

Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions

Author(s): Robert E. Cumby and John Huizinga

Source: *Econometrica*, Vol. 60, No. 1 (Jan., 1992), pp. 185-195

Published by: [The Econometric Society](#)

Stable URL: <http://www.jstor.org/stable/2951684>

Accessed: 11/10/2011 06:22

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The Econometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Econometrica*.

<http://www.jstor.org>

TESTING THE AUTOCORRELATION STRUCTURE  
OF DISTURBANCES IN ORDINARY LEAST  
SQUARES AND INSTRUMENTAL  
VARIABLES REGRESSIONS

BY ROBERT E. CUMBY AND JOHN HUIZINGA<sup>1</sup>

1. INTRODUCTION

THIS PAPER DERIVES the asymptotic distribution of a vector of sample autocorrelations of regression residuals from a quite general linear regression model. The model is allowed to have a regression error that is a moving average of order  $q \geq 0$  with possibly conditionally heteroscedastic innovations; to have strictly exogenous, predetermined, and/or endogenous regressors; and to be estimated by a variety of Generalized Method of Moments estimators, such as ordinary least squares, two-stage least squares, or two-step two-stage least squares.<sup>2</sup>

One important use of the distribution derived here is to form the basis for a simple test of the null hypothesis that the regression error is a moving average of known order  $q \geq 0$  against the general alternative that autocorrelations of the regression error are nonzero at lags greater than  $q$ . The test—denoted the  $l$  test—is thus general enough to test the null hypothesis that the regression error has no serial correlation ( $q = 0$ ) or the null hypothesis that serial correlation in the regression error exists, but dies out at a known finite lag ( $q > 0$ ).

The  $l$  test is especially attractive because it can be used in at least three frequently-encountered situations where such popular tests as the Box-Pierce (1970) test, Durbin's (1970)  $h$  test, and the Lagrange multiplier test described by Godfrey (1978b) either are not applicable or are costly to compute.

The first situation is when the regression contains endogenous variables. The three popular tests listed above are not valid when the regression has been estimated by instrumental variables, and the Box-Pierce test is further restricted to having only lagged dependent variables.<sup>3</sup> In contrast, the  $l$  test can be used not only with ordinary least squares but also with a wide class of instrumental variables estimators.

A second situation,  $q > 0$ , arises in studies of asset returns over holding periods which differ from the observation interval and in studies where time aggregated data are used.<sup>4</sup> In this situation, existing tests that investigate the serial correlation of the regression error require estimating the parameters of the moving average error process, and

<sup>1</sup> We would like to thank seminar participants at the University of Chicago and Lars Hansen, in particular, as well as two anonymous referees for helpful comments. Huizinga gratefully acknowledges funding from the University of Chicago Graduate School of Business and the Sloan Foundation.

<sup>2</sup> See Hansen (1982) for a description of Generalized Method of Moments estimators. Cumby, Huizinga, and Obstfeld (1983) describe the two-step two-stage least squares estimator.

<sup>3</sup> Godfrey (1978a) describes a test that is valid with some instrumental variables estimators, but the test is not valid in the presence of conditionally heteroscedastic errors or with instrumental variables estimators such as two-step two-stage least squares. The test, like Durbin's  $h$  test, is also restricted to testing the significance of the first autocorrelation of the regression error.

<sup>4</sup> See, for example, work on returns in the foreign exchange market by Hansen and Hodrick (1980), the study of real interest rates by Huizinga and Mishkin (1984), the investigation of stock returns by Fama and French (1988), and work on the term structure of interest rates by Mishkin (1990). Hall (1988), Hansen and Singleton (1990), and Christiano, Eichenbaum, and Marshall (1991) address the issue of time aggregated data.

therefore necessitate nonlinear estimation.<sup>5</sup> In contrast, the  $l$  test described in this paper avoids the use of nonlinear estimation because it is based solely on the sample autocorrelations of regression residuals and a consistent measure of their asymptotic covariance matrix. The  $l$  test thus reflects a desire for simplicity, and for ensuring that regression diagnostics do not become more costly or more difficult to compute than the original regression.

The third situation is conditional heteroscedasticity of the error term, a situation that is frequently detected in empirical studies.<sup>6</sup> As discussed below, the presence of conditional heteroscedasticity can seriously undermine tests for serial correlation of regression errors that ignore its presence. The  $l$  test can be used with either conditionally heteroscedastic or homoscedastic errors.

The outline of the paper is as follows. In Section 2, we derive the asymptotic distribution of the sample autocorrelations at lags  $q+1$  to  $q+s$  of regression residuals from a model where the regression errors are a  $q$ th order moving average with possibly conditionally heteroscedastic innovations. The regression is assumed to be estimated by instrumental variables, with instruments that are predetermined, but not necessarily strictly exogenous. We note how the distribution simplifies when the regression errors are conditionally homoscedastic and when all regressors are predetermined or strictly exogenous variables so that ordinary least squares is appropriate. Based on this asymptotic distribution of the sample autocorrelations of regression residuals, a test of the hypothesis that the true regression errors are a  $q$ th order moving average process is presented in Section 3. Section 4 contains summary remarks.

## 2. DISTRIBUTION OF SAMPLE AUTOCORRELATIONS OF REGRESSION RESIDUALS

The regression equation to be considered in this paper is

$$(1) \quad y_t = X_t \delta + \varepsilon_t \quad (t = 1, \dots, T),$$

where  $y_t$  and  $\varepsilon_t$  are scalar random variables,  $X_t$  is a  $1 \times k$  vector of the  $k$  scalar random variables  $X_{1,t}, X_{2,t}, \dots, X_{k,t}$ , and  $\delta$  is a  $k \times 1$  vector of unknown parameters. The vector of regressors,  $X_t$ , may include jointly endogenous variables (those contemporaneously correlated with  $\varepsilon_t$ ), predetermined variables (those uncorrelated with  $\varepsilon_{t+j}$  for  $j \geq 0$  but correlated with  $\varepsilon_{t-j}$  for some  $j > 0$ ), or strictly exogenous variables (those uncorrelated with  $\varepsilon_{t+j}$  for all  $j$ ).

The regression errors  $\varepsilon_t$  are assumed to have mean zero and satisfy two other conditions. First, though they are allowed to be conditionally heteroscedastic, they are assumed to be unconditionally homoscedastic. Second, for a known  $q \geq 0$ , their autocorrelations at all lags greater than  $q$  are required to be zero.

It is also assumed that there exists a  $1 \times h$  vector of instrumental variables  $Z_t$ , comprised of  $h \geq k$  scalar random variables  $Z_{1,t}, Z_{2,t}, \dots, Z_{h,t}$ , each of which is uncorrelated with  $\varepsilon_t$ .  $Z_t$  is required to be predetermined, but not necessarily strictly exogenous,

<sup>5</sup> This is true of the Box-Pierce test, the likelihood ratio test, and, as discussed in Godfrey (1978c), the Lagrange multiplier test. It is also true of a GMM approach that jointly estimates the parameters of primary interest and the residual autocorrelations.

<sup>6</sup> Diebold (1986) proposes a test for serially correlated errors in the presence of the ARCH distribution described by Engle (1982). The  $l$  test discussed here is more generally applicable than Diebold's test because it does not assume any particular functional form for the conditional heteroscedasticity.

with respect to  $\varepsilon_t$ . These assumptions are summarized by

$$(2) \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma_\varepsilon^2,$$

$$(3) \quad E(\varepsilon_t \varepsilon_{t-n}) / \sigma_\varepsilon^2 = \rho_n,$$

and

$$(4) \quad E(\varepsilon_t | Z_t, Z_{t-1}, \dots, \varepsilon_{t-q-1}, \varepsilon_{t-q-2}, \dots) = 0.$$

Furthermore, the  $h \times h$  matrix

$$(5) \quad \Omega = \lim_{T \rightarrow \infty} (1/T) E(Z' \varepsilon \varepsilon' Z)$$

is assumed to exist and be of full rank.

It is assumed that the  $k \times 1$  parameter vector  $\delta$  in equation (1) has been estimated using a root- $T$  consistent estimator of the form

$$(6) \quad d = (X' Z A_T^{-1} Z' X)^{-1} X' Z A_T^{-1} Z' y$$

for some observable matrix  $A_T$ . This formulation is general enough to include ordinary least squares ( $A_T = (X'X/T)$  and  $Z = X$ ), two-stage least squares ( $A_T = (Z'Z/T)$ ), and two-step two-stage least squares ( $A_T$  is proportional to a consistent estimate of  $\Omega$ ). The asymptotic covariance matrix for the estimator  $d$  is denoted  $V_d$ , with

$$(7) \quad V_d = D \Omega D'$$

and the  $k \times h$  matrix  $D$  given by

$$(8) \quad D = \text{plim } T (X' Z A_T^{-1} Z' X)^{-1} X' Z A_T^{-1}.$$

The objective of this section of the paper is to derive, within the framework of the model described by equations (1)–(8), the asymptotic covariance matrix of the sample autocorrelations of the regression residuals,  $\hat{\varepsilon}_t = y_t - X_t d$ . In the following section we show how a consistent estimate of this covariance matrix can be used to test the hypothesis that the  $s \times 1$  vector  $\rho = (\rho_{q+1}, \dots, \rho_{q+s})' = 0$ .

Let  $\hat{r} = [\hat{r}_{q+1}, \hat{r}_{q+2}, \dots, \hat{r}_{q+s}]'$  and

$$(9) \quad \hat{r}_n = \frac{\sum_{t=n+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-n}}{\sum_{t=1}^T \hat{\varepsilon}_t^2}.$$

By the mean value theorem,

$$(10) \quad \sqrt{T} \hat{r} = \sqrt{T} r + \frac{\partial r}{\partial \delta} \sqrt{T} (d - \delta),$$

where the  $s \times 1$  vector  $r = [r_{q+1}, r_{q+2}, \dots, r_{q+s}]'$ ,

$$(11) \quad r_n = \frac{\sum_{t=n+1}^T \varepsilon_t \varepsilon_{t-n}}{\sum_{t=1}^T \varepsilon_t^2},$$

and the  $j$ th row of the  $s \times k$  matrix  $\partial r / \partial \delta$  is evaluated at  $d_j^*$ , which lies between  $d$  and  $\delta$ . Equation (10) shows that the asymptotic covariance matrix of  $\hat{r}$  (the vector of sample autocorrelations of the regression residuals) can be derived as the asymptotic covariance matrix for the sum of  $r$  (the vector of sample autocorrelations of the true disturbances) and  $\partial r / \partial \delta (d - \delta)$ . Only when  $\partial r / \partial \delta$  can safely be ignored will the sampling variation in the estimation of  $\delta$  not affect the sampling variation in the estimation of  $\rho$ .

Let the  $s \times k$  matrix  $B$  have  $i, j$ th element,

$$(12) \quad B(i, j) = - \left[ E(\varepsilon_{t-q-i} X_{j,t}) + E(\varepsilon_t X_{j,t-q-i}) \right] / E(\varepsilon_t^2).$$

We show in the Appendix that  $B = \text{plim } \partial r / \partial \delta$  and thus that  $BV_d B'$  is the asymptotic covariance matrix of  $\partial r / \partial \delta (d - \delta)$ . In most models, the implication of equation (4) that  $E(\varepsilon_t | \varepsilon_{t-q-1}, \varepsilon_{t-q-2}, \dots) = 0$  will be sufficient to ensure  $X_{j,t-q-i}$  is predetermined with respect to  $\varepsilon_t$  and thus that the second term of the sum in equation (12) is zero.

To complete the notation, let  $\xi_{i,t} = \varepsilon_t \varepsilon_{t-q-i}$  for  $i = 1, \dots, s$ ,  $\omega_{j,t} = \varepsilon_t Z_{j,t}$  for  $j = 1, \dots, h$ , the  $ij$ th element of the  $s \times s$  matrix  $V_r$  be given by

$$(13) \quad V_r(i, j) = \sigma_\varepsilon^{-4} \sum_{n=-q}^q E(\xi_{i,t} \xi_{j,t-n}),$$

and the  $ij$ th element of the  $s \times h$  matrix  $C$  be given by

$$(14) \quad C(i, j) = \sigma_\varepsilon^{-2} \sum_{n=-q}^q E(\xi_{i,t} \omega_{j,t-n}).$$

In the Appendix we show that  $V_r$  is the asymptotic covariance matrix of  $r$  and that the asymptotic covariance matrix of  $r$  with  $\partial r / \partial \delta (d - \delta)$  is  $BDC'$ . Proposition 1 combines these findings in giving the key result of this section.<sup>7</sup>

**PROPOSITION 1:** *Given equations (1) through (14) and the regularity conditions stated in the Appendix,  $\sqrt{T} \hat{r} \overset{A}{\sim} N(0, V_\rho)$ , where  $V_\rho = V_r + BV_d B' + CD' B' + BDC'$ .*

Proposition 1 states that, in general, having to estimate the residuals will affect the asymptotic distribution of their sample autocorrelations. The following special cases of the general model provide further insight into Proposition 1 and help clarify the relationship between tests based on the asymptotic distribution of  $\hat{r}$  and tests of residual autocorrelation proposed elsewhere in the literature.

*Case (i): Strictly Exogenous Regressors.* Since  $B = 0$  when the regressors are strictly exogenous,  $V_\rho = V_r$  and one can safely ignore the fact that the true residuals are unavailable.

*Case (ii): Conditionally Homoscedastic Residuals.* We show in the Appendix that when the residuals are conditionally homoscedastic,  $V_r$  and  $C$  can be rewritten as

$$(15) \quad V_r(i, j) = \sum_{n=-q}^q \rho_{n-i+j} \rho_n$$

and

$$(16) \quad C(i, j) = \sum_{n=-q}^q \rho_n E(\varepsilon_{t-q-i} Z_{j,t-n}).$$

<sup>7</sup> The proof of Proposition 1 can be found in the Appendix.

The well known result that the sample autocorrelations of a serially uncorrelated series are independent and asymptotically normal with variance  $1/T$  follows from (15) with  $q = 0$ . When  $q > 0$  the sample autocorrelations are not independent and, though asymptotically normal, do not have variance  $1/T$ .

*Case (iii):* Conditionally Homoscedastic Residuals, Predetermined Regressors, and  $q = 0$ . When the regressors are predetermined, ordinary least squares yields consistent estimates of  $\delta$ ; we can set  $Z = X$ ,  $A_T = X'X/T$ , and the second term of  $B$  will be zero. Combining this with the assumption of conditional homoscedasticity (so that equation (16) is valid) and  $q = 0$  (so that  $\rho_n = 0$  for  $n \neq 0$ ) yields  $C = -\sigma_\varepsilon^2 B$ . Furthermore,  $V_d = \sigma_\varepsilon^2 \text{plim}(X'X/T)^{-1} = \sigma_\varepsilon^2 D$  so that  $BDC' = -BV_dB'$ . Finally, it follows from equation (15) that in this case  $V_r = I$ , and thus  $V_f = I - BV_dB'$ .

Unlike the case of strictly exogenous regressors, when regressors are merely predetermined one cannot safely ignore the use of regression residuals rather than the true disturbances in estimating autocorrelations. The expression  $V_f = I - BV_dB'$  can be used to derive the well-known Durbin (1970)  $h$  test. Durbin (1970) considers testing whether the autocorrelation of the error term at lag one is zero in a model with lagged dependent variables and strictly exogenous variables as regressors. In this case  $B$  will contain all zeros except a single value of minus one in the position corresponding to the dependent variable lagged once. Using  $V_{d1}$  to denote the estimated variance of the coefficient on this variable, the asymptotic variance of the first autocorrelation of the regression residuals is seen to be  $1/T - V_{d1}$ , which matches the formula given by Durbin (1970).<sup>8</sup>

*Case (iv):* Conditionally Homoscedastic Residuals, Only Lagged Dependent Variables, and  $q = 0$ . When the regression error is conditionally homoscedastic,  $X_t$  contains only  $k$  lagged values of  $y_t$ , and  $q = 0$ , we have a special case of the model considered by Box and Pierce (1970), who propose testing the hypothesis of zero correlation in the regression error by comparing  $Q_s = T\hat{r}'\hat{r}$  to the critical value of a chi-squared random variable with  $s - k$  degrees of freedom. Understanding the logic behind the Box-Pierce test and why the test in general fails when regressors other than lagged dependent variables are present becomes quite simple using the result from case (iii) that  $V_f = I - BV_dB'$ .

Specifically, it can be shown that when  $X_t$  contains only lagged values of  $y_t$ ,  $V_d$  approaches  $(B'B)^{-1}$  as  $s$  increases. It follows that as  $s$  increases,  $V_f$  approaches  $I - B(B'B)^{-1}B'$ , an idempotent matrix of rank  $s - k$ . Hence, for both large  $s$  and large  $T$ ,  $Q_s$  will be approximately distributed as a chi-square with  $s - k$  degrees of freedom.<sup>9,10</sup> If, however,  $X_t$  contains any variables other than lagged dependent variables,  $V_d$  will not in general approach  $(B'B)^{-1}$  and it is unlikely, though not impossible, that  $I - BV_dB'$  will be an idempotent matrix.

### 3. TESTING RESIDUAL AUTOCORRELATIONS EQUAL TO ZERO

The results presented in Section 2 can be used to develop a Wald test of the null hypothesis that the regression error in equation (1) is uncorrelated with itself at lags

<sup>8</sup> Godfrey (1978b) also considers the case of lagged endogenous and/or strictly exogenous regressors, conditionally homoscedastic errors, and  $q = 0$ . Among other things, he extends Durbin (1970) by showing that the asymptotic covariance matrix for a vector of sample autocorrelations of regression residuals is  $I - BV_dB'$ , the formula derived above.

<sup>9</sup> If  $W$  is an  $n \times 1$  random normal vector with mean 0 and  $n \times n$  covariance matrix  $V$  whose trace is nonzero, then  $W'W$  is distributed as a chi-square random variable with  $n - m$  degrees of freedom if and only if  $V$  is idempotent and has rank  $n - m$ . See Johnson and Kotz (1970, pages 177–178).

<sup>10</sup> Ljung (1986) investigates how large  $s$  must be before the  $Q_s$  statistic approaches the chi-square distribution. She finds that in samples of 50 or 100 observations,  $s \geq 10$  is sufficient for all AR(1) models examined and that  $s \geq 2$  is sufficient for AR(1) models with the autoregressive parameter below .9.

$q + 1$  through  $q + s$ .<sup>11</sup> Proposition 2 presents this result.

**PROPOSITION 2:** *Let  $\hat{V}_r$ ,  $\hat{B}$ ,  $\hat{C}$ ,  $\hat{D}$ , and  $\hat{V}_d$  be consistent estimates of  $V_r$ ,  $B$ ,  $C$ ,  $D$ , and  $V_d$ . Then, given the conditions of Proposition 1,*

$$l_{q,s} = T\hat{r}' \left[ \hat{V}_r + \hat{B}\hat{V}_d\hat{B}' + \hat{C}\hat{D}'\hat{B}' + \hat{B}\hat{D}\hat{C}' \right]^{-1} \hat{r} \stackrel{A}{\sim} \chi^2(s).$$

Proposition 2 states that if  $V_r$ ,  $B$ ,  $C$ ,  $D$ , and  $V_d$  can be estimated consistently, then the  $l_{q,s}$  statistic will be asymptotically distributed as a chi-square random variable with  $s$  degrees of freedom.<sup>12</sup> In the remaining part of this section we discuss how consistent estimates of  $V_r$ ,  $B$ ,  $C$ ,  $D$ , and  $V_d$  can be formed.

Define the  $(h + s) \times 1$  vector  $\eta_t$  by

$$(17) \quad \eta_t = (\omega_{1,t}, \dots, \omega_{h,t}, \xi_{1,t}, \dots, \xi_{s,t})'.$$

Then the  $(h + s) \times (h + s)$  spectral density matrix at frequency zero of  $\eta_t$  is proportional to

$$(18) \quad \Psi = \begin{pmatrix} \Omega & \sigma_\varepsilon^2 C' \\ \sigma_\varepsilon^2 C & V_r \sigma_\varepsilon^4 \end{pmatrix}.$$

Next, define the  $(2s) \times (s + h)$  matrix  $\Phi$  by

$$(19) \quad \Phi = \begin{pmatrix} BD & 0 \\ 0 & \sigma_\varepsilon^{-2} I \end{pmatrix},$$

so that

$$(20) \quad \Phi \Psi \Phi' = \begin{pmatrix} BV_d B' & BDC' \\ CD' B' & V_r \end{pmatrix}.$$

It follows from equation (20) that a consistent estimate of the asymptotic covariance matrix of  $\hat{r}$  can be obtained from consistent estimates of  $\Phi$  and  $\Psi$ . It also follows that if the consistent estimate of  $\Psi$  is positive definite in the sample, the resulting  $l_{q,s}$  will be positive.

<sup>11</sup> In many instances, instrumental variables are chosen as lagged endogenous variables so that rejecting the null hypothesis may call into question the validity of equation (4). In such cases it may be preferable to think of the null hypothesis being tested as a joint hypothesis concerning the serial correlation of the residuals and the validity of the instruments. Viewed in this way, the test described in this paper becomes an alternative to the  $J$  statistic proposed in Hansen (1982).

<sup>12</sup> Godfrey (1978b) considers a model with lagged endogenous and strictly exogenous regressors, conditionally homoscedastic errors, and  $q = 0$ . He shows that using  $\hat{r}$  to test  $\rho = 0$  is equivalent to the Lagrange multiplier test of the null hypothesis that the error term is serially uncorrelated against the alternatives that the error is MA( $s$ ) or AR( $s$ ) for  $s > 0$ . Hence, in some models, the test described in Proposition 2 is asymptotically equivalent to a likelihood ratio test. However, Godfrey (1978c) shows that in the same model but with  $q > 0$ , computation of the Lagrange multiplier test of the null hypothesis that the error term is MA( $q$ ) against the alternatives that the error is MA( $q + s$ ) or AR( $q + s$ ) requires that the moving average parameters be estimated. In this model the test described in Proposition 2 may not possess all the desirable properties of a Lagrange multiplier or likelihood ratio test, but will be less computationally burdensome than those tests. A procedure that would not require full maximum likelihood estimation of the moving average parameters is to implement a  $C(\alpha)$  test. Such a test would be asymptotically equivalent to the likelihood ratio and Lagrange multiplier tests and would only require that the derivatives of the likelihood function be evaluated at initial consistent estimates. See Godfrey (1989, pp. 27–28).

Consistently estimating  $\Phi$  is straightforward. Let  $\hat{U}$  be the  $T \times s$  matrix,

$$(21) \quad \hat{U}(i, j) = \hat{\varepsilon}_{i-j-q} \quad \text{for} \quad i-j-q > 0, \\ = 0 \quad \text{otherwise,}$$

so that the  $j$ th column of  $\hat{U}$  is the vector of regression residuals lagged  $q+j$  times. Then,

$$(22) \quad \hat{\sigma}_\varepsilon^2 = (1/T) \sum_{t=1}^T \hat{\varepsilon}_t^2,$$

$$(23) \quad \hat{B} = -(\hat{U}'X/T)/\hat{\sigma}_\varepsilon^2,$$

and

$$(24) \quad \hat{D} = T(X'ZA_T^{-1}Z'X)^{-1}X'ZA_T^{-1}$$

are consistent estimates of  $\sigma_\varepsilon^2$ ,  $B$ , and  $D$  respectively.

Consistently estimating  $\Psi$  is also straightforward. Let the  $(s+h) \times 1$  vector  $\hat{\eta}_t$  be given by

$$(25) \quad \hat{\eta}_t = (\hat{\varepsilon}_t Z_{1,t}, \dots, \hat{\varepsilon}_t Z_{h,t}, \hat{\varepsilon}_t \hat{\varepsilon}_{t-q-1}, \dots, \hat{\varepsilon}_t \hat{\varepsilon}_{t-q-s})',$$

and the  $(s+h) \times (s+h)$  matrix  $R_n$  be given by

$$(26) \quad R_n = (1/T) \sum_{t=n+1}^T \hat{\eta}_t \hat{\eta}_{t-n}'.$$

Then, as described in Anderson (1971), there are a variety of  $(N+1) \times 1$  weighting vectors  $w^N = (w_0^N, \dots, w_N^N)'$  such that the  $(s+h) \times (s+h)$  matrix

$$(27) \quad \hat{\Psi} = \sum_{n=-N}^N w_{|n|}^N R_n$$

is a consistent estimate of  $\Psi$ . Not all choices of  $w^N$  that give a consistent estimate of  $\Psi$  will also give an estimate that is positive definite in small samples, however.

Equations (15) and (16) in the previous section showed how the matrices  $V_r$  and  $C$  could be simplified in the case of homoscedastic errors.<sup>13</sup> With conditionally homoscedastic errors, the  $s \times s$  matrix

$$(28) \quad \hat{V}_r(i, j) = \sum_{n=-q}^q \hat{r}_{n-i+j} \hat{r}_n$$

is a consistent estimate of  $V_r$ , where  $\hat{r}_j = 0$  for  $|j| > q$ . A consistent estimate of  $C$  is given by the  $s \times h$  matrix

$$(29) \quad \hat{C} = \hat{U}' \hat{V}_\varepsilon Z / T \hat{\sigma}_\varepsilon^2, \quad \text{where} \quad \hat{V}_\varepsilon(i, j) = \hat{r}_{|i-j|} \hat{\sigma}_\varepsilon^2$$

is an estimate of the covariance matrix of the error term.

<sup>13</sup> McLeod (1978) derives the asymptotic distribution of residual autocorrelations from univariate ARMA models with homoscedastic errors and, as we do here, suggests using a consistent estimate of the asymptotic covariance matrix to form a Wald test as an alternative to the Box-Pierce test. Breusch and Godfrey (1981) describe unpublished work by Sargan (1976) that suggests a test that is equivalent to the  $l_{q,s}$  test when  $q = 0$  and the residuals are conditionally homoscedastic.



While the analysis of this paper centers on the asymptotic distribution of simple autocorrelations, the results are also relevant for the asymptotic distribution of partial autocorrelations of regression residuals. Regressing  $\hat{\varepsilon}_t$  on  $\hat{\varepsilon}_{t-q-1}, \dots, \hat{\varepsilon}_{t-q-s}$  yields the estimated coefficient vector  $b = \hat{F}\hat{r}$ , where  $\hat{F} = (\hat{U}'\hat{U})^{-1}\hat{U}'\hat{\varepsilon}$  converges in probability to a  $s \times s$  matrix  $F$ . As a result,  $b$  converges in distribution to a normal random variable with mean zero and covariance matrix  $V_b = FV_\varepsilon F'$  when the null hypothesis is true, and the standard Wald statistic for testing  $b = 0$  will be numerically identical to the  $l$  statistic described in Proposition 2.<sup>14</sup> In the special case of  $q = 0$ ,  $F$  is an identity matrix so that even though  $b$  will not equal  $\hat{r}$  in finite samples, one can replace  $\hat{r}$  with  $b$  in Proposition 2 and obtain a valid test.

#### 4. CONCLUDING REMARKS

The distribution of the  $l$  test described in this paper is based on asymptotic distribution theory and thus leaves open the issue of how well the test will work in finite samples. In Cumby and Huizinga (1990) we report the results of a series of Monte Carlo experiments that investigate the performance of the  $l$  test in a number of models. In each model, the performance of the  $l$  test is evaluated both in sample sizes of 50 and 100 observations, both with errors that are conditionally homoscedastic and errors that are conditionally heteroscedastic, and both with  $q = 0$  and  $q = 2$ .

Several conclusions emerge from these Monte Carlo experiments. First, the small sample distribution of the  $l$  statistic is close to the asymptotic distribution in the tails, so that the test is reliable in terms of its size. Second, performance of the  $l$  test in the case of  $q = 2$  is sensitive to how the matrix  $\Psi$  in equation (18) is estimated. The best results were obtained using equation (27) and the "Gaussian" weights  $w_i^N = \exp(-i^2/2N^2)$ .<sup>15</sup> Third, while the presence of conditionally heteroscedastic errors does not present any problems for the  $l$  test, it often makes the popular portmanteau tests suggested by Box and Pierce (1970) and Ljung and Box (1978) quite unreliable.

*Stern School of Business, New York University, 100 Trinity Place, New York, NY 10006, U.S.A.*

*and*

*Graduate School of Business, University of Chicago, 1101 E. 58th Street, Chicago, IL 60637, U.S.A.*

*Manuscript received January, 1989; final revision received March, 1991.*

#### APPENDIX

This appendix provides a proof of the main proposition in the text. Let  $y_t, \varepsilon_t, X_{1,t}, \dots, X_{k,t}, Z_{1,t}, \dots, Z_{h,t}$  be scalar random variables on which we have observations for  $t = 1, \dots, T$ . Define  $X_t'$  and  $Z_t'$  to be the  $1 \times k$  and  $1 \times h$  vectors  $(X_{1,t}, \dots, X_{k,t})$  and  $(Z_{1,t}, \dots, Z_{h,t})$ , and define  $y, X$ , and  $Z$  to be the  $T \times 1$ ,  $T \times k$ , and  $T \times h$  matrices  $(y_1, \dots, y_T)'$ ,  $(X'_1, \dots, X'_T)'$ , and  $(Z'_1, \dots, Z'_T)'$ . Define  $\eta_t = (\omega_{1,t}, \dots, \omega_{h,t}, \xi_{1,t}, \dots, \xi_{s,t})'$  for  $\omega_{j,t} = \varepsilon_t Z_{j,t}$  and  $\xi_{i,t} = \varepsilon_t \varepsilon_{t-q-i}$ , and let  $A_T$  be an observable

<sup>14</sup> Since the estimated covariance matrix for  $b$  reported by standard regression packages will not in general be a consistent estimate of  $V_b$ , testing  $b = 0$  with the typical  $F$  test reported by these packages is not an asymptotically valid procedure.

<sup>15</sup> See Brillinger (1975, p. 55) for a discussion of the Gaussian weighting scheme. Details on the optimal choice of  $N$  are discussed in Cumby and Huizinga (1990). The performance of the  $l$  test was substantially poorer when  $\Psi$  was estimated with equation (27) and the modified Bartlett weights discussed by Anderson (1971) and Newey and West (1987).

$h \times h$  matrix. We assume that for a known constant  $q$  and unknown  $k \times 1$  vector of constants  $\delta$ ,

$$(A1) \quad \{X_t, Z_t, \varepsilon_t\} \text{ is wide sense stationary and ergodic,}$$

$$(A2) \quad y_t = X_t \delta + \varepsilon_t,$$

$$(A3) \quad E(\varepsilon_t) = 0,$$

$$(A4) \quad E(\varepsilon_t | Z_t, Z_{t-1}, \dots, \varepsilon_{t-q-1}, \varepsilon_{t-q-2}, \dots) = 0,$$

$$(A5) \quad E(\eta_t \eta'_{t-i}) \text{ is finite for } i = 0, \dots, q,$$

$$(A6) \quad \Psi = E(\eta_t \eta'_{t-q}) + E(\eta_t \eta'_{t-q+1}) + \dots + E(\eta_t \eta'_{t+q-1}) + E(\eta_t \eta'_{t+q}) \text{ is positive definite,}$$

$$(A7) \quad (1/T) \text{ plim } X'Z \text{ exists and has rank } k,$$

and

$$(A8) \quad \text{plim } A_T = A \text{ exists and is nonsingular.}$$

Define  $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ ,  $E(\varepsilon_t \varepsilon_{t-n}) / \sigma_\varepsilon^2 = \rho_n$ ,  $r_n$  to be the sample autocorrelation of  $\varepsilon_t$  at lag  $n$ , the  $s \times 1$  vector  $r = (r_{q+1}, r_{q+2}, \dots, r_{q+s})'$ , the  $k \times 1$  vector  $d = (X'Z A_T^{-1} Z'X)^{-1} X'Z A_T^{-1} Z'y$ ,  $\hat{r}_n$  to be the sample autocorrelation of  $\hat{\varepsilon}_t = y_t - X_t d$  at lag  $n$ , and the  $s \times 1$  vector  $\hat{r} = (\hat{r}_{q+1}, \hat{r}_{q+2}, \dots, \hat{r}_{q+s})$ . We also define  $C$  to be the  $s \times h$  matrix that has  $\sigma_\varepsilon^{-2}$  times the sum from  $n$  equals  $-q$  to  $q$  of  $E(\xi_{i,t} \omega_{j,t-n})$  as its  $ij$ th element,  $V_r$  to be the  $s \times s$  matrix that has  $\sigma_\varepsilon^{-4}$  times the sum from  $n$  equals  $-q$  to  $q$  of  $E(\xi_{i,t} \xi_{j,t-n})$  as its  $ij$ th element,  $U$  to be the  $T \times s$  matrix that has  $q+j$  zeros followed by  $\varepsilon_t$ ,  $t = 1, \dots, T - q - j$  as its  $j$ th column, and  $B$  to be the  $s \times k$  matrix that has  $-[E(\varepsilon_{t-q-i} X_{j,t}) + E(X_{j,t-q-i} \varepsilon_t)] / \sigma_\varepsilon^2$  as its  $ij$ th element.

LEMMA 1: Given the assumptions (A1)–(A8) and the definitions stated above,  $d$  is a consistent estimate of  $\delta$  and  $\sqrt{T}(d - \delta) \overset{A}{\sim} N(0, V_d)$  where  $V_d = D\Omega D'$ ,  $D = \text{plim } T(X'Z A_T^{-1} Z'X)^{-1} X'Z A_T^{-1}$ , and  $\Omega = E(\omega_t \omega'_{t-q}) + E(\omega_t \omega'_{t-q+1}) + \dots + E(\omega_t \omega'_{t+q-1}) + E(\omega_t \omega'_{t+q})$ .

PROOF: The proof can be found in Cumby, Huizinga, and Obstfeld (1983).

LEMMA 2: Given the assumptions (A1)–(A8) and the definitions stated above,  $\text{plim } \partial r / \partial \delta = B$ .

PROOF:  $\partial r / \partial \delta$  has as its  $ij$ th element

$$(\partial r_{i,t} / \partial \delta_j) \big|_{\delta=d_i^*} = \frac{\partial}{\partial \delta_j} \frac{\sum e_t e_{t-q-i}}{\sum e_t^2}, \quad \text{where} \quad e_t = y_t - X_t d_i^*$$

and  $d_i^*$  lies between  $d$  and  $\delta$ . Differentiating, we obtain

$$-\frac{\sum [X_{j,t} e_{t-q-i} + X_{j,t-q-i} e_t]}{\sum e_t^2} + 2 \frac{\sum e_t e_{t-q-i} \sum X_{j,t} e_t}{[\sum e_t^2]^2}.$$

Therefore, using (A1), the fact that  $d$  is a consistent estimate of  $\delta$  (Lemma 1), and the fact that  $d_i^*$  lies between  $d$  and  $\delta$ , we get

$$\text{plim } \frac{\partial r_{i,t}}{\partial \delta_j} \bigg|_{\delta=d_i^*} = -\frac{E(X_{j,t} \varepsilon_{t-q-i})}{\sigma_\varepsilon^2} - \frac{E(X_{j,t-q-i} \varepsilon_t)}{\sigma_\varepsilon^2} + 2\rho_{q+i} \frac{E(X_{j,t} \varepsilon_t)}{\sigma_\varepsilon^2}.$$

Since  $\rho_j = 0$  for  $j > q$ , the third term in this sum is zero and the lemma is proved.

Q.E.D.

The proof of Proposition 1 is now straightforward.

PROPOSITION 1: *Given the assumptions (A1)–(A8) and the definitions above,  $\sqrt{T}\hat{r} \stackrel{A}{\sim} N(0, V_r)$  with  $V_r = V_r + BV_dB' + BDC' + CD'B'$ .*

PROOF: By the mean-value theorem,

$$\sqrt{T}\hat{r} = \sqrt{T}r + \sqrt{T} \frac{\partial r}{\partial \delta}(d - \delta),$$

where the  $i$ th row of  $\partial r/\partial \delta$  is evaluated at  $d_i^*$ , which lies between  $d$  and  $\delta$ . Stacking the terms on the right-hand side of this expression and substituting the definitions of  $d$  and  $U$  gives

$$\sqrt{T} \begin{bmatrix} \partial r/\partial \delta \\ r \end{bmatrix} (d - \delta) = \begin{bmatrix} \partial r/\partial \delta T (X'ZA_T^{-1}Z'X)^{-1}X'ZA_T^{-1} & 0 \\ 0 & (\varepsilon'\varepsilon/T)^{-1}I \end{bmatrix} \begin{bmatrix} Z'\varepsilon/\sqrt{T} \\ U'\varepsilon/\sqrt{T} \end{bmatrix},$$

where  $I$  is an  $s \times s$  identity matrix. By Lemma 2, (A7), and (A8),

$$\text{plim} \begin{bmatrix} \partial r/\partial \delta T (X'ZA_T^{-1}Z'X)^{-1}X'ZA_T^{-1} & 0 \\ 0 & (\varepsilon'\varepsilon/T)^{-1}I \end{bmatrix} = \begin{bmatrix} BD & 0 \\ 0 & \sigma_\varepsilon^{-2}I \end{bmatrix} = \Phi$$

and by a central limit theorem in Hannan (1973) (see Hansen (1982)),

$$\begin{bmatrix} Z'\varepsilon/\sqrt{T} \\ U'\varepsilon/\sqrt{T} \end{bmatrix} \stackrel{A}{\sim} N(0, \Psi)$$

for

$$\Psi = \begin{bmatrix} \Omega & C'\sigma_\varepsilon^2 \\ C\sigma_\varepsilon^2 & V_r\sigma_\varepsilon^4 \end{bmatrix}.$$

Thus,

$$\sqrt{T} \begin{bmatrix} \partial r/\partial \delta \\ r \end{bmatrix} (d - \delta) \stackrel{A}{\sim} N(0, \Phi\Psi\Phi')$$

where

$$\Phi\Psi\Phi' = \begin{bmatrix} BV_dB' & BDC' \\ CD'B' & V_r \end{bmatrix}.$$

Since  $\sqrt{T}\hat{r}$  is the sum of the two random vectors that are asymptotically normally distributed with covariance matrix  $\Phi\Psi\Phi'$ , it follows that  $\sqrt{T}\hat{r}$  is asymptotically normally distributed with covariance matrix given by  $BV_dB' + BDC' + CD'B' + V_r$  and the proof is complete. Q.E.D.

In the text, we discuss how the asymptotic distribution of  $\sqrt{T}\hat{r}$  is affected when the assumption that  $\varepsilon_t$  is conditionally homoscedastic,

$$(A9) \quad E(\varepsilon_t \varepsilon_{t-n} | Z_t, Z_{t-1}, \dots, \varepsilon_{t-q-1}, \varepsilon_{t-q-2}, \dots) = E(\varepsilon_t \varepsilon_{t-n}), \quad 0 \leq n \leq q,$$

is added to assumptions (A1)–(A8) above. In particular, equations (15) and (16) give forms of equations (13) and (14) which are claimed to be valid when this assumption is added. To verify that equation (15) is in fact correct, note that when (A9) holds and  $-q \leq n \leq q$ ,

$$\begin{aligned} \sigma_\varepsilon^{-4} E(\xi_{i,t} \xi_{j,t-n}) &= \sigma_\varepsilon^{-4} E(\varepsilon_{t-q-i} \varepsilon_{t-n} \varepsilon_{t-n-q-j}) \\ &= \sigma_\varepsilon^{-4} E\{\varepsilon_{t-q-i} \varepsilon_{t-n-q-j} E(\varepsilon_{t-n} | \varepsilon_{t-q-i}, \varepsilon_{t-n-q-j})\} \\ &= \sigma_\varepsilon^{-4} E(\varepsilon_{t-q-i} \varepsilon_{t-n-q-j}) E(\varepsilon_t \varepsilon_{t-n}) = \rho_{n+j-i} \rho_n. \end{aligned}$$

Equation (16) can be verified in a similar manner.

## REFERENCES

- ANDERSON, T. W. (1971): *The Statistical Analysis of Time Series*. New York: John Wiley and Sons.
- BOX, G. E. P., AND D. A. PIERCE (1970): "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *Journal of the American Statistical Association*, 65, 1509–1526.
- BREUSCH, T. S., AND L. GODFREY (1981): "A Review of Recent Work on Testing for Autocorrelation in Dynamic Simultaneous Models," in *Macroeconomic Analysis*, ed. by D. Currie, R. Nobay, and D. Peel. London: Croon-Helm.
- BRILLINGER, D. R. (1975): *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart, and Winston.
- CHRISTIANO, L. J., M. EICHENBAUM, AND D. MARSHALL (1991): "The Permanent Income Hypothesis Revisited," *Econometrica*, 59, 397–423.
- CUMBY, R. E., AND J. HUIZINGA (1990): "Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions," NBER Technical Working Paper No. 90.
- CUMBY, R. E., J. HUIZINGA, AND M. OBSTFELD (1983): "Two-Step Two-Stage Least Squares Estimation in Models with Rational Expectations," *Journal of Econometrics*, 21, 333–355.
- DIEBOLD, F. X. (1986): "Testing for Serial Correlation in the Presence of ARCH," *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, 323–328.
- DURBIN, J. (1970): "Testing for Serial Correlation in Least Squares Regression Models when some of the Regressors Are Lagged Dependent Variables," *Econometrica*, 38, 410–421.
- ENGLE, R. F. (1982): "Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1008.
- FAMA, E. F., AND K. R. FRENCH (1988): "Permanent and Transitory Components of Stock Prices," *Journal of Political Economy*, 96, 246–273.
- GODFREY, L. G. (1978a): "A Note on the Use of Durbin's  $h$  Test when the Equation Is Estimated by Instrumental Variables," *Econometrica*, 46, 225–228.
- (1978b): "Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables," *Econometrica*, 46, 1293–1301.
- (1978c): "Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables," *Econometrica*, 46, 1303–1310.
- (1989): *Misspecification Tests in Econometrics*. New York: Cambridge University Press.
- HALL, R. E. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96, 339–357.
- HANNAN, E. J. (1973): "Central Limit Theorems for Time Series Regressions," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 26, 157–170.
- HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.
- HANSEN, L. P., AND R. J. HODRICK (1980): "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829–853.
- HANSEN, L. P., AND K. J. SINGLETON (1990): "Efficient Estimation of Linear Asset Pricing Models with Moving Average Errors," NBER Technical Working Paper No. 86.
- HUIZINGA, J., AND F. S. MISHKIN (1984): "Inflation and Real Interest Rates on Assets with Different Risk Characteristics," *Journal of Finance*, 34, 699–714.
- JOHNSON, N. L., AND S. KOTZ (1970): *Continuous Univariate Distributions—2*. New York: John Wiley and Sons.
- LJUNG, G. M. (1986): "Diagnostic Testing of Univariate Time Series Models," *Biometrika*, 73, 725–730.
- LJUNG, G. M., AND G. E. P. BOX (1978): "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.
- MCLEOD, A. I. (1978): "On the Distribution of Residual Autocorrelations in Box-Jenkins Models," *Journal of the Royal Statistical Society*, 40, 296–302.
- MISHKIN, F. S. (1990): "What Does the Term Structure Tell Us About Future Inflation?" *Journal of Monetary Economics*, 25, 77–95.
- NEWBY, W. K., AND K. D. WEST (1987): "A Simple, Positive Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.
- SARGAN, J. D. (1976): "Testing for Misspecification After Estimating Using Instrumental Variables," London School of Economics Working Paper.