\Gamma(t/2)\Gamma((t-1)/2)]$. In the evaluation we make 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 = $2w^2/vr^3$, is

$$f(r,v,w) = k(2w^2/vr^3)(w^2/r^2-w^2)^{(t-3)/2}$$
$$\exp\{-w(v/\sigma_{11}r^2 - 2\phi/\sqrt{\sigma_{11}\sigma_{22}} + 1/\sigma_{22}v)/2(1-\phi^2)\} \tag{3.2}$$

when w, $v \in R$, and $1 \le r \le +1$.

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.

To obtain the joint density of r and w, we integrate out w by using the gamma function

$$f(r,v) = 2k(1-r^2)^{(t-3)/2}\Gamma(t)/((v/r^2\sigma_{11}-2\phi\sqrt{\sigma_{11}\sigma_{22}}+1/v\sigma_{22})/2(1-\phi^2))^t vr^t \qquad (3.3)$$

If we define

$$g = 1/2(1-\phi^2)\sigma_{11} > 0,$$

$$h = -\phi/(1-\phi^2)\sqrt{\sigma_{11}\sigma_{22}} \in R,$$

$$\text{and } q = 1/2(1-\phi^2)\sigma_{22} > 0, \qquad (3.4)$$

the density in (3.3) may be written compactly as

$$f(r,v) = 2k(1-r^2)^{(t-3)/2}\Gamma(t)v^{t-1}/r^t((gv^2/r^2)+hv+q)^t$$

$$= 2k(1-r^2)^{(t-3)/2}\Gamma(t)v^{t-1}/r^t g^t((v^2/r^2)+(hv/g)+q/g)^t$$

$$= 2k(1-r^2)^{(t-3)/2}\Gamma(t)v^{t-1}/r^t g^t[((v/r)+hr/2g)^2+(q/g)-h^2r^2/4g^2]^t \qquad (3.5)$$

This completes the derivation of the joint density of r and v. To obtain the marginal distribution of r we define $m(r) = ((q/g)-h^2r^2/4g^2)^{1/2}$ and make a change of variable in v, $x = v + hr^2/2g$ and $r = r$. This gives

$$f(r,x) = 2k(1-r^2)^{(t-3)/2}\Gamma(t)(x-hr^2/2g)^{t-1}r^t/g^t(x^2+r^2m(r)^2)^t$$

$$= 2k\Gamma(t)(1-r^2)^{(t-3)/2}r^t \sum_{j=0}^{t-1} \binom{t-1}{j}(-hr^2/2g)^{t-1-j}x^j/g^t(x^2+r^2m(r)^2)^t$$

where $x > hr^2/2g$ (3.6)

Next we substitute $x = rm(r) \tan \Theta$ and obtain

$$f(r,\Theta) = \frac{2k\Gamma(t)(1-r^2)^{(t-3)/2}}{g^t} \sum_{j=0}^{t-1} \binom{t-1}{j} \frac{(-hr^2/2g)^{t-1-j}\sin^j\Theta\cos^{w-j}\Theta}{m(r)^{w-j+1}r^{w/2-j}}$$

where $w = 2t-2$ and $\mathrm{arctg}(hr/2gm(r)) < \Theta < \pi/2$. (3.7)

To integrate out Θ, we use successive integration by parts. This method depends heavily on j being even since the reduction from the integration by parts is by two at each step. Hence we distinguish two cases i) j is even, and ii) j is odd. The value of the integrals for even j is given by

$$I_e = \int_{\Theta^*}^{\pi/2} (\sin \Theta)^j(\cos \Theta)^{w-j}d\Theta$$

$$= \sum_{i=1}^{j/2} (j-1)!!(-1)/(j-2i+1)!!$$

$$\times ((w-j-1)!!/(w-j-1+2i)!!)\sin(\Theta^*)^{j+1-2i}\cos(\Theta^*)^{w-j-1+2i}$$

$$+ (j-1)!!(w-j-1)!!/(w-1)!!\int_{\Theta^*}^{\pi/2} (\cos \Theta)^w d\Theta \quad (3.8)$$

where $\Theta^* = \text{arctg } hr/2gm(r)$ and !! means double factorial. The integral in (3.8) can be evaluated by using the value given in Gradshteyn and Ryznik (1980).

$$\int_{\Theta^*}^{\pi/2} (\cos\Theta)^w d\Theta = ((1/2^w)\binom{w}{t-1}\Theta^* + 1/2^{w-1}) \sum_{k=0}^{t-2} \binom{w}{k} \sin\{(w-2k)\Theta^*/(w-2k)\} \quad (3.9)$$

When j is odd, the odd terms of the summation indexed by j in (3.7) can be integrated using

$$I_o = \int_{\Theta^*}^{\pi/2} (\sin\Theta)^j (\cos\Theta)^{w-j} d\Theta$$

$$= \sum_{i=1}^{\frac{j+1}{2}} ((-1)(j-1)!!/(j-2i+1)!!)$$

$$((w-j-1)!!/(w-j-1+2i)!!)\sin(\Theta^*)^{j+1-2i}\cos(\Theta^*)^{w-j-1+2i} \quad (3.10)$$

Finally using I_e or I_o depending on whether j is even or odd, Θ in (3.7) can be integrated out to compute the marginal distribution of the test statistic. This is given by

$$f(r) = \frac{2(1-r^2)^{(t-3)/2}\Gamma(t)(1-\phi^2)^{t/2} \sum_{j=0}^{t-1} \binom{t-1}{j} (\phi r^{t-1-j}(I_e, I_o, j)/(1-\phi^2 r^2)^{t-1/2-3/2}}{\sqrt{\pi}\,\Gamma(t/2)\Gamma((t-1)/2)} \quad (3.11)$$

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 ϕ. 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 ϕ gets larger there is more probability to the right. For example, $P(r>0|\phi=.2)=.72$, whereas $(P(r>0|\phi=.4)=.88$.

4. The Risk of the Pre-test Estimator (SURPE)

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 $\underset{\sim}{\alpha}^*(1)$, $\underset{\sim}{\alpha}^*(3)$ and $\underset{\sim}{\alpha}^*(4)$ are (pxl) vectors of estimators of the coefficients of the first sample only. Under squared error loss the risk of the SURPE is given by

$$\begin{aligned}\rho(\underset{\sim}{\alpha}^*(3),\underset{\sim}{\alpha}_1) &= \mathrm{tr}E||I_{[-1,c]}(r)\underset{\sim}{\alpha}^*(1) + I_{(c,+1]}(r)\underset{\sim}{\alpha}^*(4)-\underset{\sim}{\alpha}_1||^2 \\ &= \mathrm{tr}E||[I_{[-1,c]}(r)(X_1'X_1)^{-1}X_1'\underset{\sim}{y}_1 - I_{[-1,c]}(r)\underset{\sim}{\alpha}_1] \\ &\quad + [I_{(c,+1]}(r)\{(X_1'X_1)^{-1}X_1'\underset{\sim}{y}_1 - v(X_1'X_1)^{-1}X_1'\underset{\sim}{y}_2\} \\ &\quad - I_{(c,+1]}(r)\underset{\sim}{\alpha}_1]||^2 \qquad (4.1)\end{aligned}$$

Using $(X_1'X_1)^{-1}X_1'\underset{\sim}{y}_1 = \underset{\sim}{\alpha}_1+(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_1$ and $(X_1'X_1)^{-1}X_1'\underset{\sim}{y}_2 = (X_1'X_1)^{-1}X_1'\underset{\sim}{e}_2$ we have

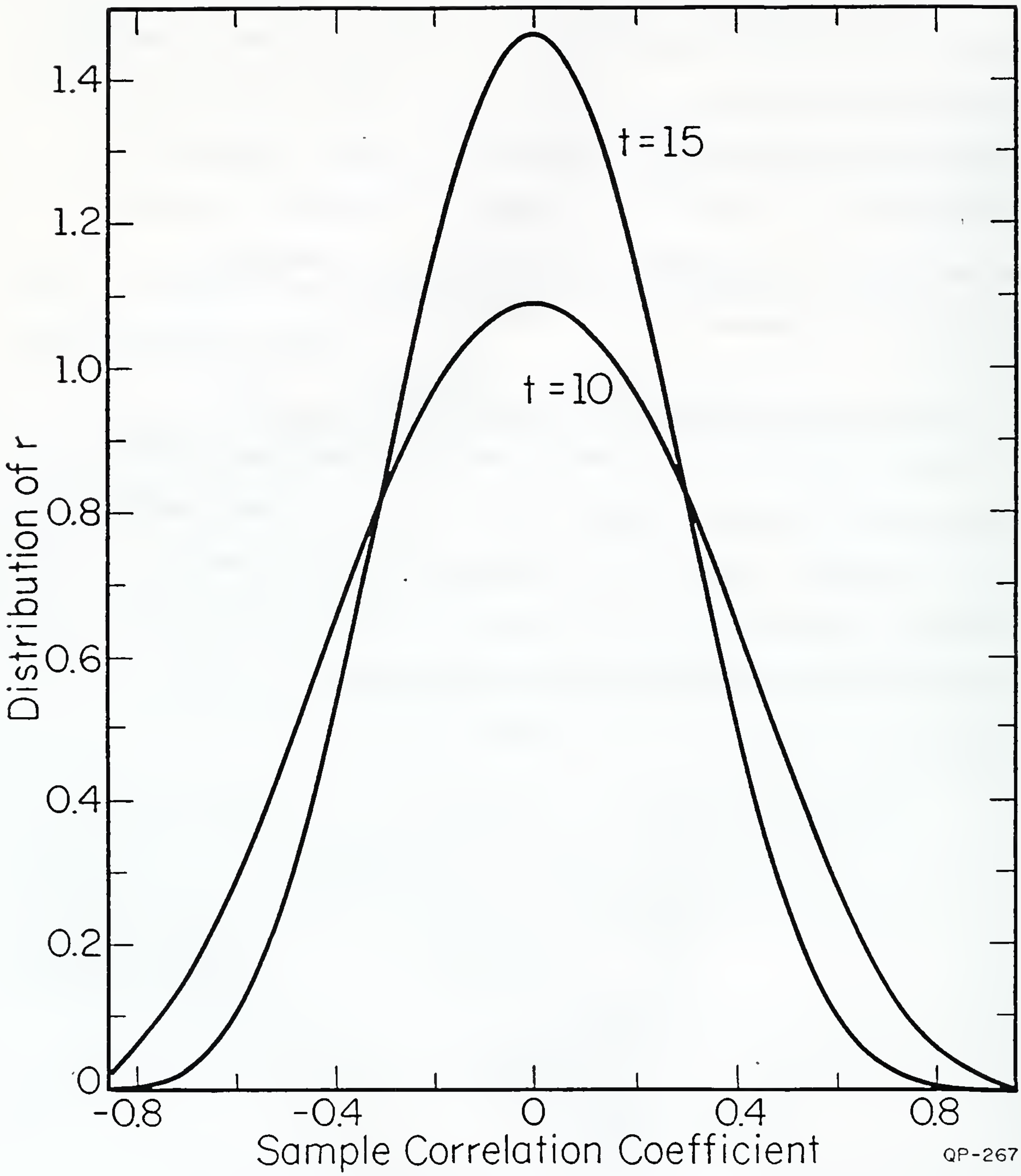

FIG.1. THE SMALL SAMPLE DISTRIBUTION OF R (t=10, 15: ϕ=0)

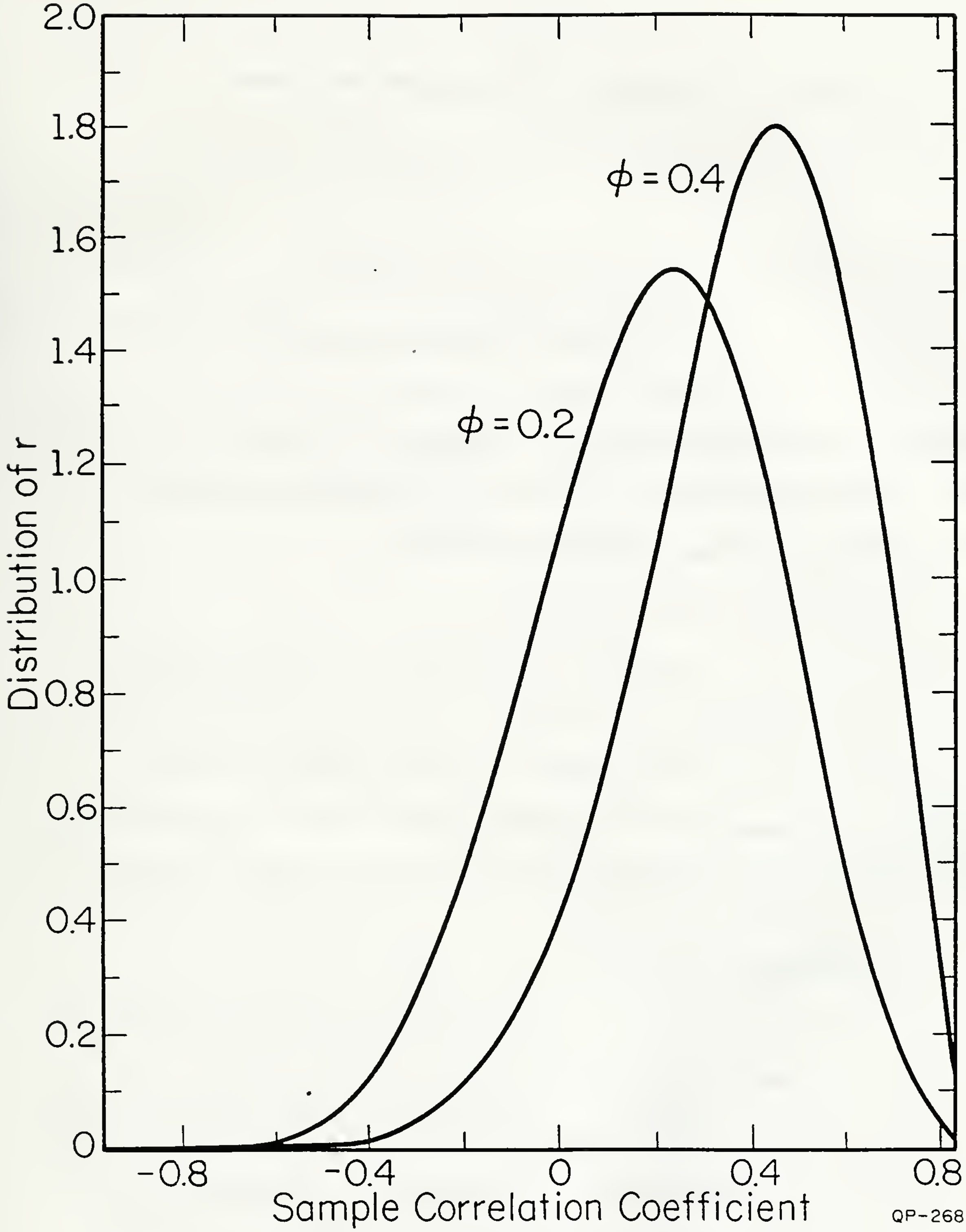

FIG.2. THE SMALL SAMPLE DISTRIBUTION OF R (t=15: ϕ=0.2, 0.4)

$$\rho(\underset{\sim}{\alpha}^*(3),\underset{\sim}{\alpha}_1) = \mathrm{tr}E||[I_{[-1,c]}(r)(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_1$$

$$+ I_{(c,+1]}(r)(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_1$$

$$- I_{(c,+1]}(r)v(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_2]||^2$$

$$= \mathrm{tr}E||(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_1 - I_{(c,+1]}(r)v(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_2||^2 \quad (4.2)$$

where we use the fact that $I_{[-1,c]}(r) + I_{(c,+1]}(r) = 1$, 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

$$\rho(\underset{\sim}{\alpha}^*(3),\underset{\sim}{\alpha}_1) = \sigma_{11}\mathrm{tr}(X_1'X_1)^{-1}$$

$$- 2\mathrm{tr}E\{I_{(c,+1]}(r)v(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_1\underset{\sim}{e}_2'X_1(X_1'X_1)^{-1}\}$$

$$+\mathrm{tr}E\{I_{(c,+1]}(r)v^2(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_2\underset{\sim}{e}_2'X_1(X_1'X_1)^{-1}\} \quad (4.3)$$

Using the independence of the following vectors, ($\underset{\sim}{\alpha}^*(1)$, $(X_1'X_1)^{-1}X_1'\underset{\sim}{y}_2$, $(X_2'X_2)^{-1}X_2'\underset{\sim}{y}_1$) and the scale parameter estimates (s_{11}, s_{22}, s_{12}), yields

$$\rho(\underset{\sim}{\alpha}^*(3),\underset{\sim}{\alpha}_1) = \sigma_{11}\mathrm{tr}(X_1'X_1)^{-1}$$

$$- 2E\{I_{(c,+1]}(r)v\}\mathrm{tr}E\{(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_1\underset{\sim}{e}_2'X_1(X_1'X_1)^{-1}\}$$

$$+ E\{I_{(c,+1]}(r)v^2\}\mathrm{tr}E\{(X_1'X_1)^{-1}X_1'\underset{\sim}{e}_2\underset{\sim}{e}_2'X_1(X_1'X_1)^{-1}\}$$

$$= \sigma_{11}\mathrm{tr}(X_1'X_1)^{-1} - 2\sigma_{12}E\{I_{(c,+1]}(r)v\}\mathrm{tr}(X_1'X_1)^{-1}$$

$$+ \sigma_{22}\mathrm{tr}(X_1'X_1)^{-1}E\{I_{(c,+1]}(r)v^2\} \quad (4.4)$$

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

$$\frac{\rho(\underset{\sim}{\alpha}^*(3),\underset{\sim}{\alpha}_1)}{\rho(\underset{\sim}{\alpha}^*(1),\underset{\sim}{\alpha}_1)} = 1-2\ E\{I_{(c,+1]}(r)v\}(\sigma_{12}/\sigma_{11}) + E\{I_{(c,+1]}(r)v^2\}(\sigma_{22}/\sigma_{11}) \quad (4.5)$$

Here we should note that the 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 ϕ and the critical value of the test c, in Table 1, for t = 10, 15, 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 that is derived in Section 3.

From the tabled values of the relative risk of SURPE, that is a function of ϕ and the critical value c used in the preliminary testing, we notice that over the range of the (ϕ,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 of the SURPE to decrease for ϕ close to zero, and to increase for ϕ close to one. 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

TABLE 1

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

		ϕ				
	c	.1	.3	.5	.7	.9
	.9	1.0004	1.0009	1.0002	0.9775	0.5551
	.8	1.0040	1.0072	0.9967	0.8753	0.3030
	.7	1.0133	1.0180	0.9803	0.7652	0.2413
t = 10	.6	1.0273	1.0273	0.9517	0.6837	0.2247
	.5	1.0425	1.0303	0.9187	0.6332	0.2196
	.4	1.0552	1.0263	0.8887	0.6050	0.2179
	.3	1.0630	1.0178	0.8660	0.5907	0.2174
	.0	1.0648	0.9997	0.8426	0.5815	0.2172
	.9	1.0000	1.0000	1.0000	0.9924	0.5623
	.8	1.0001	1.0005	0.9870	0.8163	0.2563
	.7	1.0017	1.0041	0.9807	0.7554	0.2129
t = 15	.6	1.0064	1.0085	0.9436	0.6459	0.2128
	.5	1.0146	1.0085	0.8967	0.5880	0.2048
	.4	1.0240	1.0011	0.8553	0.5626	0.2047
	.3	1.0310	0.9885	0.8271	0.5530	0.2046
	.0	1.0307	0.9651	0.8049	0.5491	0.2046
	.9	1.0000	1.0000	1.0000	0.9972	0.5665
	.8	1.0000	1.0002	0.9987	0.9192	0.2348
	.7	1.0004	1.0015	0.9848	0.7528	0.2200
t = 20	.6	1.0022	1.0040	0.9450	0.6266	0.2195
	.5	1.0070	1.0031	0.8979	0.5675	0.2135
	.4	1.0143	0.9942	0.8413	0.5465	0.2090
	.3	1.0207	0.9790	0.8107	0.5402	0.2088
	.0	1.0212	0.9524	0.7907	0.5376	0.2086

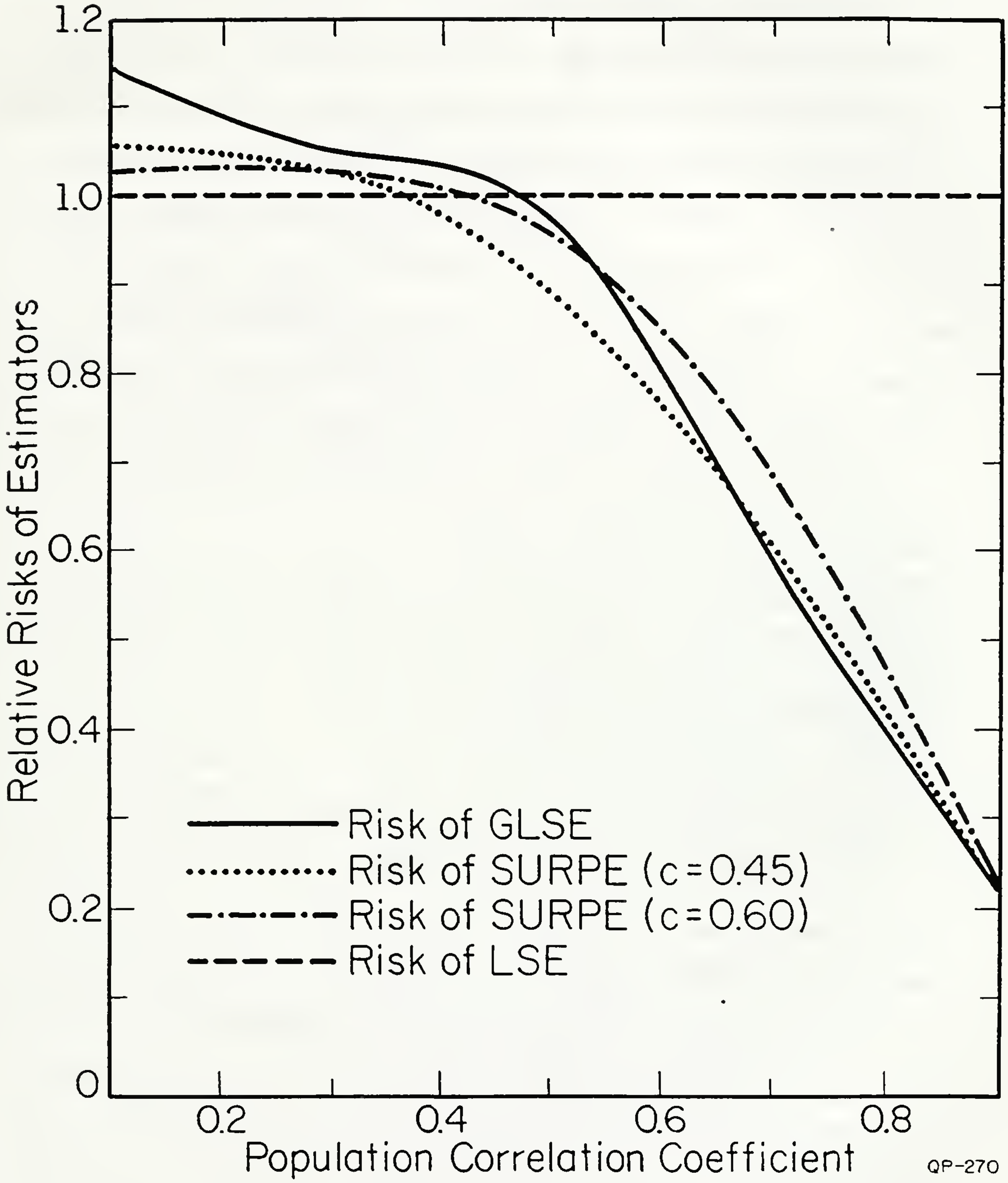

FIG.3. RISK VALUES OF SURPE ESTIMATORS (t=10)

of Zellner's estimator are taken from Zellner's paper (1963, p. 983). We observe that the relative risk of the SURPE with c = .60, starts below that of c = .45, crosses the latter around ϕ = .3, and remains above for all $\phi > .3$. This means that throughout the (c,v) parameter space, no one SURPE is risk superior to the other. The SURPE with c = .6 is risk superior to SURPE with c = .45, for ϕ close to zero. In turn it is risk inferior once ϕ 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 ϕ 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 ϕ, and then crosses the risks of LSE, SURPE (c=.6) and finally SURPE (c=.45) as ϕ 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 ϕ where SURPE is better than both LSE and GLSE. This is not the case in the regression coefficient pretesting. 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 ϕ, and the risk loss is modest when ϕ 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 $t = 20$ 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 Conclusions

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 ϕ is close to zero, and the GLSE is used when ϕ is close to 1. Alternatively, we can determine the risk consequences of always using the pre-test rule. Finally, we examined the distribution of the test statistic r, evaluated some probabilities by numerical methods and found that the distribution when $\phi = 0$, is symmetric around zero, but skewed to the left for $\phi > 0$.

The applied statistician can gain insight into the nature of the correlation of disturbances of the underlying model, by conducting a preliminary test. Consequently in many situations, the risk advantages of the SUR pre-test estimator over the LSE and the GLSE can be exploited. For example in a somewhat different context, Stanek (1988) considers an experimental design which permits a variety of hypotheses

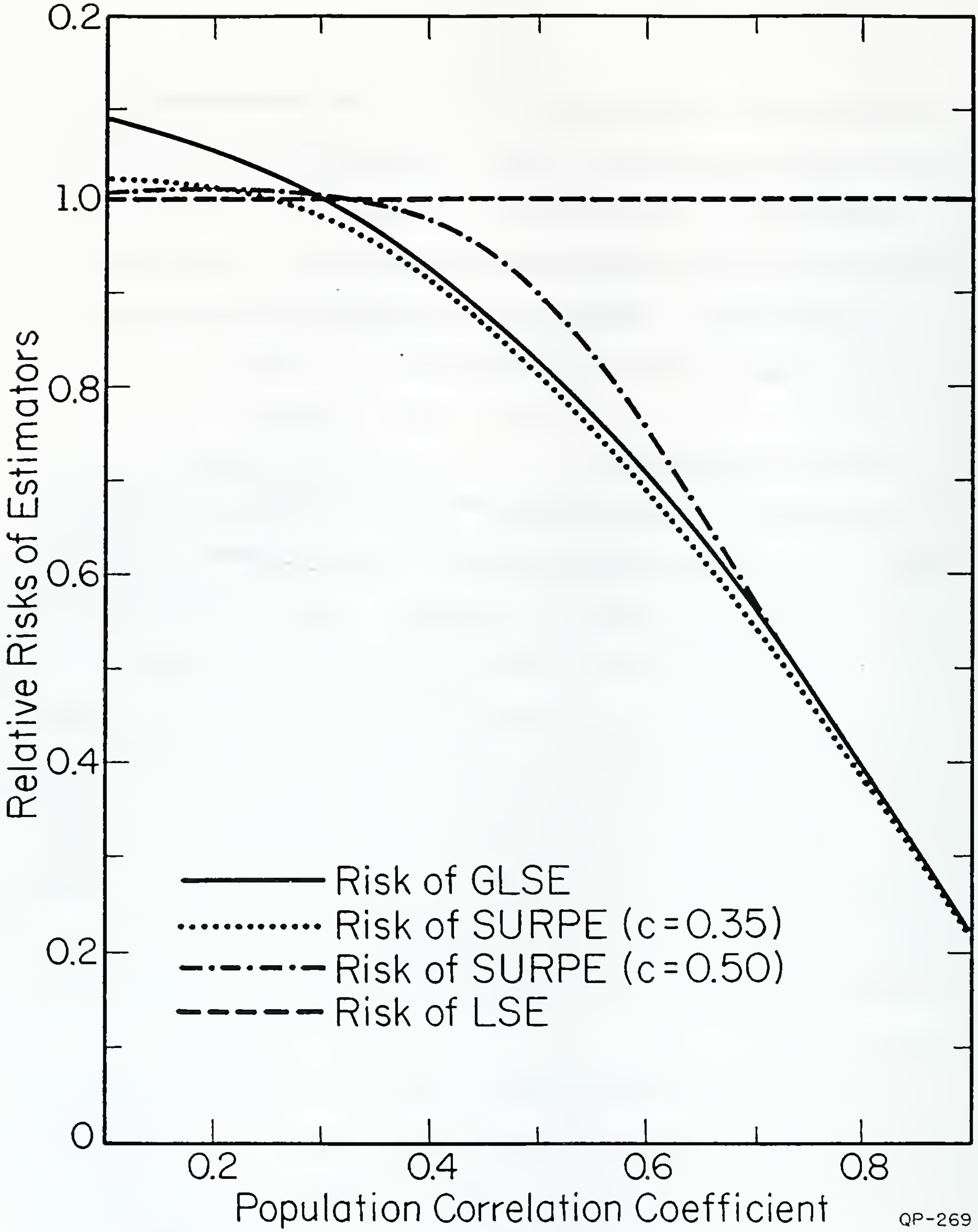

FIG.4. RISK VALUES OF SURPE ESTIMATORS (t=20)

to be tested some of which use SUR estimation to reduce variances. The SURPE procedure could be used to determine if SUR estimation is justified.

References

Breusch, T. S. and A. P. Pagan (1980). "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics," The Review of Economic Studies, 47, 239-254.

Grandshteyn, I. S. and I. M. Ryznik (1980). "Table of Integrals, Series and Products," New York: Academic Press.

Judge, G. G. and M. E. Bock (1978). "The Statistical Implications of Pretest and Stein-rule Estimators in Econometrics," Amsterdam: North-Holland.

Kariya, T. (1981). "Tests for the Independence Between Two Seemingly Unrelated Regression Equations," Annals of Statistics, 9, 381-390.

Kunitomo, N. (1977). "A Note on the Efficiency of Zellner's Estimator for the Case of Two Seemingly Unrelated Regression Equations," Economic Studies Quarterly, 28, 73-77.

Mehta, J. S. and P. A. V. B. Swamy (1976). "Further Evidence on the Relative Efficiencies of Zellner's Seemingly Unrelated Regressions Estimator," Journal of American Statistical Association, 71, 634-639.

Shiba, T. and H. Tsurumi (1988). "Bayesian and Non-Bayesian Tests of Independence in Seemingly Unrelated Regressions," International Economic Review, 20, 377-395.

Stanek, Edward J. (1988). "Choosing a Pretest-Posttest Analysis," The American Statistician, 42, 178-183.

Zellner, A. (1962). "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias," Journal of American Statistical Association, 57, 348-368.

__________ (1963). "Estimators of Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results," Journal of American Statistical Association, 58, 977-992.