Western University
Scholarship@Western

Department of Economics Research Reports | Economics Working Papers Archive

1971

A Measure of Correlation for Simultaneous Equation Systems

R. A.L. Carter

A. L. Nagar

Follow this and additional works at: https://ir.lib.uwo.ca/economicsresrpt

Part of the Economics Commons

Citation of this paper:

Carter, R. A.L., A. L. Nagar. "A Measure of Correlation for Simultaneous Equation Systems." Department of Economics Research Reports, 7133. London, ON: Department of Economics, University of Western Ontario (1971).

20304

RESEARCH REPORT 7133

A MEASURE OF CORRELATION FOR SIMULTANEOUS EQUATION SYSTEMS

by

R.A.L. Carter

and

A.L. Nagar

December, 1971

A Measure of Correlation for Simultaneous Equation Systems[1]

R. A. L. Carter and A. L. Nagar

I. Introduction

In single equation regression models the coefficient of multiple correlation (R^2) is used as a measure of the proportion of explained variation in the dependent variable. The sampling distribution of R^2 is well known and tests are available for testing

i) the significance of an observed multiple corrleation, and

ii) the significance of difference between two observed multiple correlations.

In the case of a simultaneous equations model the multiple correlation coefficient computed from each structural equation does not follow the same probability distribution as in the case of a single equation model. In fact the sampling distribution of the multiple correlation coefficient in the simultaneous equations case is not known and is perhaps very difficult to derive.

As shown by Hooper [5] one may employ the theory of canonical correlations (developed by Hotelling [6]) and use the vector correlation and vector alienation coefficients in the context of simultaneous equations. However, besides other difficulties pointed out by Hooper the sampling distribution of the vector correlation is not known. Hooper also proposed a "trace correlation" in this context, and he analyzed the asymptotic sampling variances of this index which may be used to apply some approximate tests of significance.

---

[1]We would like to express appreciation to the Canada Council and to I.B.M. for financial support and to T. H. Wonnacott for helpful discussions. Residual errors are ours alone.

The purpose of this note is to point out that Hooper's trace correlation does not make use of the covariance structure of the reduced form disturbances and in this sense it is not an efficient measure. Using the covariance structure of these disturbances we arrive at an alternative measure--the asymptotic distribution of which can be obtained in a straightforward manner.

2. Notation and Assumptions

We write the linear simultaneous equation system as

(1) $Y\beta + X\Gamma + U = 0$

where Y is a T x G matrix of T observations on G endogenous variables, X is a T x K matrix of T observations on K predetermined variables, β and Γ are G x G and K x G matrices respectively of unknown parameters and U is a T x G matrix of unobservable random disturbances. Y and X are measured in deviations from their sample means.

We make the following assumptions

(2) β is non-singular. Therefore, we can write the reduced form of (1) as

(3) $Y = X\Pi + V$, where $\Pi = \Gamma\beta^{-1}$ and $V = U\beta^{-1}$.

(4) The T rows of U are independent random drawings from a G variate normal distribution with means zero and a positive definite covariance matrix Σ i.e., $U \sim N(0,\Sigma)$.

(5) $\text{plim}(1/T)X'U = 0$.

(6) $\text{plim}(1/T)X'X = \Sigma_X$, a matrix of constants.

Assumption (4) implies

(7) $V \sim N(0,\Omega)$ with $\Omega = \beta'^{-1}\Sigma\beta^{-1}$.

(8) $\text{plim}(1/T)X'V = 0$.

(9) $Y \sim N(X\Pi,\Omega)$

3. The Coefficient of Correlation

We define the population coefficient of correlation for the whole model to be the positive square root of

$$(10) \qquad \rho_w^2 = 1 - \frac{tr(V\,\Omega^{-1} V')}{tr(Y\,\Omega^{-1} Y')} .$$

An estimate of ρ_w^2 is provided by

$$(11) \qquad R_w^2 = 1 - \frac{tr(\hat{V}\,\hat{\Omega}^{-1} \hat{V}')}{tr(Y\,\hat{\Omega}^{-1} Y')}$$

where V and Ω in (10) have been replaced by their consistent estimates. It is then clear that R_w^2 is a consistent estimator of ρ_w^2. ρ_w^2 and R_w^2 have the same interpretation as the familiar coefficient of determination for single equations. They show, for the population and the sample, respectively, what proportion of the total variation in all the endogenous variables is accounted for by the systematic variation in the reduced form. At one extreme, when there is no random variation at all, V (or $\hat{V}$) is zero and the measures go to their maximum value of one. At the other extreme, when all the variation in the reduced form is random $V = Y$ (or $\hat{V} = Y$)and the measures go to their lower limit of zero.

The measure defined in (10) is to be compared with those of Hotelling [6]

$$(12) \qquad q^2 = (-1)^G \frac{\begin{vmatrix} 0 & Y'X \\ X'Y & X'X \end{vmatrix}}{|Y'Y||X'X|}$$

and Afriat [1] and Hooper [5]

$$(13) \qquad \bar{\rho}^2 = \frac{1}{G}\, tr\left[I - (Y'Y)^{-1} V'V\right]$$

It is clear that the presence of Ω^{-1} in (10) and its estimator $\hat{\Omega}^{-1}$ in (11) gives our measure the advantage, over those in (12) and (13), of containing information on the covariances of the reduced form disturbances.

For purposes of computation R_w^2 can be rearranged to

$$(14) \qquad R_w^2 = 1 - \frac{tr(\hat{\Omega}^{-1}\ \hat{V}'\hat{V})}{tr(\hat{\Omega}^{-1}\ Y'Y)} = 1 - \frac{tr(T\ (\hat{V}'\hat{V})^{-1}\hat{V}'\hat{V})}{tr(T\ (\hat{V}'\hat{V})^{-1} Y'Y)} = 1 - \frac{G}{tr\left[(\hat{V}'\hat{V})^{-1} Y'Y\right]}$$

where $\Omega = (1/T)\hat{V}'\hat{V}$.

4. Transforming the Model

In order to derive the asymptotic distribution of R_W^2 it is useful to transform the model so that $\Omega = I$. As shown by Basmann [2] this can always be achieved, without loss of generality.

Since Σ (defined in assumption 4) is a positive definite matrix, we can always obtain a nonsingular square matrix Ψ such that

(15) $\Psi'\Psi = \Sigma^{-1}$

and then let us define

(16) $P = \beta\,\Psi'$.

The transformed model may be written as

(17) $Y_* \beta_* + X\Gamma + U = 0$ where $Y_* = YP$ and $\beta_* = P^{-1}\beta$.

The transformed reduced form is

(18) $Y_* = X\Pi_* + V_*$ where $\Pi_* = \Pi P$ and $V_* = VP$.

It follows that for the transformed model

(19) $\Omega_* = \frac{1}{T} E\, V_*'V_* = \frac{1}{T} E\, P'V'VP = P'\Omega P = P'\beta^{-1}\Sigma\beta'^{-1}P = \Psi\,\Sigma\,\Psi' = I$.

The derivation of the probability distribution of R_W^2 will be considerably simplified if we write the reduced form of the system in a slightly different form as below. First we define

$$(20)\quad y_* = \begin{bmatrix} y_{*1} \\ y_{*2} \\ \vdots \\ y_{*G} \end{bmatrix} ; \quad \pi_* = \begin{bmatrix} \pi_{*1} \\ \pi_{*2} \\ \vdots \\ \pi_{*G} \end{bmatrix} ; \quad v_* = \begin{bmatrix} v_{*1} \\ v_{*2} \\ \vdots \\ v_{*G} \end{bmatrix}$$

y_* and v_* are TG order column vectors made up of the stacked columns of Y_* and V_*. π_* is a KG order column vector of the stacked columns of Π_*

$$F = \begin{bmatrix} X & 0 \dots 0 \\ 0 & X & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \dots & X \end{bmatrix}$$

is a GT x GK matrix with X in the main diagonal blocks and zero elsewhere.

Now we write the reduced form as

(21) $y_* = F\,\pi_* + v_*$

Assumption (6) implies

(22) $\text{plim}(1/GT)F'F = \Sigma_F$, a matrix of constants.

From (7), (9) and (19), we see that

(23) $V_* \sim N(0, I_G)$

(24) $Y_* \sim N(X\Pi_*, I_G)$

so that

(25) $v_* \sim N(0, I_{GT})$

(26) $y_* \sim N(F\pi_*, I_{GT})$

Now we can consider the numerator and denominator of the second term on the right side of (10)

(27) $\text{tr}(V\,\Omega^{-1}V') = \text{tr}(\Omega^{-1}V'V) = \text{tr}(\beta\,\Sigma^{-1}\beta'V'V) = \text{tr}(\beta\,\Psi'\Psi\beta'V'V) = \text{tr}(P\,P'V'V)$

$= \text{tr}(P'V'VP) = \text{tr}\,V_*'V_* = v_*'v_*$

Similarly

(18) $\text{tr}(Y\,\Omega^{-1}Y') = \text{tr}(P'Y'YP) = \text{tr}\,Y_*'Y_* = y_*'y_*$

Therefore, we now rewrite (10) and (11) as

(29) $\rho_w^2 = 1 - \dfrac{v_*'\,v_*}{y_*'\,y_*}$

and

(30) $R_w^2 = 1 - \dfrac{\text{tr}(T\,I_G)}{\hat{y}_*'\,\hat{y}_*} = 1 - \dfrac{T\,G}{\hat{y}_*'\,\hat{y}_*}$

5. <u>The Asymptotic Distribution of R_w^2</u>

It is convenient here to consider the asymptotic distribution of:

(31) $$1 - R_w^2 = \frac{T\ G}{\hat{y}_*'\ \hat{y}_*}$$

If we restrict ourselves to consistent estimators, the asymptotic distribution of $\hat{y}_*$ will be the same as that of y_* as given in (26). Therefore, $\hat{y}_*'\hat{y}_*$ has the same asymptotic distribution as $y_*'y_*$, which is non-central χ^2 with GT degrees of freedom and a non-centrality parameter λ given by

(32) $$\lambda = \frac{1}{2}(E\ y_*)'E\ y_* = \frac{1}{2}\ \pi_*'F'F\pi_*$$

Then the probability limit of $1 - R_w^2$ is

(33) $$\text{plim}(1 - R_w^2) = \frac{GT}{GT + 2\lambda} = \frac{1}{1 + \pi_*'\ \Sigma_F\pi_*} = \text{plim}(1 - \rho_w^2)$$

and its asymptotic variance is (cf. [7])

(34) $$\bar{V}(1 - R_w^2) = \frac{GT}{2GT + 8\lambda} = \frac{1}{2 + 4\pi_*'\ \Sigma_F\pi_*} .$$

We can make probability statements about R_w^2 by using

(35) $$\Pr(y_*'y_* \le k) = \Pr(R_w^2 \le 1 - \frac{GT}{k})$$

It is useful to rewrite the non-centrality parameter λ in terms of the original model

(36) $$\lambda = \frac{1}{2}\ \text{tr}(EY_*'\ EY_*) = \frac{1}{2}\ \text{tr}(P'EY'EYP) = \frac{1}{2}\ \text{tr}(\Omega^{-1}EY'EY) = \frac{1}{2}\ \text{tr}(\Omega^{-1}\Pi'X'X\Pi)$$

An estimate of λ is provided by

(37) $$\hat{\lambda} = \frac{T}{2}\ \text{tr}[(\hat{V}'\hat{V})^{-1}\hat{\Pi}'X'X\hat{\Pi}]$$

6. Applications

In practical work with small samples equations (33) to (35) are approximations which can be expected to improve as the sample size increases.

Consider, as an example, Tintner's model of the American meat market [10]. We will follow Hooper's procedure and use least squares on the reduced form to obtain:

(38) $(\hat{V}'\hat{V})^{-1} = \begin{bmatrix} .00484 & .00422 \\ .00422 & .00522 \end{bmatrix}$

(39) $R_w^2 = 1 - \frac{2}{11.9} = .832$

(40) $\hat{y}_*'\hat{y}_* = \chi'^2 = 262$

(41) $f = 44$ degrees of freedom

(42) $\hat{\lambda} = 109$

Using Pearson's approximation [8] we see that approximately 95% of a non-central χ^2 distribution with the λ and f given above lies left of $\chi'^2 = 193$ and approximately 99% lies left of $\chi'^2 = 211$. These points translate into values of $R_w^2 = .772$ and $R_w^2 = .791$ respectively. Therefore, our computed value of $R_w^2 = .832$ is significant at both the 5% and 1% significance levels.

A second example is provided by Klein's Model I [4]. Using the moment matrices and two stage least squares coefficient estimates supplied by Goldberger [3] we obtain a value for Hooper's coefficient of

(43) $\bar{R}^2 = 869$

In contrast, our coefficient is

(44) $R_w^2 = .999$

(45) $\hat{y}_*'\hat{y}_* = \chi'^2 = 60000$

(46) $f = 60$ degrees of freedom

(47) $\hat{\lambda} = 33456$

Because the value of $\hat{\lambda}$ is so high in this case we must use, in addition to Pearson's approximation given above, the approximation given by Pearson and Hartley ([9], p. 137) to obtain the critical values of χ'^2. In this case we see that approximately 95% of the distribution lies left of $\chi'^2 = 34061$ and approximately 99% lies left of $\chi'^2 = 34438$. These points correspond to values of $R_w^2 = .99824$ and $R_w^2 = .99826$ respectively. Once again the computed $R_w^2 = .999$ is significant at both 5% and 1%.

References

[1] Afriat, S. N.: "Orthogonal and Oblique Projectors and the Characteristics of Vector Spaces," Proceedings of the Cambridge Philosophical Society 53 (1957), pp. 800-816.

[2] Basmann, R. L.: "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in a Leading Three-Equation Case," Journal of the American Statistical Association 58 (1963).

[3] Goldberger, A. S.: Econometric Theory, New York: John Wiley and Sons, 1964.

[4] Klein, L. R.: Economic Fluctuations in the United States 1924-41, Cowles Commission Monograph 11, New York: John Wiley and Sons, 1950.

[5] Hooper, J. W.: "Simultaneous Equations and Canonical Correlation Theory," Econometrica, 27 (1959), pp. 245-256.

[6] Hotelling, H.: "Relations Between Two Sets of Variates," Biometrika 28 (1936), pp. 321-377.

[7] Lancaster, H.O.: The Chi-Squared Distribution, New York: John Wiley and Sons, 1969.

[8] Pearson, E. S.: "Note on an Approximation to the Distribution of Non-Central χ^2," Biometrika 46 (1959), p. 364.

[9] Pearson, E. S. and H. O. Hartley: Biometrika Tables for Statisticians, Cambridge: Cambridge University Press, 1958.

[10] Tintner, G.: Econometrics, New York: John Wiley and Sons, 1952.