A NOTE ON EXACT TESTS FOR SERIAL CORRELATION

by

Christopher A. Sims

Discussion Paper No. 26, March 1973

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

A NOTE ON EXACT TESTS FOR SERIAL CORRELATION

by

Christopher A. Sims

Introduction

Durbin (1970) and Abrahamse and Koerts (1971) have proposed linear transforms of the OLS residual vector which give it a convenient covariance matrix. Though both proposed transforms achieve the same covariance matrix on the null hypothesis,[1] they differ in important respects.

[1] Both techniques allow specification of an arbitrary idempotent covariance matrix for the transformed residuals.

Durbin presents his transform in easily computable form, exploiting standard statistics printed out in OLS computer programs, and his formulas involve matrix inversions of order no greater than 2k, where k is the number of independent variables. On the other hand, Durbin makes no claims related to the power of tests based on his transformed residuals, though it is easy to see that his transformed residual vector is, like the original OLS residual vector, elementwise consistent as sample size expands.

Abrahamse and Koerts (A-K) optimize the choice of linear transform of the OLS residuals, over a class of transforms including Durbin's, minimizing

$$Q = E[(u-B'\hat{u})'(u-B'\hat{u})],$$

where u is the true residual vector, $\hat{u}$ is the OLS estimate, and $B'\hat{u}$ is the transformed vector. Minimizing Q under the null hypothesis is not the same thing as maximizing power of tests based on $B'\hat{u}$ by an means; but for alternative hypotheses near H_o, minimizing Q and maximizing power are probably closely related. On the other hand, A-K's formula involves taking the square root of an $(n-k)\times(n-k)$ matrix, where n is sample size, and therefore appears much less easy to implement than Durbin's formula.

In this note we will see that a) Durbin's transform behaves noticeably worse than A-K's in a certain important class of cases and b) Durbin's computational ideas can be applied to make A-K's transform nearly as simple as Durbin's own.

1. A Special Case

Suppose we have the regression

1) $\underset{n\times 1}{y} = \underset{n\times 1}{X_1}\,\beta + \underset{n\times 1}{u}$.

Let $X_1 = aZ_1 + Z_2$,

where Z_1 and Z_2 are orthonormal vectors, and suppose we wish to transform the OLS residual vector $\hat{u}$ so it will have the same covariance matrix as residuals from an OLS regression of u on Z_2. The following paradoxes emerge from examining Durbin's procedure: A) Durbin's transform is different if we substitute $Z_2^* = -Z_2$ for Z_2, so that $X_1 = aZ_1 - Z_2^*$, despite the fact that Z_2 itself is a "dummy" vector whose sign does not affect the problem; B) In the problem as first formulated, with Z_2, Durbin's transformed residuals do not approach $\hat{u}$ as $a \to 0$. Verification of these points is left to the reader. Neither of these two paradoxes arises for the A-K transform.

Paradox (B) is disturbing, since it is often argued[/]

[/] I am not sure this argument is of much practical relevance, myself.

that in econometric practice we frequently meet cases where independent variables are close in mean square to low-order trigonometric polynomials. Generalizing the example to $k > 1$, this suggests that small a's are likely in practice. It is known for this case that $\hat{u}$ itself has nearly the convenient covariance matrix the transforms aim to achieve. A-K exploits this fact and leaves $\hat{u}$ nearly unchanged, while Durbin's transform may substantially alter $\hat{u}$, probably noticeably reducing power in the process.

2. Computing the Abrahamse and Koerts Transform

In the A-K notation, our problem is to find an $n\times n$ matrix B such that

2) $B'B = \Omega$,

3) $B'X_1 = 0$,

and tr B is a maximum.

In Durbin's notation, $\Omega = I - X_2(X_2'X_2)^{-1} X_2'$,[/] X_2 being $n\times k$.

[/] We ignore for the time being Durbin's X_3, which consists of variables included in the original regression and also in the dummy regression of u on "convenient" variables.

A-K begin by choosing P an arbitrary $n\times(n-k)$ matrix of orthonormal eigenvectors of $M = I - X_1(X_1'X_1)^{-1} X_1'$ such that $PP' = M$, while K is an arbitrary $n\times(n-k)$ matrix of orthonormal eigenvectors of $\Omega = I - X_2(X_2'X_2)^{-1} X_2'$ such that $KK' = \Omega$. They then show that any B satisfying the restrictions (2) and (3) has the form $B = PHK'$, where $HH' = I$. They then choose H optimally.

The difficulty of A-K's suggested computations comes from choosing K and P too arbitrarily. Let P_3 be an $n\times(n-2k)$ matrix of orthonormal vectors spanning the space orthogonal to both X_1 and X_2. Let $P_{1.2}$ span the subspace of the space spanned by X_1 and X_2 which is orthogonal to X_2. Let $P_{2.1}$ be defined analogously. Then we can take $P = [P_{2.1}, P_3]$, $K = [P_{1.2}, P_3]$.

Applying A-K's procedure, we now find that H has the form

$$H = \begin{bmatrix} H_k & 0 \\ 0 & I \end{bmatrix}$$

where H_k is $k\times k$. Using A-K's method for choosing H_k and applying Durbin's computational ideas, we write

$$4) \quad B'y = \hat{u}_{12} + X_{1.2}\, G_{12}(P_2^{-1})'(P_2^{-1}\, G_{21}\, G_1^{-1}\, G_{12}(P_2^{-1})')^{-\frac{1}{2}}$$

$$P_2'\, X_{2.1}'\, y.$$

Here $\hat{u}_{12}$ is the residual vector from OLS regression of y on X_1 and X_2 together. Following Durbin's notation, $X_{i.j}$ is that part of X_i not explained by linear regression on X_j, $G_1 = (X_{1.2}'X_{1.2})^{-1}$, $G_2 = (X_{2.1}'X_{2.1})^{-1}$, $G_{12} = G_{21}' = G_1\, X_{1.2}'X_{2.1}\, G_2$, and P_2 is any square matrix satisfying $P_2P_2' = G_2$. The exponent $\frac{1}{2}$ for a positive definite matrix $Q = WDW'$, where D is diagonal and $WW' = I$, is defined by $Q^{\frac{1}{2}} = WD^{\frac{1}{2}}W'$, where $D^{\frac{1}{2}}$ is the element-wise square root of D. G_1 and G_2 are, as Durbin pointed out, scalar multiples of the covariance matrices of c_1 and c_2, respectively, where c_1 and c_2 are the coefficients of X_1 and X_2 in the OLS estimate of the joint regression of y on these two variables. G_{12}, which does not enter Durbin's formula, is proportional to the matrix of covariances between c_1 and c_2. Finally, explicit computation of the vector $X_{2.1}$ can be avoided by noting that the term $P_2' X_{2.1}' y$, appearing at the end of (4), can be rewritten in terms of c_2, to yield

$$5) \quad B'y = \hat{u}_{12} + X_{1.2}\, G_{12}(P_2^{-1})'(P_2^{-1}\, G_{21}\, G_1^{-1}\, G_{12}(P_2^{-1})')^{-\frac{1}{2}} \, P_2^{-1}\, c_2 \, .$$

Formula (5) requires taking one symmetric square root of a k×k matrix and inversion of the k×k matrix G_1, while avoiding the taking of a triangular square root of G_1 as is required for Durbin's transform. This increased computational requirement is modest, and for small k it is negligible. For $k = 1$ it is only a matter of finding the right sign for X_2.

What if there is a set of variables X_3 included in both the original regression matrix and in the dummy regression? In this case P_3 could not be chosen as $n\times(n-2k)$. However, appropriate modification of formula (5) is simple. Now X_1 becomes the matrix of variables appearing only in the original regression, X_2 the matrix of variables appearing only in the dummy regression. The auxiliary regression from which G_1, G_2, G_{12}, c_2, and $\hat{u}_{12}$ are drawn is now a joint regression on X_1, X_2, and X_3, and $X_{1.2}$ is replaced by $X_{1.23}$.

REFERENCES

Durbin, J. (1970), "An Alternative to the Bounds Test for Serial Correlation in Least-Squares Regression", _Econometrica_, _38_, 422-429.

Abrahamse, A.P.J. and J. Koerts (1971), "New Estimators of Disturbances in Regression Analysis", _JASA_, _66_, 71-74.