Testing Vector Error Autocorrelation and Heteroscedasticity[1]

BY JURGEN A. DOORNIK
Nuffield College, Oxford OX1 1NF, UK

22nd September 1996

ABSTRACT

This paper investigates tests for serial correlation and heteroscedasticity which can be applied to VARs and simultaneous equations models.

An F-type test for vector error autocorrelation is considered. The test is conceptually simple and reduces to the familiar single equation test in a one-equation model. A wide range of Monte Carlo experiments is set up to investigate size and power, and compare the test to the standard chi-squared and multivariate portmanteau tests. The non-centrality of the chi-squared test gives asymptotic power similar to that found in the simulations.

Next, it is shown that under simplifying assumptions, Kelejian's test for heteroscedasticity reduces to a multivariate extension of White's test, which consists of a multivariate regression of the residual variances and correlations on the squares and cross-products of the original regressors. Various forms of the test are considered in simulation experiments.

Key words: Lagrange multiplier test; Likelihood ratio test; Non-centrality; Rao's F-approximation; Vector autocorrelation test; White's heteroscedasticity test.

[1] Revised version of a paper presented at the Econometric Society 7th World Congress, Tokio. All comments are welcome. Please direct them to Jurgen Doornik at jurgen.doornik@nuffield.ox.ac.uk.

Testing Vector Autocorrelation and Heteroscedasticity in Dynamic Models

BY JURGEN A. DOORNIK

Nuffield College, Oxford OX1 1NF, UK

22nd September 1996

SUMMARY

An F-type test for vector error autocorrelation is considered. The test is conceptually simple and reduces to the familiar single equation test in a one-equation model. A wide range of Monte Carlo experiments is set up to investigate size and power, and compare the test to the standard chi-squared and multivariate portmanteau tests. The non-centrality of the chi-squared test gives asymptotic power similar to that found in the simulations.

Next, it is shown that under simplifying assumptions, Kelejian's test for heteroscedasticity reduces to a multivariate extension of White's test, which consists of a multivariate regression of the residual variances and correlations on the squares and cross-products of the original regressors. Various forms of the test are considered in simulation experiments.

Key words: Lagrange multiplier test; Likelihood ratio test; Non-centrality; Rao's F-approximation; Vector autocorrelation test.

1 INTRODUCTION

Following the example of earlier versions of PcGive, most computer packages now have a range of diagnostic tests preprogrammed. As a consequence, it is noe a routine operation to subject single equation regression equations to a battery of tests, e.g. for autocorrelation, heteroscedasticity, normality and parameter constancy. The current popularity of multivariate methods, such as vector autoregressions, requires a corresponding range of diagnostic test procedures. The subject of this paper is to investigate two such vector tests, namely for error autocorrelation and heteroscedasticity (testing normality is considered in a separate paper, see Doornik and Hansen, 1994). To facilitate applied research, we require acceptable small sample behaviour, but also aim to have conceptually simple tests, which reduce to their univariate counterparts in a single equation setting.

The single equation Lagrange multiplier (LM) test for autocorrelation developed by Breusch (1978) and Godfrey (1978) (reviewed in Godfrey, 1988), has become a standard tool in applied econometrics. The pervasiveness of this test procedure derives from its simplicity, wide applicability (unlike e.g. the Durbin-Watson statistic which needs to be adjusted for dynamic models) and flexibility (it can be used to check for any order of autocorrelation), also see Breusch and Godfrey (1981, §2.2). Finally, the F-form of the test performs well in small samples.

The test is performed through an auxiliary regression of the residuals on their lags and the original regressors. Then the significance of all regressors is tested. Two forms are generally computed:

(1) TR^2, where T is the sample size, and R^2 the coefficient of multiple correlation in the auxiliary regression. This statistic has an asymptotic $\chi^2(s)$ distribution under the null of no serial correlation

when s lagged residuals are used.

(2) the F-test on the lagged residuals in the auxiliary regression:

$$\frac{R^2}{1-R^2}\frac{T-k-s}{s} \sim F(s, T-k-s), \tag{1}$$

where k is the number of regressors in the original regression. This modified LM procedure was first suggested by Durbin (1970) (also see Osborn, 1981).

In both cases it is convenient to set missing observations for the lagged residuals to zero, so that no observations are lost. Identical statistics can be obtained when using the original dependent variable as the dependent variable in the auxiliary regression; then R^2 should be defined relative to the original regressors: $R^2 = (\text{RSS}_0)^{-1}(\text{RSS}_0 - \text{RSS})$, where RSS_0 and RSS are the residual sums of squares of the original and the auxiliary regression respectively.

Kiviet (1986) compared a large number of tests for autocorrelation in a Monte Carlo study. He found that the F-version is better behaved in small samples and retains a correct size in an overspecified model (where TR^2 overrejects). Mizon and Hendry (1980) compared the LM form with the Wald and likelihood ratio (LR) version in a model with first order autocorrelation (satisfying common factor restrictions) and found evidence favouring the LM statistic. Godfrey (1981) also finds that the LM test is effective relative to the LR test.

For testing heteroscedasticity we focus on White's test, see White (1980), primarily because it does not require explicit formulation of the form of heteroscedasticity. In a single equation setting the test amounts to adding the squares and cross-products of the original regressors to an auxiliary regression of the squared residuals on a constant term and testing the significance of these added terms. Under the null hypothesis of homoscedasticity (normality is not required, the assumption is that the errors are $\mathsf{IID}(0, \sigma^2)$ with constant kurtosis and the first eight moments exist), the squares and cross-products of the original regressors have a coefficient of zero. In general, it will be necessary to remove redundant variables. Let s denote the number of added terms; when there are k regressors in the original equation, including the constant term, and no other redundancies: $s = \frac{1}{2}k(k-1)$. The test can be computed as TR^2 from the auxiliary regression and will have an asymptotic $\chi^2(s)$ distribution under the null. As with the LM test for autocorrelation, an F-form may be considered which could potentially achieve better small sample behaviour.

White's test is just one from a whole spectrum of tests proposed in the literature. A comparative study of Ali and Giaccotto (1984) shows that the size of White's test is robust against some non-normal error distributions and is among the best tests in terms of power. Godfrey and Orme (1994) on the other hand find significant deviations from the nominal size when the errors are generated from a log-normal or $\chi^2(2)$. test. They also show that White's heteroscedasticity test does not have power against omitted variables, and criticise PcGive (Doornik and Hendry, 1994b) for calling this a test for functional form. (PcGive reports two forms of the test: one involving cross-products and squares, one using squares only.)

The rest of the paper is organised as follows. I first briefly review specification tests in multivariate systems to introduce notation. Section 3 then discusses the vector version of the LM test, and introduces an approximate F-version, corresponding to the modified LM test for the single equation model. The subsequent two sections consider the vector portmanteau test, and testing in the simultaneous equations model. Then the small sample properties of these two tests are investigated in a set of Monte Carlo experiments, and compared to the multivariate portmanteau test. Finally, the non-centrality of two of the power simulations is computed.

2 System specification tests

Consider the n-dimensional multivariate linear regression model:

$$\boldsymbol{y}_t = \boldsymbol{\Pi} \boldsymbol{w}_t + \boldsymbol{v}_t, \quad \mathsf{E}[\boldsymbol{v}_t] = \boldsymbol{0}, \quad \mathsf{E}[\boldsymbol{v}_t \boldsymbol{v}_t'] = \boldsymbol{\Omega}.$$

In matrix form:

$$\boldsymbol{Y}' = \boldsymbol{\Pi} \boldsymbol{W}' + \boldsymbol{V}', \tag{2}$$

in which $\boldsymbol{Y}'$ is $n \times T$, $\boldsymbol{W}'$ is $k \times T$ and $\boldsymbol{\Pi}$ is $n \times k$. $\boldsymbol{W}'$ may include lagged dependent variables; I assume it includes a constant term. The multivariate least squares (MLS) estimates of the coefficients and residual covariance are:

$$\hat{\boldsymbol{\Pi}}' = (\boldsymbol{W}'\boldsymbol{W})^{-1} \boldsymbol{W}'\boldsymbol{Y} \text{ and } \tilde{\boldsymbol{\Omega}} = \hat{\boldsymbol{V}}'\hat{\boldsymbol{V}} / (T-k),$$

where the residuals are defined by:

$$\hat{\boldsymbol{V}} = \boldsymbol{Y} - \boldsymbol{W} \hat{\boldsymbol{\Pi}}'.$$

When $\boldsymbol{v}_t \sim \mathsf{N}(\boldsymbol{0}, \boldsymbol{\Omega})$, the maximum likelihood estimates are $\hat{\boldsymbol{\Pi}}'$ as before, and $\hat{\boldsymbol{\Omega}} = T^{-1} \hat{\boldsymbol{V}}'\hat{\boldsymbol{V}}$.

Tests whether columns of $\boldsymbol{\Pi}$ are zero can be based on the Wald, LM and LR principle. Partitioning the coefficients as $\boldsymbol{\Pi} = (\boldsymbol{\Pi}_1 : \boldsymbol{\Pi}_2)$, and $\boldsymbol{W} = (\boldsymbol{W}_1 : \boldsymbol{W}_2)$ accordingly, we may write this test as:

$$\mathsf{H}_0 : \boldsymbol{\Pi}_2 = \boldsymbol{0} \text{ versus } \mathsf{H}_1 : \boldsymbol{\Pi}_2 \neq \boldsymbol{0},$$

with the maintained hypothesis given in (2). The matrices $\boldsymbol{\Pi}_i$ are $n \times k_i$, so that $k = k_1 + k_2$. The likelihood ratio is:

$$\hat{\lambda} = \left(\left| \hat{\boldsymbol{V}}'\hat{\boldsymbol{V}} \right| \left| \hat{\boldsymbol{V}}_0'\hat{\boldsymbol{V}}_0 \right|^{-1} \right)^{T/2},$$

where $\hat{\boldsymbol{V}}_0$ are the residuals from regressing $\boldsymbol{Y}$ on $\boldsymbol{W}_1$ (that is, under H_0), whereas $\hat{\boldsymbol{V}}$ results from the unrestricted system (2). Minus twice the logarithm of $\hat{\lambda}$ is asymptotically $\chi^2(nk_2)$ distributed under the null hypothesis. Anderson gives small-sample correction factors for the χ^2 test; Anderson (1984, §8.4) and Rao (1973, §8b.2) derive the exact distribution of $\lambda^{2/T}$ for fixed $\boldsymbol{W}$, which is called the U-test.[1] The corresponding LM-test is:

$$\hat{\mu} = T \operatorname{tr} \left\{ \left(\hat{\boldsymbol{V}}_0'\hat{\boldsymbol{V}}_0 - \hat{\boldsymbol{V}}'\hat{\boldsymbol{V}} \right) \left(\hat{\boldsymbol{V}}_0'\hat{\boldsymbol{V}}_0 \right)^{-1} \right\},$$

which is also asymptotically $\chi^2(nk_2)$ distributed under the null hypothesis.

Based on these, it is convenient to define two R^2-type measures of goodness of fit:

$$\begin{aligned} R_r^2 &= 1 - \left| \hat{\boldsymbol{V}}'\hat{\boldsymbol{V}} \right| \left| \hat{\boldsymbol{V}}_0'\hat{\boldsymbol{V}}_0 \right|^{-1} \\ R_m^2 &= 1 - \frac{1}{n} \operatorname{tr} \left\{ \left(\hat{\boldsymbol{V}}'\hat{\boldsymbol{V}} \right) \left(\hat{\boldsymbol{V}}_0'\hat{\boldsymbol{V}}_0 \right)^{-1} \right\}. \end{aligned}$$

In a one-equation system ($n = 1$), or when $\boldsymbol{V}$ is diagonal, these two measures are identical. In the first case they correspond to the traditional R^2 if the constant term is the only variable excluded in the specification test.

[1] Guilkey (1974) and Deschamps (1994) consider this form of the vector error autocorrelation test, but only in models excluding lagged dependent variables.

3 Vector error autocorrelation tests

Consider the augmented system with vector autoregressive errors:

$$\boldsymbol{Y}' = \boldsymbol{\Pi}\boldsymbol{W}' + \boldsymbol{U}' \quad \text{where} \quad \boldsymbol{U}' = \sum_{i=1}^{s} \boldsymbol{R}_i \boldsymbol{U}'_{-i} + \boldsymbol{V}'. \tag{3}$$

We wish to test the null hypothesis:

$$\mathsf{H}_0 : \boldsymbol{R}_1 = \cdots = \boldsymbol{R}_s = \boldsymbol{0}. \tag{4}$$

Godfrey (1981) showed that this test can be implemented by testing the significance of the lagged residuals (obtained under the null hypothesis) in the auxiliary system:

$$\boldsymbol{Y}' = \boldsymbol{\Pi}\boldsymbol{W}' + \boldsymbol{R}_1 \hat{\boldsymbol{U}}'_{-1} + \cdots + \boldsymbol{R}_s \hat{\boldsymbol{U}}'_{-s} + \boldsymbol{V}'. \tag{5}$$

This is most easily implemented by partialling out lagged residuals from the original regressors, and re-estimating the original system using the new regressors. Again we set missing observations to zero. Godfrey also showed that testing for vector MA(s) residuals is locally equivalent and leads to the same procedure.

When $\hat{\boldsymbol{V}}$ denotes the residuals from the auxiliary regression, and $\hat{\boldsymbol{V}}_0 = \hat{\boldsymbol{U}}$ the residuals of the system under the null, both R^2_m and R^2_r of the previous section can be computed. The χ^2 form of the LM test is given by:

$$LM = TnR^2_m, \tag{6}$$

with an asymptotic $\chi^2(sn^2)$ distribution.

Using an F-approximation to the likelihood-ratio form of the LM test which is due to Rao (see Rao, 1973, §8c.5, or Anderson, 1984, §8.5.4):

$$LMF = \frac{1 - \left(1 - R^2_r\right)^{1/r}}{\left(1 - R^2_r\right)^{1/r}} . \frac{Nr - q}{np}, \tag{7}$$

with:

$$r = \left(\frac{n^2p^2 - 4}{n^2 + p^2 - 5}\right)^{1/2}, \; q = \tfrac{1}{2}np - 1, \; N = T - k - p - \tfrac{1}{2}\left(n - p + 1\right)$$

and:

k	number of regressors in original system,
n	dimension of system,
T	number of observations,
p	number of regressors added in auxiliary system $(= ns)$.

LMF has an approximate $F(np, Ns - q)$ distribution (the F-approximation is exact for fixed regressors when $p \leq 2$ or $n \leq 2$).

A larger number of transformations of the LM and LR statistics for testing the significance of the lagged residuals is considered by Edgerton and Shukur (1995).

Both *LM* and *LMF* have the attractive property of reducing to the single equation LM and modified LM tests: $n = 1$ in (7) results in $p = s$, $r = 1$, $q = \frac{1}{2}s - 1$, $N = T - k - \frac{1}{2}s - 1$ so that $Nr - q = T - k - s$ and (7) reduces to (1). As in the univariate case we may use an auxiliary regression with $\hat{\boldsymbol{U}}$ replacing $\boldsymbol{Y}$ as regressand.

Testing in the simultaneous equations model (SEM) proceeds similarly. Under the null hypothesis the model is:

$$\boldsymbol{B}\boldsymbol{Y}' + \boldsymbol{C}\boldsymbol{W}' = \boldsymbol{H}', \tag{8}$$

with

$$\boldsymbol{H}' = \sum_{i=1}^{s} \boldsymbol{R}_i \boldsymbol{H}'_{-i} + \boldsymbol{E}'. \text{ and } \mathsf{H}_0 : \boldsymbol{R}_1 = \cdots = \boldsymbol{R}_s = \boldsymbol{0},$$

where the $n \times n$ matrix $\boldsymbol{B}$ is non-singular and has unity on the diagonal. Assume the model is estimated by full information maximum likelihood (FIML), and the FIML residuals $\hat{\boldsymbol{H}}$ have been obtained, together with the restricted reduced form (RRF) residuals $\hat{\boldsymbol{V}}_0' = \hat{\boldsymbol{B}}^{-1}\hat{\boldsymbol{H}}'$. Next, to test for vector error autocorrelation of order s, estimate the auxiliary model

$$\boldsymbol{B}\boldsymbol{Y}' + \boldsymbol{C}\boldsymbol{W}' - \boldsymbol{R}_1\hat{\boldsymbol{H}}'_{-1} - \cdots - \boldsymbol{R}_s\hat{\boldsymbol{H}}'_{-s} = \boldsymbol{E}' \tag{9}$$

by FIML,[2] giving structural residuals $\hat{\boldsymbol{E}}$, and reduced form residuals $\hat{\boldsymbol{V}}' = \hat{\boldsymbol{B}}^{-1}\hat{\boldsymbol{E}}'$. The $\hat{\boldsymbol{V}}_0'$ and $\hat{\boldsymbol{V}}'$ thus created can now be used to compute R_m^2 and R_r^2, and from that the χ^2 and F forms of the LM test for the hypothesis $\boldsymbol{R}_1 = \cdots = \boldsymbol{R}_s = \boldsymbol{0}$. When modelling starts with the URF, testing for autocorrelation will also start there.

4 Vector portmanteau test

In the Monte Carlo comparisons we also consider two forms of the multivariate portmanteau statistic, as introduced by Hosking (1980). This statistic serves as a goodness-of-fit test in multivariate stationary ARMA processes, and is an extension to the tests of Box and Pierce (1970) and Ljung and Box (1978).

Define:

$$\hat{\boldsymbol{C}}_{rs} = \frac{1}{T}\hat{\boldsymbol{U}}'_{-r}\hat{\boldsymbol{U}}_{-s},$$

with $\hat{\boldsymbol{U}}_{-i}$ the $T \times n$ residual matrix i periods lagged, where missing values are set to zero. Then $\hat{\boldsymbol{C}}_{00} = \hat{\boldsymbol{\Omega}}$. The vector portmanteau statistic is:

$$Q(s) = T\sum_{j=1}^{s} \operatorname{tr}\left(\hat{\boldsymbol{C}}'_{0j}\hat{\boldsymbol{C}}_{00}^{-1}\hat{\boldsymbol{C}}_{0j}\hat{\boldsymbol{C}}_{00}^{-1}\right).$$

Hosking argues that, as in the univariate case, a modified form might be better behaved in small samples:

$$Q^*(s) = T^2\sum_{j=1}^{s} \frac{1}{T-j} \operatorname{tr}\left(\hat{\boldsymbol{C}}'_{0j}\hat{\boldsymbol{C}}_{00}^{-1}\hat{\boldsymbol{C}}_{0j}\hat{\boldsymbol{C}}_{00}^{-1}\right).$$

For an n-dimensional VAR with lag length m, both statistics are asymptotically $\chi^2(n^2(s-m))$ distributed under the assumptions of the test (one of them being that s is large: $s = O(T^{1/2})$). Ahn (1988) showed that the test statistic is also valid in a stationary VAR with parameter restrictions (in that case n^2m should be replaced by the appropriate number of estimated coefficients). However, as Breusch and Pagan (1980) argue, the portmanteau statistic is invalid in a model including both non-modelled and lagged endogenous variables. Monte Carlo evidence in Kiviet (1986) supports this.

[2] As pointed out by Godfrey (1988), since the lagged residuals enter unrestrictedly, $\boldsymbol{Y}'$ and $\boldsymbol{W}'$ can be replaced by the residuals from regression on the lagged residuals, resulting in the same model structure as the restricted model.

5 VECTOR HETEROSCEDASTICITY TEST

The objective here is to test the residuals of the system (2) for heteroscedasticity:

$$\mathsf{H}_0 : \mathsf{E}\left[\boldsymbol{v}_t \boldsymbol{v}_t'\right] = \boldsymbol{\Omega}.$$

Kelejian (1982) extended White's test to the simultaneous equations framework. However, the test has not found widespread application, probably because it was considered rather cumbersome (see e.g. the footnote on p.187 in Godfrey, 1988).

In a two-equation system Kelejian's procedure would amount to applying GLS to the following regression equation:

$$\begin{pmatrix} \hat{v}_{11}^2 \\ \vdots \\ \hat{v}_{T1}^2 \\ \hat{v}_{12}^2 \\ \vdots \\ \hat{v}_{T2}^2 \\ \hat{v}_{11}\hat{v}_{12} \\ \vdots \\ \hat{v}_{T1}\hat{v}_{T2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ & \vdots & \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ & \vdots & \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ & \vdots & \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_1^2 \\ \omega_2^2 \\ \omega_{12} \end{pmatrix} + \begin{pmatrix} \boldsymbol{P}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{P}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{P}_3 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \boldsymbol{\alpha}_3 \end{pmatrix} + \boldsymbol{\epsilon}. \tag{10}$$

In this expression, the notation becomes more compact as we move from left to right, $\boldsymbol{P}_i$ is $(3T \times p_i)$, $\boldsymbol{\alpha}_i$ is $(p_i \times 1)$ and $\boldsymbol{\epsilon}$ is $(3T \times 1)$. Each $\boldsymbol{P}_i$ could consist of different functions of the regressors $\boldsymbol{w}_t$. Taking deviations from means (10) may be written as

$$\text{vec}\left(\check{\boldsymbol{\Psi}}'\right) = \check{\boldsymbol{P}}^* \boldsymbol{\alpha} + \boldsymbol{\epsilon} \tag{11}$$

in which

$$\check{\boldsymbol{\Psi}}' = \left(\check{\boldsymbol{\psi}}_1 \cdots \check{\boldsymbol{\psi}}_T\right), \ \check{\boldsymbol{\psi}}_t = \boldsymbol{\psi}_t - \bar{\boldsymbol{\psi}}_t, \ \boldsymbol{\psi}_t = \text{vech}\boldsymbol{\Omega}_t.$$

Writing $g = \frac{1}{2}n\left(n+1\right)$ and $p^* = \sum p_i$, vec$\left(\check{\boldsymbol{\Psi}}'\right)$ is of dimension $(gT \times 1)$ and $\boldsymbol{\alpha}$ is $(p^* \times 1)$. Under the null hypothesis of homoscedasticity:

$$\mathsf{E}\left[\boldsymbol{\epsilon}\right] = \mathbf{0}, \ \mathsf{V}\left[\boldsymbol{\epsilon}\right] = \mathsf{V}\left[\boldsymbol{\psi}_t\right] \otimes \boldsymbol{I}_T.$$

Kelejian suggested to test for $\hat{\boldsymbol{\alpha}} = 0$ in (11) using:

$$\widehat{\mathsf{V}\left[\boldsymbol{\psi}_t\right]} = \frac{1}{T} \check{\boldsymbol{\Psi}}' \check{\boldsymbol{\Psi}}. \tag{12}$$

Using a common set of regressors: $\check{\boldsymbol{P}}_i = \check{\boldsymbol{P}}$, it is easy to show that the test procedures simplifies to a multivariate regression of $\check{\boldsymbol{\Psi}}$ on $\check{\boldsymbol{P}}$. Writing $\check{\boldsymbol{P}}^* = \boldsymbol{I}_r \otimes \check{\boldsymbol{P}}$, $\check{\boldsymbol{P}}$ is $(T \times h)$ we find:

$$\hat{\boldsymbol{\alpha}} = \left\{ \left(\boldsymbol{I}_r \otimes \check{\boldsymbol{P}}\right)' \left(\widehat{\mathsf{V}\left[\boldsymbol{\psi}_t\right]} \otimes \boldsymbol{I}_T\right)^{-1} \left(\boldsymbol{I}_r \otimes \check{\boldsymbol{P}}\right) \right\}^{-1} \left(\boldsymbol{I}_r \otimes \check{\boldsymbol{P}}\right)' \left(\widehat{\mathsf{V}\left[\boldsymbol{\psi}_t\right]} \otimes \boldsymbol{I}_T\right)^{-1} \text{vec}\left(\check{\boldsymbol{\Psi}}'\right),$$

and

$$\begin{aligned} \hat{\boldsymbol{\alpha}}' \left(\check{\boldsymbol{P}}^{*\prime} \widehat{\mathsf{V}\left[\boldsymbol{\epsilon}\right]} \check{\boldsymbol{P}}^*\right) \hat{\boldsymbol{\alpha}} &= \text{vec}\left(\check{\boldsymbol{\Psi}}'\right)' \left\{ \widehat{\mathsf{V}\left[\boldsymbol{\psi}_t\right]}^{-1} \otimes \check{\boldsymbol{P}} \left(\check{\boldsymbol{P}}'\check{\boldsymbol{P}}\right)^{-1} \check{\boldsymbol{P}}' \right\} \text{vec}\left(\check{\boldsymbol{\Psi}}'\right) \\ &= \text{tr}\left\{ \widehat{\mathsf{V}\left[\boldsymbol{\psi}_t\right]}^{-1} \check{\boldsymbol{\Psi}}' \check{\boldsymbol{P}} \left(\check{\boldsymbol{P}}'\check{\boldsymbol{P}}\right)^{-1} \check{\boldsymbol{P}}' \check{\boldsymbol{\Psi}} \right\} \qquad (13) \\ &= T\,\text{tr}\left\{ \left(\check{\boldsymbol{\Psi}}'\check{\boldsymbol{\Psi}}\right)^{-1} \check{\boldsymbol{\Psi}}' \check{\boldsymbol{P}} \left(\check{\boldsymbol{P}}'\check{\boldsymbol{P}}\right)^{-1} \check{\boldsymbol{P}}' \check{\boldsymbol{\Psi}} \right\}. \end{aligned}$$

The last line uses (12). The test statistic is asymptotically $\chi^2(gh)$ distributed.

So, writing (11) in system form, under the assumption (12):

$$\check{\boldsymbol{\Psi}}' = \boldsymbol{\beta}\check{\boldsymbol{P}}' + \boldsymbol{E}' \tag{14}$$

with vec$(\boldsymbol{\beta}') = \boldsymbol{\alpha}$, shows that (13) is an LM test for vec$(\boldsymbol{\beta}') = \mathbf{0}$ in this system. This test may be written as $LM = TgR_m^2$. If $\boldsymbol{p}_t$ (row t of $\boldsymbol{P}$) consists of the original w_{it}'s and their squares and cross-products, this test is the system analogue of White (1980)'s test for heteroscedasticity. For numerical accuracy reasons, it might be advisable to estimate the system (14) as it stands: first partial out the constant term, and then estimate the system (or better, use a QR-based method).

Analogous to the vector autocorrelation test, we use the F-approximation (7):

$$LMF = \frac{1-\left(1-R_r^2\right)^{1/r}}{\left(1-R_r^2\right)^{1/r}}.\frac{Nr-q}{gh} \underset{app}{\widetilde{\ \ }} F(gh, Ns-q), \tag{15}$$

with k, n, p in (7) replaced by k_1, g, h:

k_1	degrees of freedom in restricted auxiliary system, see Table 1
g	dimension of auxiliary system,
h	number of regressors tested for exclusion.

Assuming that $\boldsymbol{w}_t$ includes a column of ones for the intercept, and none of the w_{it}'s are redundant when squared, we have $h = 2(k-1)$ for the heteroscedasticity test, and $h = \frac{1}{2}k(k-1) + 2(k-1)$ for the form involving all cross-products. Since n is the dimensionality of the original system: $g = \frac{1}{2}n(n+1)$.

We may wish to consider a variant of this test. Under the null hypothesis of multivariate normality, we can transform the residuals to independent normal. Write $\boldsymbol{\Omega} = \boldsymbol{T}\boldsymbol{T}'$ and $\boldsymbol{v}_t \sim \mathsf{N}(\mathbf{0}, \boldsymbol{\Omega})$ then $\boldsymbol{u}_t = \boldsymbol{T}^{-1}\boldsymbol{v}_t \sim \mathsf{N}(\mathbf{0}, \boldsymbol{I})$. So the tests could be based on the transformed residuals $\boldsymbol{u}_t$, omitting the cross-products, and hence reducing the dimensionality of the test. For $\boldsymbol{T}$ we use a symmetric square root as in Doornik and Hansen (1994); the resulting tests are labelled *ZLM* and *ZLMF*.

In all, we consider 10 forms of the heteroscedasticity test, listed in Table 1. The last two entries in the table also subtract the degrees of freedom of the system which is tested for heteroscedasticity. This is the procedure used in PcGive.

Table 1. Forms of the heteroscedasticity test.

	form	g	k_1	h
HET	LM	$\frac{1}{2}n(n+1)$	1	$2(k-1)$
$HETX$	LM	$\frac{1}{2}n(n+1)$	1	$\frac{1}{2}k(k-1)+2(k-1)$
$ZHET$	ZLM	n	1	$2(k-1)$
$ZHETX$	ZLM	n	1	$\frac{1}{2}k(k-1)+2(k-1)$
$HET\text{-}F$	LMF	$\frac{1}{2}n(n+1)$	1	$2(k-1)$
$HETX\text{-}F$	LMF	$\frac{1}{2}n(n+1)$	1	$\frac{1}{2}k(k-1)+2(k-1)$
$ZHET\text{-}F$	$ZLMF$	n	1	$2(k-1)$
$ZHETX\text{-}F$	$ZLMF$	n	1	$\frac{1}{2}k(k-1)+2(k-1)$
$GIV\text{-}F$	LMF	$\frac{1}{2}n(n+1)$	$1+k$	$2(k-1)$
$GIVX\text{-}F$	LMF	$\frac{1}{2}n(n+1)$	$1+k$	$\frac{1}{2}k(k-1)+2(k-1)$

6 Monte Carlo design

The Monte Carlo design is based on the n-variate version of the PcNaive data generation process (DGP), see Hendry, Neale and Ericsson (1991):

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{A}_0 \boldsymbol{y}_t + \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \boldsymbol{A}_2 \boldsymbol{z}_t + \boldsymbol{u}_t, \\ \boldsymbol{u}_t &= \boldsymbol{B}_0 \boldsymbol{u}_{t-1} + \boldsymbol{e}_t + \boldsymbol{B}_1 \boldsymbol{e}_{t-1}, \\ \boldsymbol{z}_t &= \boldsymbol{C}_0 \boldsymbol{z}_{t-1} + \boldsymbol{c}_1 + \boldsymbol{c}_2 t + \boldsymbol{v}_t. \end{aligned} \tag{16}$$

The vectors $\boldsymbol{y}_t, \boldsymbol{u}_t, \boldsymbol{e}_t$ are $n \times 1$, so that the coefficient matrices $\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{B}_0$ are $n \times n$. $\boldsymbol{I} - \boldsymbol{A}_0$ corresponds to $\boldsymbol{B}$ in (8). The $\boldsymbol{z}_t$ vector is $q \times 1$. The zs are always fixed for each experiment, and 20 initial observations of the DGP are discarded (19 or 18 when one or two lags of $\boldsymbol{y}$ are used in the model).

All experiments have a DGP with $n = q = 3$, and always have $\boldsymbol{A}_0 = \boldsymbol{0}$, $\boldsymbol{C}_0 = \boldsymbol{\Gamma}_1$, $\boldsymbol{c}_1 = \boldsymbol{c}_2 = \boldsymbol{0}$ and $\boldsymbol{v}_t \sim \mathsf{N}(\boldsymbol{0}, \boldsymbol{I})$, $\boldsymbol{e}_t \sim \mathsf{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_1)$. The on coefficient matrices $\boldsymbol{\Gamma}_i$ are:

$$\boldsymbol{\Gamma}_0 = \begin{pmatrix} 0.3 & 0.1 & 0.1 \\ 0.1 & 0.3 & 0.1 \\ 0.1 & 0.1 & 0.3 \end{pmatrix}, \quad \boldsymbol{\Gamma}_1 = \begin{pmatrix} 0.5 & 0.1 & 0 \\ 0.1 & 0.5 & 0 \\ 0 & 0 & 0.3 \end{pmatrix}, \quad \boldsymbol{\Gamma}_2 = \begin{pmatrix} 0.5 & 0.3 & 0 \\ 0.3 & 0.7 & 0 \\ 0 & 0 & 0.7 \end{pmatrix}.$$

Table 2 summarizes the DGPs and models used in the Monte Carlo experiments for the autocorrelation tests. The DGPs in Table 2 are stationary: all the eigenvalues of the companion matrix are inside the unit circle.

Table 2. Design of experiments for the autocorrelation test.

case	DGP	Model	description
(a)	$\boldsymbol{A}_1 = \boldsymbol{A}_2 = \boldsymbol{B}_0 = \boldsymbol{B}_1 = \boldsymbol{0}$	$\boldsymbol{1}$	$\boldsymbol{y}_t$ white noise
(b)	$\boldsymbol{A}_1 = \boldsymbol{0}, \boldsymbol{A}_2 = \boldsymbol{I}, \boldsymbol{B}_0 = \boldsymbol{B}_1 = \boldsymbol{0}$	$\boldsymbol{1}, \boldsymbol{z}_t$	no dynamics
(c)	$\boldsymbol{A}_1 = \boldsymbol{\Gamma}_0, \boldsymbol{A}_2 = \boldsymbol{B}_0 = \boldsymbol{B}_1 = \boldsymbol{0}$	$\boldsymbol{1}, \boldsymbol{y}_{t-1}$	VAR(1)
(d)	as (c)	$\boldsymbol{1}, \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}$	overspecified dynamics
(e)	$\boldsymbol{A}_1 = \boldsymbol{\Gamma}_0, \boldsymbol{A}_2 = \boldsymbol{I}, \boldsymbol{B}_0 = \boldsymbol{B}_1 = \boldsymbol{0}$	$\boldsymbol{1}, \boldsymbol{y}_{t-1}, \boldsymbol{z}_t$	VARX(1)
(f)	as (e)	$\boldsymbol{1}, \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \boldsymbol{z}_t$	overspecified dynamics
(g)	$\boldsymbol{A}_1 = \boldsymbol{\Gamma}_0, \boldsymbol{A}_2 = \boldsymbol{0}, \boldsymbol{B}_0 = \boldsymbol{\Gamma}_1, \boldsymbol{B}_1 = \boldsymbol{0}$	$\boldsymbol{1}, \boldsymbol{y}_{t-1}$	power: autocorrelation
(h)	$\boldsymbol{A}_1 = \boldsymbol{\Gamma}_0, \boldsymbol{A}_2 = \boldsymbol{0}, \boldsymbol{B}_0 = \boldsymbol{\Gamma}_2, \boldsymbol{B}_1 = \boldsymbol{0}$	$\boldsymbol{1}, \boldsymbol{y}_{t-1}$	power: autocorrelation
(i)	$\boldsymbol{A}_1 = \boldsymbol{\Gamma}_0, \boldsymbol{A}_2 = \boldsymbol{0}, \boldsymbol{B}_0 = \boldsymbol{0}, \boldsymbol{B}_1 = \boldsymbol{\Gamma}_1$	$\boldsymbol{1}, \boldsymbol{y}_{t-1}$	power: moving average

The heteroscedasticity experiments use cases (b) and (e) from Table 2 for size. The power experiments are based on case (h), and on case (e) but with ARCH or heteroscedastic errors:

$$\begin{aligned} &(e1) \quad \boldsymbol{e}_t \sim \mathsf{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_1 \boldsymbol{e}_{t-1} \boldsymbol{e}'_{t-1} \boldsymbol{\Gamma}'_1) && \text{ARCH,} \\ &(e2) \quad \boldsymbol{e}_t \sim \mathsf{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 \boldsymbol{e}_{t-1} \boldsymbol{e}'_{t-1} \boldsymbol{\Gamma}'_2) && \text{ARCH,} \\ &(e3) \quad \boldsymbol{e}_t \sim \mathsf{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_1 \boldsymbol{y}_{t-1} \boldsymbol{y}'_{t-1} \boldsymbol{\Gamma}'_1) && \text{heteroscedasticity.} \end{aligned}$$

In the second set of experiments (autocorrelation tests only), we use the DGP on which the tutorials for PcFiml are based (see Doornik and Hendry, 1994a, pp. 132–133). In terms of (16) this corresponds for the $\boldsymbol{y}$-equation to:

$$\boldsymbol{A}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & -0.3 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{A}_1 = \begin{pmatrix} 0.9 & 0 & 0 & 0.1 \\ 0 & 0.75 & -0.25 & 0 \\ 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{A}_2 = \boldsymbol{I},$$

and for the $\boldsymbol{z}$-equation:

$$\boldsymbol{C}_0 = \mathbf{0}, \quad \boldsymbol{c}_1' = (0.01, 0.02, 0.02, 0.01), \quad \boldsymbol{c}_2' = (\beta_0, 0, 0, 0),$$

with error distributions:

$$\boldsymbol{B}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \beta_1 & 0 & 0 \\ 0 & 0 & \beta_2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \boldsymbol{B}_1 = \mathbf{0}, \quad \boldsymbol{e}_t \sim \mathsf{N}\left[\mathbf{0}, \begin{pmatrix} 0.06 & 0 & 0 & 0 \\ 0 & 0.015 & 0.057 & 0 \\ 0 & 0.057 & 0.05 & 0 \\ 0 & 0 & 0 & 0.15 \end{pmatrix}\right], \quad \boldsymbol{v}_t = \mathbf{0}.$$

Table 3 specifies the values chosen for β_i in the DGP, and formulates the specifications used to model the DGP. M_1 is the unrestricted reduced form, which can be estimated by multivariate least squares. M_2 is a simultaneous equations model which has M_1 as its URF, and will be estimated by FIML. The DGPs in Table 3 have two roots on the unit circle.

Table 3. DGPs and models (M_1, M_2) for the second group of experiments.

$case$	DGP	Model	description
(j)	$\beta_0 = 0.004, \beta_1 = \beta_2 = 0$	M_1	size
(k)	as (i)	M_2	size
(l)	$\beta_0 = 0.004, \beta_1 = 0.5, \beta_2 = 0.4$	M_1	power
(m)	as (k)	M_2	power
(n)	$\beta_0 = 0.004, \beta_1 = 0.25, \beta_2 = 0.2$	M_1	power
(o)	as (m)	M_2	power

$$(M_1) \quad \boldsymbol{y}_t = \boldsymbol{\Pi}_1 \boldsymbol{y}_{t-1} + \boldsymbol{\alpha} + \boldsymbol{\beta} t + \boldsymbol{u}_t$$

$$(M_2) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ \gamma_1 & 1 & \gamma_2 & 0 \\ 0 & 0 & 1 & \gamma_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{y}_t = \begin{pmatrix} \gamma_4 & 0 & 0 & \gamma_5 \\ 0 & \gamma_6 & \gamma_7 & 0 \\ 0 & \gamma_8 & \gamma_9 & 0 \\ 0 & 0 & 0 & \gamma_{10} \end{pmatrix} \boldsymbol{y}_{t-1} + \boldsymbol{\alpha} + \begin{pmatrix} \gamma_{11} \\ 0 \\ 0 \\ 0 \end{pmatrix} t + \boldsymbol{u}_t$$

7 MONTE CARLO RESULTS FOR THE VECTOR AUTOCORRELATION TESTS

The Monte Carlo results are all based on $M = 1000$ replications and only presented graphically[3]. The reported simulations were done in Ox version 0.50 (see Doornik, 1996), using the internal random number generator.

[3] In an attempt to circumvent a drawback of many presentations of Monte Carlo results, described by Mizon and Hendry (1980) as 'tabulation strains the memory without producing much insight'.

The first case considered is that of regressing a white noise vector on a constant term, labelled case (a). The results are in Figs. 1 and 2, showing the empirical rejection frequencies for the *LM* and *LMF* tests at sample sizes $T = 25, 50, 100$. The approximate Monte Carlo standard error, defined as $\sqrt{\{p(1-p)/M\}}$, where p is the tail probability of interest, is in this case $(0.013, 0.009, 0.007, 0.003)$ for $p = (0.2, 0.1, 0.05, 0.01)$. The tests involved are for first up to fifth order vector error autocorrelation (Fig. 1) and both forms of the portmanteau statistic (Fig. 2). We see that the size of the *LMF* test remains somewhat more constant as we increase the AR-order. Also note that the size of the small sample corrected Q^* test behaves much better for increasing lag length. This is all the more important because these tests are usually performed with long lag lengths.

In cases (b)–(f), the *LM* and *LMF* tests behave in a similar fashion across cases. As a typical example we present in Fig. 3 the results of case (e), a first order VAR with 3 exogenous variables. The size of the F-type *LMF*-test remains constant and remarkably close to the nominal size. In contrast, the size of the χ^2 form has a tendency to 'walk away', while already starting above the nominal size. The same happens somewhat less pronounced in (b)–(d), but more severely in (e). There, at a sample sizes of $T = 25, 50, 100$, the $LM(1)$ test rejects $41\%, 16\%, 9\%$ at a nominal size of 5%. Fortunately, as so often the case for such χ^2 tests, the empirical size tends to be closest to the nominal size in the neighbourhood of 5%. The portmanteau test is valid in cases (b) and (c), and the results are similar to (a), Figure 4 only shows $T = 50$. In (d) both Q statistics start at a higher level, but drop down to their nominal size. The test is not valid in the presence of exogenous variables, and this is borne out by the third column in Fig. 4. This mimics the well known result about the single equation Box-Pierce and Ljung-Box statistics, referred to in §4.

Figure 5 presents the power at nominal rejection frequencies for cases (g), (h) and (i). Here all tests behave in a similar way, but note the nominal critical values for *LM* were overrejecting. The good small sample properties of *LMF* are not at the expense of power.

Even though the DGP for cases (j)–(o) are in I(1) space, the size of the tests is largely unaffected: the innovation process $\boldsymbol{u}_t$ is isomorphic under transformation to equilibrium correction form, both in the original and the auxiliary regression. *LM* is still overrejecting, see Fig. 6. *LMF* is clearly preferable in size terms, and power is only presented for the latter in Fig. 6.

8 Non-centrality and the effect of dimension

Mizon and Hendry (1980) showed that a useful complement to Monte Carlo experiments for tests of dynamic specification is provided by computing the asymptotic power function of the tests. This approach was also used by Godfrey (1981). It is achieved by computing the non-centrality of the χ^2–distribution. We consider the *LM* test of the hypothesis (4) in the system (3) with first order autocorrelation ($s = 1$).

Write $\boldsymbol{\theta} = \mathrm{vec}(\boldsymbol{R}_1')$ and $\boldsymbol{\theta}_p$ for the population value of $\boldsymbol{\theta}$. The LM test will have a $\chi^2\left(n^2\right)$ distribution under $\mathsf{H}_0 : \boldsymbol{\theta}_p = \mathbf{0}$. When H_0 is false, $\boldsymbol{\theta}_p \neq \mathbf{0}$, under a sequence of local alternatives around **0**, *LM* will have a non-central χ^2–distribution, see e.g. Hendry (1995, §13.2–13.3):

$$LM \sim \chi^2\left(n^2, T\delta\right).$$

The asymptotic formula for the non-centrality is derived in the Appendix for the Monte Carlo experiments (g) and (h), which are used for power simulations. The obtained values are:

$$\begin{array}{ll} \text{(g)} & \delta = 0.080, \\ \text{(h)} & \delta = 0.212. \end{array}$$

The rejection frequencies at 5% from the non-central χ^2 distribution are graphed in Fig. 7 as P_{as}. In comparison, the empirical rejection frequencies from the Monte Carlo are given as lines P_{mc}. The lines are close together, and the theoretical results correspond to the obtained Monte Carlo results. The small discrepancy, which decreases as T gets larger, is due to the fact that nominal critical values are used in the Monte Carlo experiment.

Another issue is the effect of dimension on the asymptotic power. Here, both T and n are varied. Figure 8 shows the asymptotic power, for $n = 3(1)7$ when $T = 50$ and for $n = 3(1)10$ when $T = 100$, using a DGP similar to that in cases (g) and (h):

$$
\begin{array}{ll}
(h1) & \boldsymbol{A}_1 = 0.2\boldsymbol{I} + 0.1\boldsymbol{E}, \boldsymbol{A}_2 = \mathbf{0}, \boldsymbol{B}_0 = 0.6\boldsymbol{I} + 0.1\boldsymbol{E}, \boldsymbol{B}_1 = \mathbf{0} \\
(h2) & \boldsymbol{A}_1 = 0.3\boldsymbol{I}, \boldsymbol{A}_2 = \mathbf{0}, \boldsymbol{B}_0 = 0.5\boldsymbol{I} + 0.2\boldsymbol{E}, \boldsymbol{B}_1 = \mathbf{0} \\
(h3) & \boldsymbol{A}_1 = 0.3\boldsymbol{I}, \boldsymbol{A}_2 = \mathbf{0}, \boldsymbol{B}_0 = 0.7\boldsymbol{I}, \boldsymbol{B}_1 = \mathbf{0} \\
(h4) & \boldsymbol{A}_1 = 0.2\boldsymbol{I} + 0.1\boldsymbol{E}, \boldsymbol{A}_2 = \mathbf{0}, \boldsymbol{B}_0 = 0.5\boldsymbol{I}, \boldsymbol{B}_1 = \mathbf{0}
\end{array}
$$

where **E** is a matrix of ones. The variance of $\boldsymbol{e}_t$ is set to the identity matrix (note that the power is unaffected by the value of the variance). Figure 8 shows that when the equations are independent, as in case (h3) and less so in (h2), power is largely uneffected as the number of equations increases.[4] Whereas the solid lines in the figure are arrived from the derived non-central χ^2 distribution, the dashed lines are empirical sizes at 5% from a Monte Carlo experiment with $M = 1000$. These report the size of case (h1) when $\boldsymbol{B}_0$ is zero (i.e. in a VAR(1)). We see that the F-form is well behaved, whereas the χ^2 statistic increasingly overreject. Mudholkar and Trivedi (1980) report that LMF can also overreject, namely increasingly so in cases when the number of equations (n), or number of added regressors tested for significance (p) gets larger and total the number of included regressors is large relative to the sample size (the degrees of freedom of error are small). This behaviour is not evident in Fig. 8. However, repeating this with additional regressors ($\boldsymbol{C_0} = \boldsymbol{I}_n$) will give rejection frequencies of *LMF* up to 10%. Further evidence of such behaviour is given in Edgerton and Shukur (1995). A potential solution could be to adopt the normal approximation suggested in Mudholkar and Trivedi (1980).

9 MONTE CARLO RESULTS FOR THE VECTOR HETEROSCEDASTICITY TESTS

The simulations were done in Ox, using $M = 1000$ replications and sample sizes $T = 50, 100, 200$. The results are presented in Table 4. At $T = 50$ all the χ^2-tests are somewhat undersized, especially in case (e) when the cross-products are included. The *GIV* forms of the F-tests are seriously undersized in case (e), in contrast to *HET-F* and *HETX-F*, which overreject. Results for the normality test E_p (see Doornik and Hansen, 1994) are presented for comparison.

The large number of tests makes graphical presentation less convenient. Table 5 has the outcomes regarding power, with the rejection frequency based on the nominal critical values (which, as Table 4 showed, are still somewhat out for some tests at a sample size of 100). Not surprisingly, none of the heteroscedasticity tests have power against autocorrelation. The heteroscedasticity in case (e3) is picked up very well because $\boldsymbol{y}_1$ is one of the regressors in the auxiliary system. There is less power against ARCH for all tests; the F-tests have more power, but most of that can be ascribed to higher nominal rejection frequencies. The forms using cross-products do better than when using squares and levels only.

10 CONCLUSION

We have studied an F-version of the test for vector residual autocorrelation which has been available for some time to practitionars (in the computer program PcFiml). Initially, the tests rejected several models which were developed using single equation methods and tests. A concern was that the test might have

[4] The gaps in the lines correspond to cases where numerical instability prevented computation of the non-centrality.

wrong size, thus seriously distorting inference. However, the Monte Carlo simulations showed that the test has excellent size even in small samples, and good power properties. It is clearly to be preferred to the portmanteau type statistics, and to the asymptotic (χ^2) form, unless in larger samples than those considered here. Analytic results regarding the non-centrality confirmed that both tests have good asymptotic power.

The results for the vector heteroscedasticity test are less clear cut than for the vector error autocorrelation test. The F-forms correct less well, and even over-correct in some cases. When using the F-form, I would have a slight preference for not subtracting the degrees of freedom lost in the original system (thus preferring *HET-F* and *HETX-F* to *GIV-F* and *GIVX-F*). The *ZHET* forms, which first transform the residuals to approximate normality provide a viable alternative.

Acknowledgements

I wish to thank David Hendry, Grayham Mizon and Tom Rothenberg for helpful comments. Financial support from the UK Economic and Social Science Research Council is gratefully acknowledged.

References

Ahn, S. K. (1988). Distribution for residual autocovariances in multivariate autoregressive models with structured parameterization. *Biometrika*, **75**, 590–593.

Ali, M. M. and Giaccotto, C. (1984). A study of several new and existing tests for heteroscedasticity in the general linear model. *Journal of Econometrics*, **26**, 355–373.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis* 2nd edn. New York: John Wiley & Sons.

Box, G. E. P. and Pierce, D. A. (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. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, **17**, 334–355.

Breusch, T. S. and Godfrey, L. G. (1981). A review of recent work on testing for autocorrelation in dynamic simultaneous models. in Currie *et al.* (1981), Ch. 4.

Breusch, T. S. and Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies*, **47**, 239–253.

Currie, D. (1981). Some long run features of dynamic time series models. *Economic Journal*, **91**, 704–715.

Currie, D., Nobay, A. R. and Peel, D. (eds.)(1981). *Macroeconomic Analysis*. London: Croom Helm.

Deschamps, P. J. (1994). Monte Carlo methodology for LM and LR autocorrelation tests in multivariate regression. Mimeo, Université de Fribourg, Switzerland.

Doornik, J. A. (1996). *Object-Oriented Matrix Programming using Ox*. London: International Thomson Business Press and Oxford: http://www.nuff.ox.ac.uk/Users/Doornik/.

Doornik, J. A. and Hansen, H. (1994). A practical test for univariate and multivariate normality. Discussion paper, Nuffield College.

Doornik, J. A. and Hendry, D. F. (1994a). *PcFiml 8: An Interactive Program for Modelling Econometric Systems*. London: International Thomson Publishing.

Doornik, J. A. and Hendry, D. F. (1994b). *PcGive 8: An Interactive Econometric Modelling System*. London: International Thomson Publishing, and Belmont, CA: Duxbury Press.

Durbin, J. (1970). Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables. *Econometrica*, **38**, 410–421.

Edgerton, D. and Shukur, G. (1995). Testing autocorrelation in a system perspective. Mimeo, Lund University, Sweden.

Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, **46**, 1303–1313.

Godfrey, L. G. (1981). On the invariance of the Lagrange multiplier test with respect to certain changes in the alternative hypothesis. *Econometrica*, **49**, 1443–1455.

Godfrey, L. G. (1988). *Misspecification Tests in Econometrics*. Cambridge: Cambridge University Press.

Godfrey, L. G. and Orme, C. D. (1994). The sensitivity of some general checks to omitted variables in the linear model. *International Economic Review*, **35**, 489–506.

Guilkey, D. K. (1974). Alternative tests for a first order vector autoregressive error specification. *Journal of Econometrics*, **2**, 95–104.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F., Neale, A. J. and Ericsson, N. R. (1991). *PC-NAIVE, An Interactive Program for Monte Carlo Experimentation in Econometrics. Version 6.0*. Oxford: Institute of Economics and Statistics, University of Oxford.

Hosking, J. R. M. (1980). The multivariate portmanteau statistic. *Journal of the American Statistical Association*, **75**, 602–608.

Kelejian, H. H. (1982). An extension of a standard test for heteroskedasticity to a systems framework. *Journal of Econometrics*, **20**, 325–333.

Kiviet, J. F. (1986). On the rigor of some mis-specification tests for modelling dynamic relationships. *Review of Economic Studies*, **53**, 241–261.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.

Mizon, G. E. and Hendry, D. F. (1980). An empirical application and Monte Carlo analysis of tests of dynamic specification. *Review of Economic Studies*, **49**, 21–45. Reprinted in Hendry D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Mudholkar, G. S. and Trivedi, M. C. (1980). A normal approximation for the distribution of the likelihood ratio statistic in multivariate analysis of variance. *Biometrika*, **67**, 485–488.

Osborn, D. R. (1981). Discussion of: A review of recent work on testing for autocorrelation in dynamic simultaneous models. in Currie *et al.* (1981), Ch. 4.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications* 2nd edn. New York: John Wiley & Sons.

White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817–838.

APPENDIX

This section gives the derivations underpinning the analytical results of this chapter. The system with first order autocorrelation is:

$$\boldsymbol{Y}' = \boldsymbol{\Pi}\boldsymbol{W}' + \boldsymbol{U}', \quad \boldsymbol{U}' = \boldsymbol{R}\boldsymbol{U}'_{-1} + \boldsymbol{E}',$$

so

$$\boldsymbol{\Omega} = \boldsymbol{E}'\boldsymbol{E}, \quad \boldsymbol{E}' = \boldsymbol{Y}' - \boldsymbol{\Pi}\boldsymbol{W}' - \boldsymbol{R}\boldsymbol{Y}'_{-1} + \boldsymbol{R}\boldsymbol{\Pi}\boldsymbol{W}'_{-1}.$$

The concentrated log-likelihood function of this model is:

$$\ell_c(\boldsymbol{\Pi}, \boldsymbol{R}) = -\frac{T}{2}\log|\boldsymbol{\Omega}|.$$

The partial derivatives of the concentrated log-likelihood are:

$$\begin{aligned} g_1(\boldsymbol{\Pi}, \boldsymbol{R}) = \frac{\ell_c(\boldsymbol{\Pi}, \boldsymbol{R})}{\partial \operatorname{vec}(\boldsymbol{\Pi}')} &= \left[\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'\right]\operatorname{vec}\boldsymbol{E} - \left[\boldsymbol{R}'\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'_{-1}\right]\operatorname{vec}\boldsymbol{E}, \\ g_2(\boldsymbol{\Pi}, \boldsymbol{R}) = \frac{\ell_c(\boldsymbol{\Pi}, \boldsymbol{R})}{\partial \operatorname{vec}(\boldsymbol{R}')} &= \left[\boldsymbol{\Omega}^{-1} \otimes \left(\boldsymbol{Y}'_{-1} - \boldsymbol{\Pi}\boldsymbol{W}'_{-1}\right)\right]\operatorname{vec}\boldsymbol{E}. \end{aligned}$$

Treating $\boldsymbol{\Omega}^{-1}$ as fixed:

$$\begin{aligned} G_{11}(\boldsymbol{\Pi}, \boldsymbol{R}) = \frac{\ell_c(\boldsymbol{\Pi}, \boldsymbol{R})}{\partial \operatorname{vec}(\boldsymbol{\Pi}')\,\partial(\operatorname{vec}(\boldsymbol{\Pi}'))'} &= -\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'\boldsymbol{W} + \boldsymbol{R}'\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'_{-1}\boldsymbol{W} \\ &\quad + \boldsymbol{\Omega}^{-1}\boldsymbol{R} \otimes \boldsymbol{W}'\boldsymbol{W}_{-1} - \boldsymbol{R}'\boldsymbol{\Omega}^{-1}\boldsymbol{R} \otimes \boldsymbol{W}'_{-1}\boldsymbol{W}_{-1}, \\ G_{22}(\boldsymbol{\Pi}, \boldsymbol{R}) = \frac{\ell_c(\boldsymbol{\Pi}, \boldsymbol{R})}{\partial \operatorname{vec}(\boldsymbol{R}')\,\partial(\operatorname{vec}(\boldsymbol{R}'))'} &= -\boldsymbol{\Omega}^{-1} \otimes \left(\boldsymbol{Y}'_{-1}\boldsymbol{Y}_{-1} - \boldsymbol{\Pi}\boldsymbol{W}'_{-1}\boldsymbol{Y}_{-1}\right) \\ &\quad + \boldsymbol{\Omega}^{-1} \otimes \left(\boldsymbol{Y}'_{-1}\boldsymbol{W}_{-1}\boldsymbol{\Pi}' - \boldsymbol{\Pi}\boldsymbol{W}'_{-1}\boldsymbol{W}_{-1}\boldsymbol{\Pi}'\right), \\ G_{12}(\boldsymbol{\Pi}, \boldsymbol{R}) = \frac{\ell_c(\boldsymbol{\Pi}, \boldsymbol{R})}{\partial \operatorname{vec}(\boldsymbol{\Pi}')\,\partial(\operatorname{vec}(\boldsymbol{R}'))'} &= \boldsymbol{R}'\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'_{-1}\boldsymbol{Y}_{-1} - \boldsymbol{R}'\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'_{-1}\boldsymbol{W}_{-1}\boldsymbol{\Pi}' \\ &\quad + \boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'\boldsymbol{Y}_{-1} - \boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'\boldsymbol{W}_{-1}\boldsymbol{\Pi}'. \end{aligned}$$

Evaluating at $\boldsymbol{R} = \boldsymbol{0}$, and writing $\boldsymbol{U} = \boldsymbol{Y} - \boldsymbol{W}\boldsymbol{\Pi}'$, $\boldsymbol{U}_{-1} = \boldsymbol{Y}_{-1} - \boldsymbol{W}_{-1}\boldsymbol{\Pi}'$:

$$\begin{aligned} g_2 &= g_2(\boldsymbol{\Pi}, \boldsymbol{0}) = \left[\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{U}'_{-1}\right]\operatorname{vec}\boldsymbol{U}, \\ G_{11} &= G_{11}(\boldsymbol{\Pi}, \boldsymbol{0}) = -\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'\boldsymbol{W}, \\ G_{12} &= G_{12}(\boldsymbol{\Pi}, \boldsymbol{0}) = \boldsymbol{\Omega}^{-1} \otimes \boldsymbol{W}'\boldsymbol{U}_{-1}, \\ G_{22} &= G_{22}(\boldsymbol{\Pi}, \boldsymbol{0}) = -\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{U}'_{-1}\boldsymbol{U}_{-1}. \end{aligned} \tag{17}$$

Using partitioned inversion, $\boldsymbol{M}_W = \boldsymbol{I}_T - \boldsymbol{W}\left(\boldsymbol{W}'\boldsymbol{W}\right)^{-1}\boldsymbol{W}'$:

$$G^{22} = \left(G_{22} - G_{21}G_{11}^{-1}G_{12}\right)^{-1} = -\boldsymbol{\Omega} \otimes \left(\boldsymbol{U}'_{-1}\boldsymbol{M}_W\boldsymbol{U}_{-1}\right)^{-1}. \tag{18}$$

The hypothesis to test is H_0: $\boldsymbol{R} = \boldsymbol{0}$. Denoting the constrained estimates by $\hat{\boldsymbol{\Pi}}$ and $\hat{\boldsymbol{R}}$, we have $\hat{\boldsymbol{\Pi}}' = \left(\boldsymbol{W}'\boldsymbol{W}\right)^{-1}\boldsymbol{W}'\boldsymbol{Y}$ and $\hat{\boldsymbol{R}}' = \boldsymbol{0}$. The unconstrained estimates are $\tilde{\boldsymbol{\Pi}}$, $\tilde{\boldsymbol{R}}$. The derivatives of the unconstrained likelihood are g_1, G_{11} etc., where $\hat{g}_1 = g_1(\hat{\boldsymbol{\Pi}}, \boldsymbol{0})$, and $\tilde{g}_1 = g_1(\tilde{\boldsymbol{\Pi}}, \tilde{\boldsymbol{R}})$. In addition we write $\boldsymbol{\theta}_1 = \operatorname{vec}\boldsymbol{\Pi}'$ and $\boldsymbol{\theta}_2 = \operatorname{vec}\boldsymbol{R}'$, so that the null hypothesis can be expressed as H_0: $\boldsymbol{\theta}_2 = \boldsymbol{0}$. The LM test uses derivatives of the unrestricted likelihood evaluated at restricted parameter estimates:

$$LM = -\hat{g}_2'\hat{G}^{22}\hat{g}_2.$$

Using (17) and (18) this equals:

$$LM = T\operatorname{tr}\left\{\left(\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}\right)^{-1}\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}_{-1}\left(\hat{\boldsymbol{U}}'_{-1}\boldsymbol{M}_W\hat{\boldsymbol{U}}_{-1}\right)^{-1}\hat{\boldsymbol{U}}'_{-1}\hat{\boldsymbol{U}}\right\}.$$

Godfrey (1981) showed in a more general setting that this is identical to an LM test on the lagged residuals in an auxiliary regression:

$$\boldsymbol{Y}' = \boldsymbol{\Pi}\boldsymbol{W}' + \boldsymbol{R}\hat{\boldsymbol{U}}'_{-1} + \boldsymbol{E}',$$

with $\hat{\boldsymbol{U}} = \boldsymbol{M}_W \boldsymbol{Y}$. In the absence of $\boldsymbol{W}$s we find in *LM* the first term of the vector portmanteau statistic:

$$T\,\mathrm{tr}\left(\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}\right)^{-1}\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}_{-1}\left(\hat{\boldsymbol{U}}_{-1}'\hat{\boldsymbol{U}}_{-1}\right)^{-1}\hat{\boldsymbol{U}}_{-1}'\hat{\boldsymbol{U}} \approx T\,\mathrm{tr}\,\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}_{-1}\left(\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}\right)^{-1}\hat{\boldsymbol{U}}_{-1}'\hat{\boldsymbol{U}}\left(\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}\right)^{-1}$$

The non-centrality of the *LM* test with population parameter $\boldsymbol{\theta}_p$ is:

$$\psi^2 = -\boldsymbol{\theta}_p'\left(G^{22}\right)^{-1}\boldsymbol{\theta}_p = T\boldsymbol{\theta}_p'\left\{\boldsymbol{\Omega}^{-1}\otimes T^{-1}\left(\boldsymbol{U}_{-1}'\boldsymbol{M}_W\boldsymbol{U}_{-1}\right)\right\}\boldsymbol{\theta}_p.$$

To find the relevant probability limits it is required to solve for the DGP, for which we closely follow Hendry *et al.* (1991, Ch. 3). In the VARX(1) case with first-order error autocorrelation the DGP is specified as:

$$\begin{array}{rcll} \boldsymbol{D}\boldsymbol{y}_t &=& \boldsymbol{A}_1\boldsymbol{y}_{t-1} + \boldsymbol{A}_2\boldsymbol{z}_t + \boldsymbol{u}_t, & \boldsymbol{D} = \boldsymbol{I} - \boldsymbol{A}_0, \\ \boldsymbol{u}_t &=& \boldsymbol{B}_0\boldsymbol{u}_{t-1} + \boldsymbol{e}_t, & \boldsymbol{e}_t \sim \mathsf{N}\left(\mathbf{0}, \boldsymbol{E}_1\right), \\ \boldsymbol{z}_t &=& \boldsymbol{C}_0\boldsymbol{z}_{t-1} + \boldsymbol{c}_1 + \boldsymbol{c}_2 t + \boldsymbol{v}_t, & \boldsymbol{v}_t \sim \mathsf{N}\left(\mathbf{0}, \boldsymbol{E}_2\right), \end{array}$$

The companion form of this DGP is (the first equation lagged is used to substitute out $\boldsymbol{u}_t$):

$$\boldsymbol{x}_t = \boldsymbol{\Pi}\boldsymbol{x}_{t-1} + \boldsymbol{\alpha} + \boldsymbol{\beta}t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \mathsf{N}\left(\mathbf{0}, \boldsymbol{\Sigma}\right),$$

where $\boldsymbol{x}_t$ is the $2n+q=m$ vector $(\boldsymbol{y}_t' : \boldsymbol{y}_{t-1}' : \boldsymbol{z}_t')'$, $\boldsymbol{\nu}_t = (\boldsymbol{\epsilon}_t' : \boldsymbol{0}' : \boldsymbol{v}_t')'$, and both $\boldsymbol{\Pi}$ and $\boldsymbol{\Sigma}$ are $m \times m$; in terms of the original matrices:

$$\begin{array}{rcl} \boldsymbol{\Pi} &=& \begin{pmatrix} \boldsymbol{D}^{-1}\left(\boldsymbol{A}_1 + \boldsymbol{B}_0\boldsymbol{D}\right) & -\boldsymbol{D}^{-1}\boldsymbol{B}_0\boldsymbol{A}_1 & \boldsymbol{D}^{-1}\left(\boldsymbol{A}_2\boldsymbol{C}_0 - \boldsymbol{B}_0\boldsymbol{A}_2\right) \\ \boldsymbol{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{C}_0 \end{pmatrix}, \\ \boldsymbol{\Sigma} &=& \begin{pmatrix} \boldsymbol{D}^{-1}\left(\boldsymbol{A}_2\boldsymbol{E}_2\boldsymbol{A}_2' + \boldsymbol{E}_1\right)\boldsymbol{D}^{-1\prime} & \mathbf{0} & \boldsymbol{D}^{-1}\boldsymbol{A}_2\boldsymbol{E}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{E}_2\boldsymbol{A}_2'\boldsymbol{D}^{-1\prime} & \mathbf{0} & \boldsymbol{E}_2 \end{pmatrix}. \end{array}$$

Next, when $\boldsymbol{\alpha} = \boldsymbol{\beta} = \mathbf{0}$, and under stationarity:

$$\mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{\nu}_t'\right] = \boldsymbol{\Sigma}, \quad \mathsf{E}\left[\boldsymbol{x}_{t-1}\boldsymbol{\nu}_t'\right] = \mathbf{0},$$

so that

$$\begin{array}{rcl} \mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right] &=& \boldsymbol{\Pi}\,\mathsf{E}\left[\boldsymbol{x}_{t-1}\boldsymbol{x}_t'\right] + \boldsymbol{\Sigma}, \\ \mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_{t-1}'\right] &=& \boldsymbol{\Pi}\,\mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right], \\ \mathsf{E}\left[\boldsymbol{x}_{t-1}\boldsymbol{x}_{t-1}'\right] &=& \mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right]. \end{array}$$

The solution can be found by vectorizing

$$\mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right] = \boldsymbol{\Pi}\,\mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right]\boldsymbol{\Pi}' + \boldsymbol{\Sigma}$$

to give

$$\begin{array}{rcl} \mathrm{vec}\,\mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right] &=& \left(\boldsymbol{I}_{m^2} - \boldsymbol{\Pi}\otimes\boldsymbol{\Pi}\right)^{-1}\mathrm{vec}\,\boldsymbol{\Sigma}, \\ \mathrm{vec}\,\mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_{t-1}'\right] &=& \left(\boldsymbol{I}_m\otimes\boldsymbol{\Pi}\right)\mathrm{vec}\,\mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right]. \end{array}$$

We can now work out the required probability limits under the assumptions that $\boldsymbol{D} = \boldsymbol{I}_n$, $\boldsymbol{A}_2 = \mathbf{0}$, $\boldsymbol{E}_2 = \boldsymbol{I}_q$. Dropping $\boldsymbol{z}_t$ altogether:

$$\begin{array}{rcccl} \mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_t'\right] &=& \begin{pmatrix} \mathsf{E}\left[\boldsymbol{y}_t\boldsymbol{y}_t'\right] & \mathsf{E}\left[\boldsymbol{y}_t\boldsymbol{y}_{t-1}'\right] \\ \mathsf{E}\left[\boldsymbol{y}_{t-1}\boldsymbol{y}_t'\right] & \mathsf{E}\left[\boldsymbol{y}_{t-1}\boldsymbol{y}_{t-1}'\right] \end{pmatrix} &=& \begin{pmatrix} \mathsf{M}_{00} & \mathsf{M}_{01} \\ \mathsf{M}_{10} & \mathsf{M}_{11} \end{pmatrix}, \\ \mathsf{E}\left[\boldsymbol{x}_t\boldsymbol{x}_{t-1}'\right] &=& \begin{pmatrix} \mathsf{E}\left[\boldsymbol{y}_t\boldsymbol{y}_{t-1}'\right] & \mathsf{E}\left[\boldsymbol{y}_t\boldsymbol{y}_{t-2}'\right] \\ \mathsf{E}\left[\boldsymbol{y}_{t-1}\boldsymbol{y}_{t-1}'\right] & \mathsf{E}\left[\boldsymbol{y}_{t-1}\boldsymbol{y}_{t-2}'\right] \end{pmatrix} &=& \begin{pmatrix} \mathsf{M}_{01} & \mathsf{M}_{02} \\ \mathsf{M}_{11} & \mathsf{M}_{12} \end{pmatrix}, \end{array}$$

and

$$\boldsymbol{\Pi} = \begin{pmatrix} \boldsymbol{A}_1 + \boldsymbol{B}_0 & -\boldsymbol{B}_0\boldsymbol{A}_1 \\ \boldsymbol{I}_n & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{E}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

We used $\boldsymbol{Y}_{-1}$ as a regressor, so replacing $\boldsymbol{W}$ with $\boldsymbol{Y}_{-1}$ yields for the probability limits:

$$\begin{aligned}
\text{plim}_{T\to\infty}\,\hat{\boldsymbol{\Pi}} &= \mathsf{M}_{01}\mathsf{M}_{11}^{-1} = \boldsymbol{\Pi}_0, \\
\text{plim}_{T\to\infty}\,T^{-1}\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}} &= \mathsf{M}_{00} - \mathsf{M}_{01}\boldsymbol{\Pi}_0' = \boldsymbol{\Omega}_0, \\
\text{plim}_{T\to\infty}\,T^{-1}\hat{\boldsymbol{U}}_{-1}'\hat{\boldsymbol{U}}_{-1} &= \mathsf{M}_{11} - \mathsf{M}_{12}\boldsymbol{\Pi}_0' - \mathsf{P}\mathsf{M}_{12}' + \mathsf{P}\mathsf{M}_{00}\boldsymbol{\Pi}_0', \\
\text{plim}_{T\to\infty}\,T^{-1}\hat{\boldsymbol{U}}_{-1}'\boldsymbol{Y}_{-1} &= \mathsf{M}_{11} - \mathsf{P}\mathsf{M}_{12}', \\
\text{plim}_{T\to\infty}\,T^{-1}\hat{\boldsymbol{U}}'\hat{\boldsymbol{U}}_{-1} &= \mathsf{M}_{01} - \mathsf{M}_{12}\boldsymbol{\Pi}_0' - \mathsf{P}\mathsf{M}_{11}' + \mathsf{P}\mathsf{M}_{12}\boldsymbol{\Pi}_0'.
\end{aligned}$$

For the non-centrality, with $\boldsymbol{\theta} = \text{vec}\,(\boldsymbol{R}')$, first solve:

$$\begin{aligned}
\text{plim}\;\hat{\boldsymbol{U}}_{-1}'\boldsymbol{M}_{Y_{-1}}\hat{\boldsymbol{U}}_{-1} &= \boldsymbol{\Pi}_0\left(\mathsf{M}_{00} - \mathsf{M}_{12}'\mathsf{M}_{11}^{-1}\mathsf{M}_{12}\right)\boldsymbol{\Pi}_0', \\
\text{plim}\;\boldsymbol{Y}'\boldsymbol{M}_{Y_{-1}}\hat{\boldsymbol{U}}_{-1} &= -\mathsf{M}_{02}\boldsymbol{\Pi}_0' + \mathsf{M}_{01}\mathsf{M}_{11}^{-1}\mathsf{M}_{12}\boldsymbol{\Pi}_0'.
\end{aligned}$$

This allows us to find the plim of $\hat{\boldsymbol{R}}$. Remember that $\hat{\boldsymbol{R}}$ is the coefficient on $\hat{\boldsymbol{U}}_{-1}$ in a regression of $\boldsymbol{Y}$ on $\boldsymbol{Y}_{-1}$ and $\hat{\boldsymbol{U}}_{-1}$, so that:

$$\underset{T\to\infty}{\text{plim}}\;\hat{\boldsymbol{R}} = \text{plim}\;\boldsymbol{Y}'\boldsymbol{M}_{Y_{-1}}\hat{\boldsymbol{U}}_{-1}\left(\text{plim}\;\hat{\boldsymbol{U}}_{-1}'\boldsymbol{M}_{Y_{-1}}\hat{\boldsymbol{U}}_{-1}\right)^{-1} = \mathsf{R}.$$

Under stationarity $\mathsf{M}_{11} = \mathsf{M}_{00}$ and $\mathsf{M}_{12} = \mathsf{M}_{01}$, so, writing $\mathsf{r} = \text{vec}(\mathsf{R}')$ we finally find for the non-centrality:

$$\psi^2 = T\mathsf{r}'\left\{\boldsymbol{\Omega}_0^{-1} \otimes \left[\boldsymbol{\Pi}_0\left(\mathsf{M}_{00} - \mathsf{M}_{10}\mathsf{M}_{00}^{-1}\mathsf{M}_{01}\right)\boldsymbol{\Pi}_0'\right]\right\}\mathsf{r}.$$

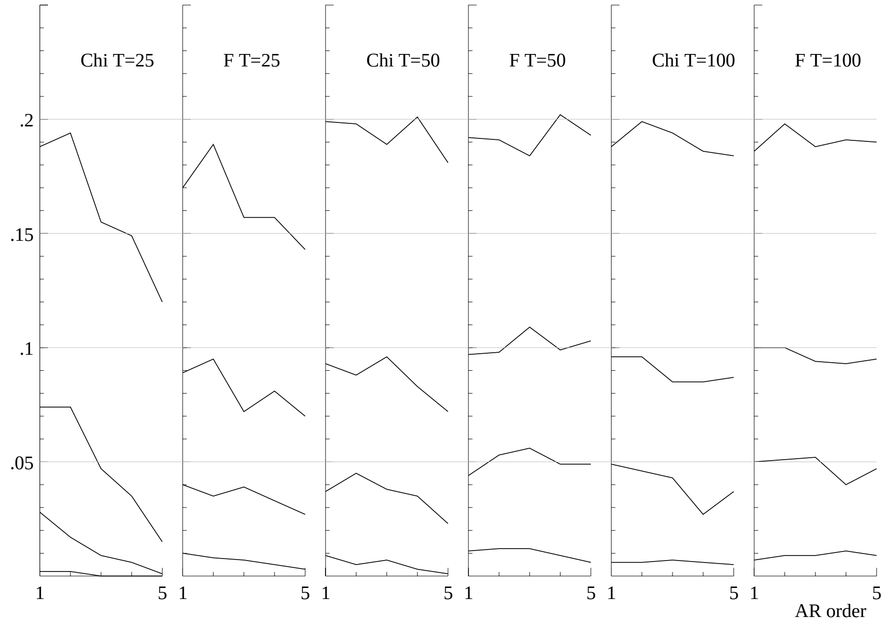

Figure 1. Rejection frequencies for *LM*,*LMF*, case (a): $\boldsymbol{y}_t$ white noise.

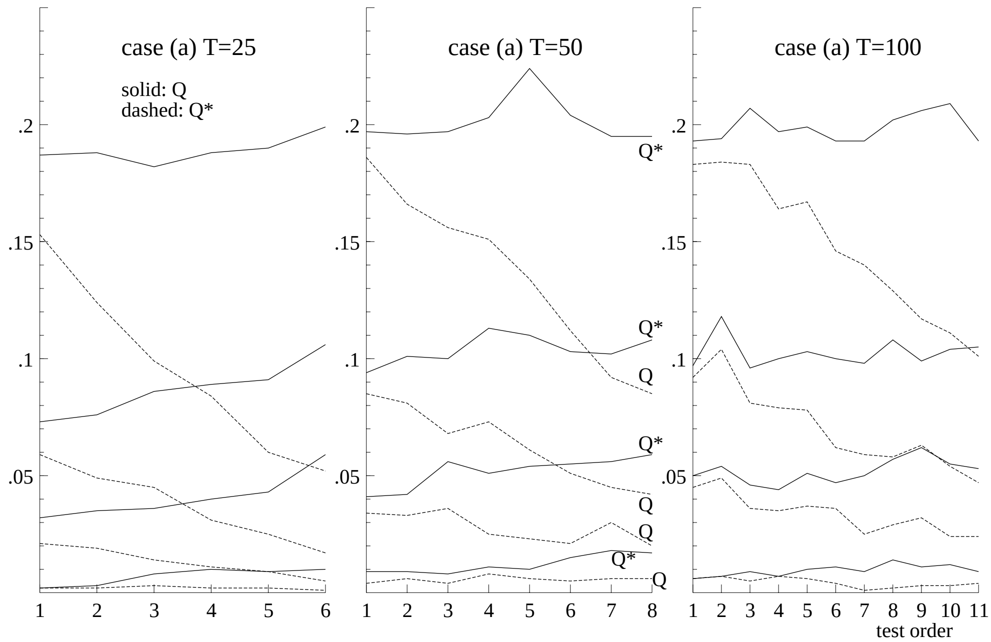

Figure 2. Rejection frequencies for Q,Q^*: case (a).

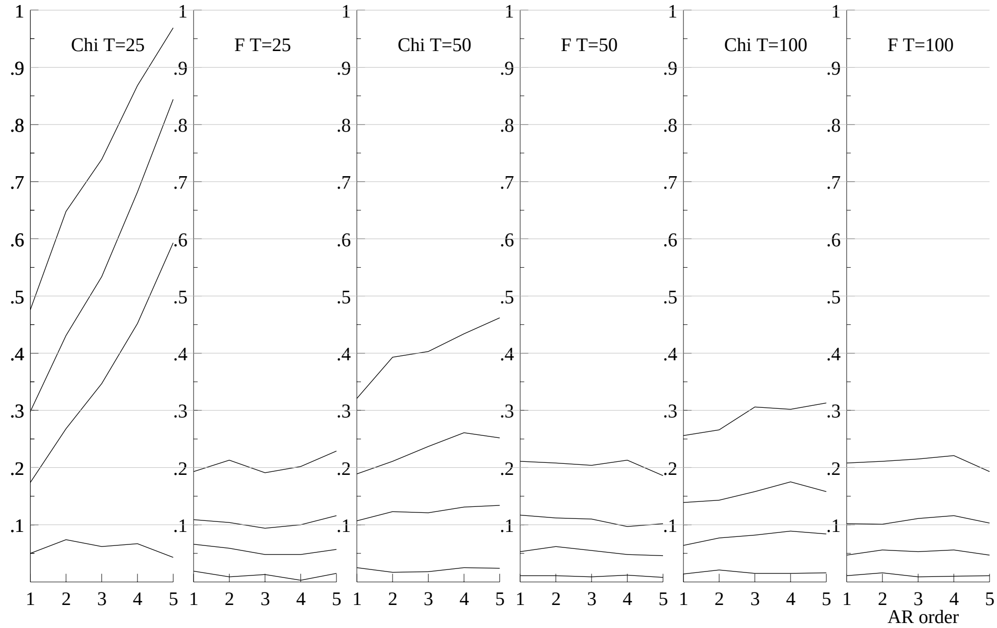

Figure 3. Rejection frequencies for *LM*,*LMF*, case (e): VARX(1).

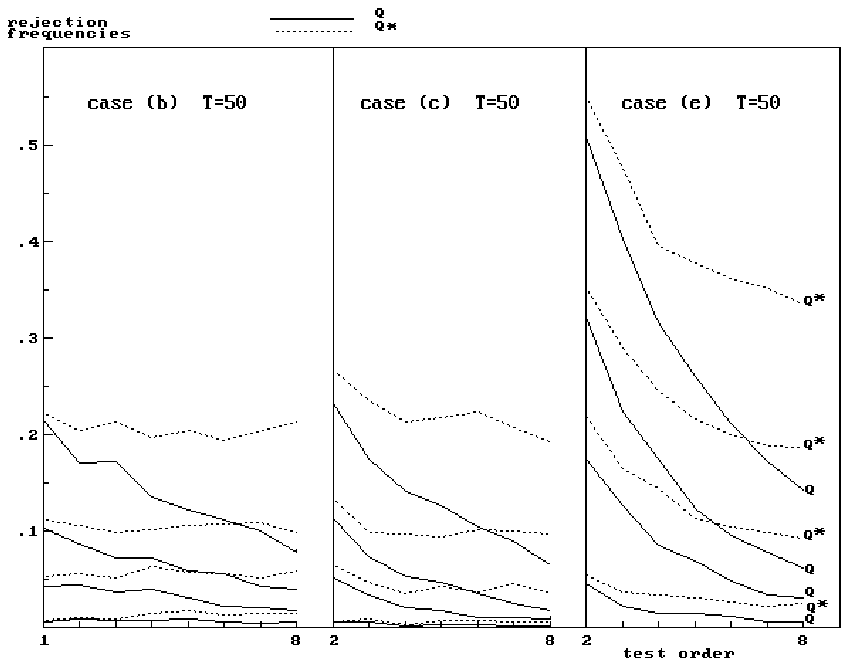

Figure 4. Rejection frequencies for Q,Q^*: cases (b),(c),(e).

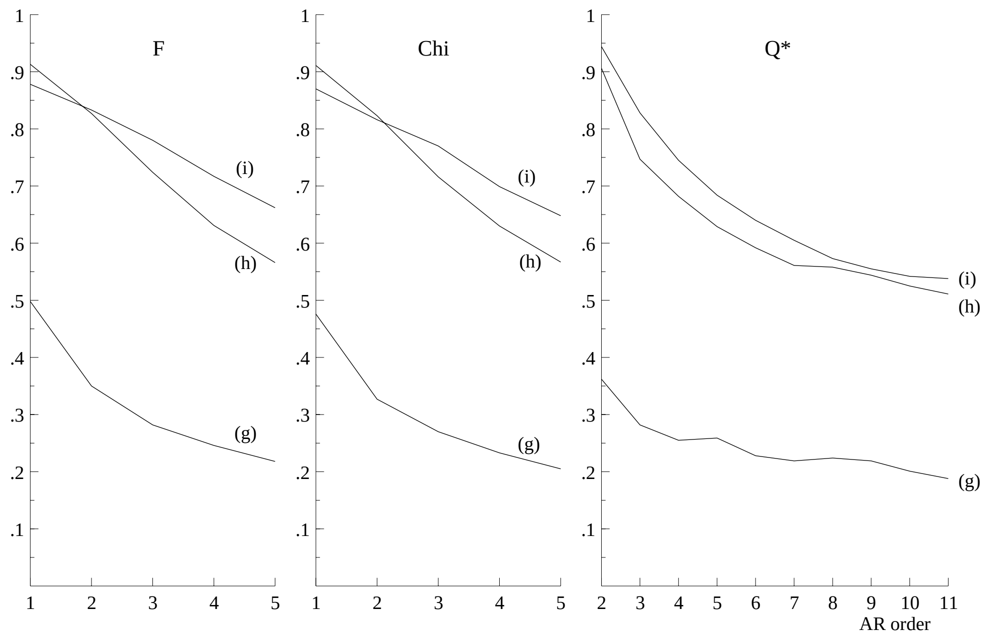

Figure 5. Rejection frequencies (power), cases (g)–(i), $T = 100$, $p = 0.05$.

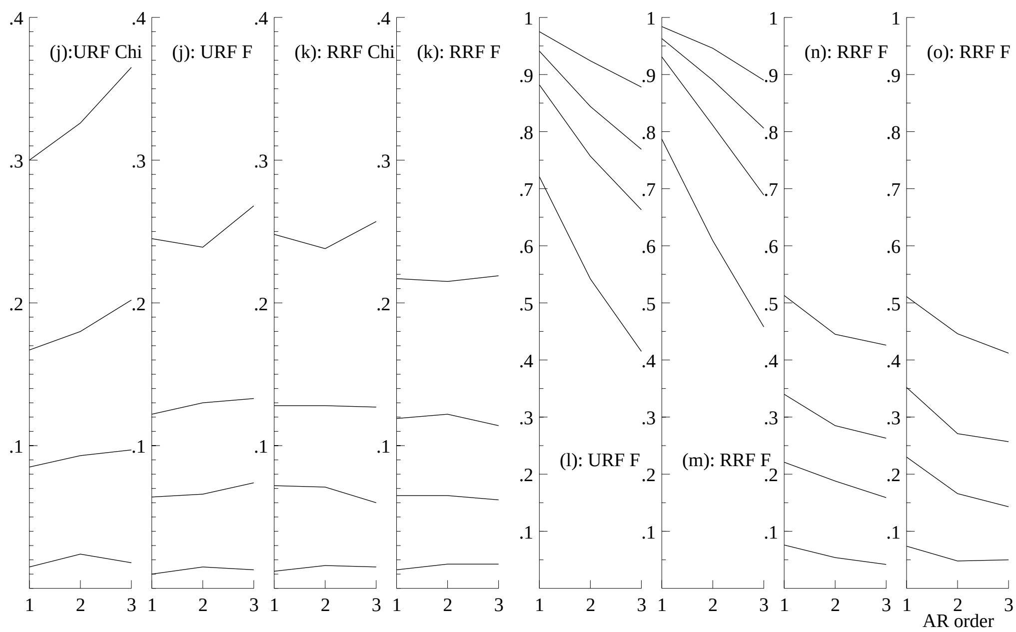

Figure 6. Rejection frequencies for *LM*,*LMF*, cases (j),(k), $T = 100$.

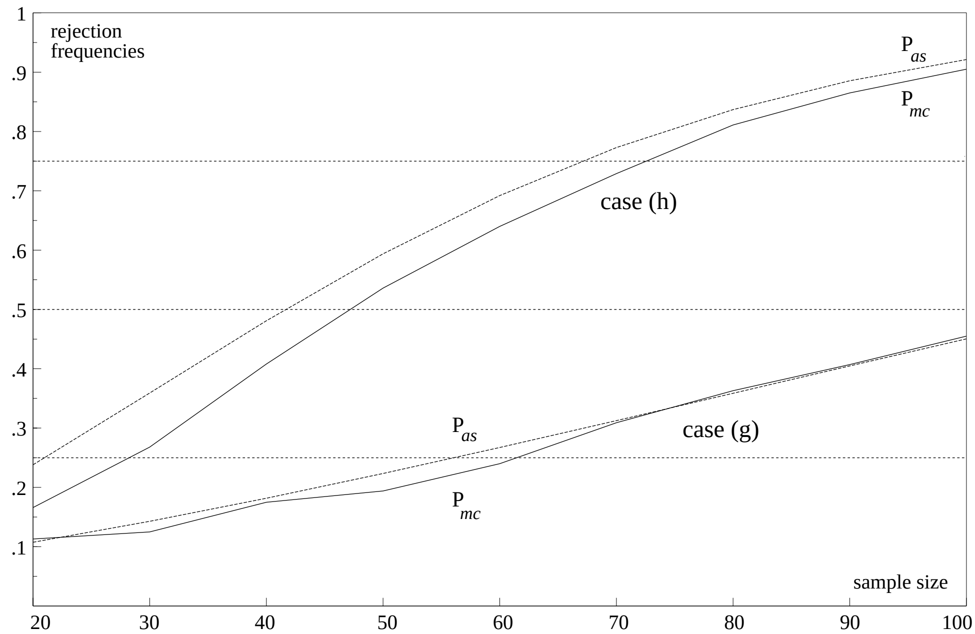

Figure 7. Asymptotic power (P_{as}) and Monte Carlo power (P_{mc}) for *LM* test: cases (g), (h).

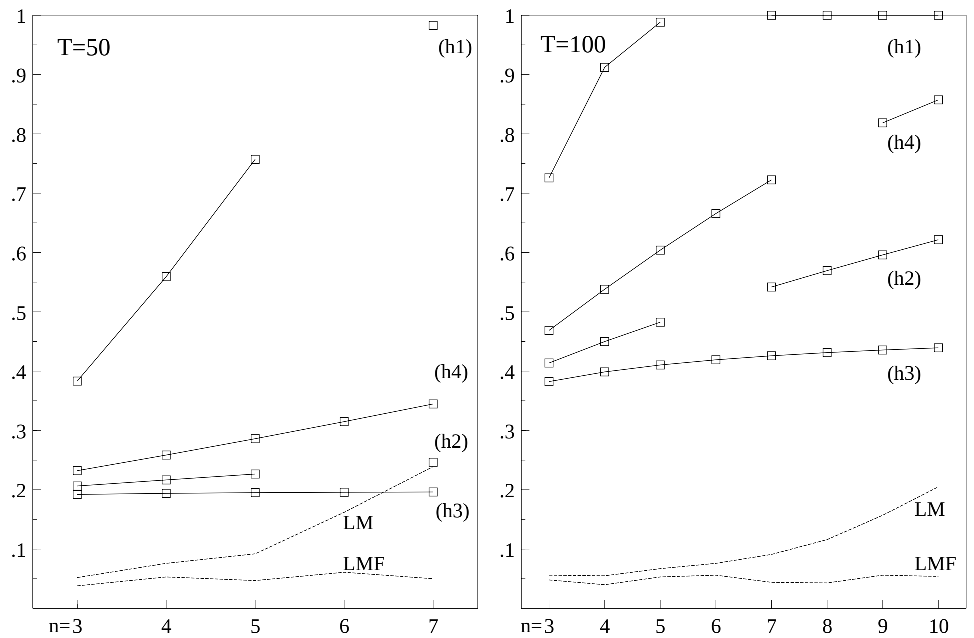

Figure 8. Asymptotic power of cases (h1) – (h4) and Monte Carlo size (dashed lines) for *LM* and *LMF* tests.

Table 4. Empirical size of heteroscedasticity tests, cases (b) and (e).

		case (b)				case (e)			
T	test	20%	10%	5%	1%	20%	10%	5%	1%
50	HET	0.181	0.100	0.048	0.014	0.161	0.072	0.037	0.004
50	$HETX$	0.182	0.091	0.058	0.009	0.136	0.049	0.014	0.001
50	HET-F	0.187	0.119	0.074	0.030	0.172	0.108	0.070	0.028
50	$HETX$-F	0.201	0.120	0.079	0.036	0.221	0.138	0.093	0.029
50	$ZHET$	0.175	0.092	0.050	0.013	0.171	0.084	0.041	0.006
50	$ZHETX$	0.174	0.085	0.042	0.009	0.130	0.047	0.015	0.002
50	$ZHET$-F	0.176	0.105	0.064	0.025	0.178	0.103	0.066	0.025
50	$ZHETX$-F	0.178	0.102	0.067	0.020	0.201	0.113	0.066	0.023
50	GIV-F	0.119	0.064	0.039	0.014	0.035	0.021	0.007	0.003
50	$GIVX$-F	0.103	0.061	0.039	0.010	0.003	0.001	0.001	0.000
50	E_p	0.187	0.089	0.045	0.010	0.181	0.087	0.046	0.013
100	HET	0.180	0.092	0.051	0.015	0.165	0.088	0.047	0.015
100	$HETX$	0.183	0.100	0.056	0.021	0.162	0.086	0.053	0.017
100	HET-F	0.182	0.100	0.056	0.023	0.163	0.101	0.063	0.021
100	$HETX$-F	0.193	0.114	0.067	0.027	0.191	0.121	0.085	0.035
100	$ZHET$	0.182	0.090	0.046	0.013	0.170	0.085	0.050	0.012
100	$ZHETX$	0.166	0.097	0.051	0.017	0.157	0.080	0.043	0.012
100	$ZHET$-F	0.178	0.096	0.053	0.015	0.177	0.092	0.057	0.021
100	$ZHETX$-F	0.168	0.105	0.062	0.023	0.178	0.106	0.069	0.024
100	GIV-F	0.147	0.075	0.041	0.015	0.093	0.049	0.028	0.011
100	$GIVX$-F	0.138	0.082	0.048	0.021	0.083	0.048	0.026	0.011
100	E_p	0.157	0.079	0.045	0.013	0.187	0.100	0.057	0.015

Table 5. Empirical power of heteroscedasticity tests, cases (h),(e1)–(e3), $T = 100$.

	case (h)				case (e3)			
test	20%	10%	5%	1%	20%	10%	5%	1%
HET	0.226	0.100	0.063	0.018	0.977	0.951	0.909	0.795
$HETX$	0.259	0.139	0.079	0.022	0.976	0.956	0.936	0.864
$HET\text{-}F$	0.222	0.111	0.068	0.023	0.978	0.960	0.935	0.866
$HETX\text{-}F$	0.256	0.162	0.095	0.033	0.983	0.978	0.961	0.938
$ZHET$	0.203	0.100	0.047	0.014	0.965	0.931	0.894	0.779
$ZHETX$	0.213	0.113	0.065	0.020	0.967	0.932	0.891	0.762
$ZHET\text{-}F$	0.203	0.105	0.054	0.018	0.971	0.944	0.918	0.832
$ZHETX\text{-}F$	0.215	0.120	0.074	0.024	0.973	0.956	0.939	0.883
E_p	0.166	0.083	0.036	0.009	0.997	0.993	0.988	0.963
	case (e1)				case (e2)			
HET	0.342	0.229	0.158	0.064	0.489	0.360	0.250	0.121
$HETX$	0.403	0.298	0.208	0.093	0.589	0.480	0.375	0.199
$HET\text{-}F$	0.354	0.245	0.183	0.094	0.500	0.387	0.286	0.166
$HETX\text{-}F$	0.450	0.354	0.286	0.174	0.638	0.557	0.488	0.349
$ZHET$	0.329	0.211	0.136	0.059	0.454	0.333	0.246	0.132
$ZHETX$	0.390	0.267	0.175	0.071	0.580	0.445	0.338	0.189
$ZHET\text{-}F$	0.338	0.226	0.155	0.071	0.458	0.348	0.275	0.167
$ZHETX\text{-}F$	0.408	0.308	0.243	0.138	0.613	0.523	0.428	0.309
E_p	0.357	0.256	0.188	0.103	0.721	0.643	0.567	0.447