MONASH
UNIVERSITY
AUSTRALIA

Pre-Test Strategies for Time-Series Forecasting in the Linear Regression Model

Maxwell L. King

and

Mei Leng Rankin

Working Paper No. 17/93

November 1993

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 7326 0385 4

Pre-Test Strategies for Time-Series Forecasting in the Linear Regression Model

Maxwell L. King

and

Mei Leng Rankin

Working Paper No. 17/93

November 1993

DEPARTMENT OF ECONOMETRICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA

Pre-Test Strategies for Time-Series Forecasting in the Linear Regression Model

by

Maxwell L. King

Department of Econometrics
Monash University
Clayton, Victoria 3168
Australia

and

Mei Leng Rankin

Department of Banking and Finance
Monash University
Caulfield, Victoria 3162
Australia

Abstract

In linear time-series regression analysis, there is typically uncertainty about which variables to include as regressors and the exact form of the disturbance process. This paper uses the Monte Carlo method to investigate the predictive performance of nine different pretesting strategies for misspecification. Data generating processes used in the study include first-order autoregressive (AR(1)) disturbances, first-order moving average disturbances, an extra exogenous regressor and the lagged dependent variable as an extra regressor. We find that remarkably robust predictions for a range of misspecified models result from applying the Durbin-Watson test for autocorrelation and correcting for AR(1) disturbances when the test is significant.

1. Introduction

In economic time series applications of the linear regression model, there is typically uncertainty about which variables to include as regressors and the exact form of the disturbance process. It has long been recognised that misspecification of the regressors or the disturbances can lead to misleading inferences and inefficient predictions. Since the 1950's, a widely adopted strategy to guard against these problems has been to test for autocorrelation in the disturbances, typically using the Durbin-Watson (DW) test. Because this test was originally designed to detect first-order autoregressive (AR(1)) disturbances, some investigators immediately assume AR(1) disturbances upon finding a significant DW statistic. As pointed out by Griliches (1966) and Carr (1972), a significant DW test can also be caused by omitted regressors. One popular interpretation is that a significant DW test indicates a problem with the dynamics of the regression that might be overcome by the inclusion of a lagged dependent variable. Furthermore, King (1983b) has shown that the DW test is approximately locally best invariant against first-order moving average (MA(1)) disturbances. This together with some support from economic theory (e.g., see Rowley and Wilton (1973), Sims (1974) and Nicholls, Pagan and Terrell (1975)) has led to an increased recognition of MA(1) disturbances as an important alternative to AR(1) errors. There is also a growing literature concerned with testing between these two competing disturbance processes (e.g., see King (1983a, 1987b), Burke, Godfrey and Tremayne (1990) and Silvapulle and King (1991)).

Thus a minimal list of possible reasons for a significant DW test would include:

(i) a Type I error has occurred,

(ii) the regression has AR(1) disturbances,

(iii) the regression has MA(1) disturbances,

(iv) an important exogenous regressor has been omitted,

(v) the lagged dependent variable has been omitted as a regressor.

If forecasting accuracy is the primary goal, it is not clear how best to proceed. It is often assumed that AR(1) disturbances help model the effects of omitted regressors. Should one therefore just correct for AR(1) disturbances in the standard way in the hope that this will result in reasonably accurate predictions? On the other hand, a case could be made for the inclusion of a lagged dependent variable. Should a series of pairwise tests be conducted in an attempt to come up with the true model which is then used to construct forecasts? How is the forecasting accuracy of this approach affected by the inevitable Type I and Type II errors that would occur in such a series of preliminary tests?

Much has been written on the effects of pre-testing on estimation and hypothesis testing (e.g., see Wallace (1977), Judge and Bock (1978), Judge (1984) and Giles and Giles (1993)). Unfortunately little is known about its effects on out of sample forecasting accuracy. King and Giles (1984) investigated the effects of pre-testing for AR(1) disturbances on estimation, testing and forecasting in the context of the linear regression model with AR(1) errors. This setting typifies much of the pre-testing literature in which only two competing models and one pre-test are considered. We are interested in a situation in which there are five possible models and even more potential pre-tests.

The aim of this paper is to investigate the effects of a range of pre-testing strategies on the predictive performance of the general linear model under the five scenarios implied by (i) to (v) above using the Monte Carlo method. To keep the study manageable, we have restricted our attention to nine possible strategies which range from ignoring the DW test and always predicting using ordinary least squares (OLS) to conducting a series of pre-tests including tests for autocorrelation, omitted regressors and a non-nested test of AR(1) disturbances against MA(1) disturbances.

The plan of the paper is as follows. Section 2 outlines the various models, estimators, tests and predictors used in the study. Section 3 discusses the nine alternative pre-test strategies the study investigates. The design of the Monte Carlo experiment is presented in Section 4 and its results are discussed in Section 5. We find that the familiar procedure of testing for autocorrelation using the DW test and correcting for AR(1) disturbances when the test is significant, results in remarkably robust predictions for a range of misspecified models. This and other conclusions are presented in Section 6.

2. Models, Estimators, Predictors and Tests

This section outlines the various models, estimators, predictors and preliminary tests that were used in the Monte Carlo experiment reported below.

2.1 Models, Estimators and Predictors

The underlying model is

$$y = X\beta + u \qquad (1)$$

where y is an $n \times 1$ vector of observations on the dependent variable, X is an $n \times k$ matrix of fixed regressors with rank $k \le n$, β is a $k \times 1$ vector of unknown parameters and u is the disturbance vector such that

$$u \sim N(0, \sigma^2 I_n) . \qquad (2)$$

The aim is to predict y_{n+i}, $i = 1,\ldots,m$, given x_{n+i} which is a $k\times1$ vector of the $(n+i)^{th}$ observations on each of the k regressors. The best linear unbiased predictor (BLUP) is

$$\hat{y}_{n+i} = x'_{n+i}\hat{\beta}\ , \qquad i = 1,\ldots,m\ , \tag{3}$$

where $\hat{\beta} = (X'X)^{-1}X'y$ is the OLS estimator of β.

The first variation of (1) and (2) is to assume that the disturbances, u_t, are generated by the stationary AR(1) process

$$u_t = \rho u_{t-1} + e_t\ , \qquad t = 1,\ldots,n \tag{4}$$

where $e = (e_1,\ldots,e_n)' \sim N(0,\sigma^2 I_n)$ and $0 \leq \rho < 1$. (4) implies (2) is replaced by

$$u \sim N(0,\sigma^2\Omega(\rho)) \tag{5}$$

where $\Omega(\rho)$ is the $n\times n$ matrix

$$\Omega(\rho) = (1-\rho^2)^{-1}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & & & \rho^{n-2} \\ \rho^2 & \rho & 1 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & & 1 & \rho \\ \rho^{n-1} & \rho^{n-2} & \cdots & \cdots & \rho & 1 \end{bmatrix}.$$

If the value of ρ is known then the BLUP (see Goldberger (1962)) of y_{n+i} is

$$\tilde{y}_{n+i} = x'_{n+i}\tilde{\beta} + \rho^i\tilde{u}_n \tag{6}$$

where

$$\tilde{\beta} = (X'\Omega^{-1}(\rho)X)^{-1}X'\Omega^{-1}(\rho)y \tag{7}$$

is the generalized least squares (GLS) estimator of β and

$$\tilde{u}_n = y_n - x_n'\tilde{\beta} \tag{8}$$

is the GLS residual for the n^{th} observation. Unfortunately, the value of ρ is typically unknown. The standard approach is to estimate β, σ^2 and ρ jointly using Beach and MacKinnon's (1978) full maximum likelihood procedure. Let $\tilde{\beta}^*$, $\tilde{\sigma}^{*2}$ and $\tilde{\rho}$ denote the resultant estimators and note that $\tilde{\beta}^*$ is given by (7) with ρ replaced by $\tilde{\rho}$. Our predictor (6) now becomes

$$\tilde{y}^*_{n+i} = x'_{n+i}\tilde{\beta}^* + \tilde{\rho}^i\tilde{u}^*_n \tag{9}$$

where $\tilde{u}^*_n$ is the n^{th} residual defined by (8) with $\tilde{\beta}$ replaced by $\tilde{\beta}^*$.

The second variation of (1) and (2) is to assume that the elements of u are generated by the MA(1) process

$$u_t = e_t + \gamma e_{t-1} \tag{10}$$

where $e^+ = (e_0, e_1, \dots, e_n)' \sim N(0, \sigma^2 I_{n+1})$ and $0 \le \gamma \le 1$. In this case (2) is replaced by

$$u \sim N(0, \sigma^2\Sigma(\gamma)) \tag{11}$$

where $\Sigma(\gamma)$ is the n×n tridiagonal matrix

$$\Sigma(\gamma) = \begin{bmatrix} 1+\gamma^2 & \gamma & 0 & \cdots & & 0 \\ \gamma & 1+\gamma^2 & \gamma & & & 0 \\ 0 & \gamma & 1+\gamma^2 & \ddots & & \vdots \\ & & & \ddots & & \vdots \\ & & & & 1+\gamma^2 & \gamma \\ 0 & 0 & \cdots & & \gamma & 1+\gamma^2 \end{bmatrix} .$$

Again the standard approach is to estimate β, σ^2 and γ jointly using full maximum likelihood. We used Pesaran's (1973) procedure which involves first reducing $\Sigma(\gamma)$ to a diagonal matrix by multiplying (1) by an orthogonal transformation comprised of the eigenvectors of $\Sigma(\gamma)$. Let $\breve{\beta}$, $\breve{\sigma}^2$ and $\breve{\gamma}$ denote the resultant estimators of β, σ^2 and γ, respectively, and observe that $\breve{\beta}$ is given by (7) with $\Omega(\rho)$ replaced by $\Sigma(\breve{\gamma})$. The predictor analogous to (9) is

$$\begin{aligned} \breve{y}_{n+1} &= x'_{n+1}\breve{\beta} + \breve{\gamma}\, \breve{e}_n , \\ \breve{y}_{n+i} &= x'_{n+i}\breve{\beta} , \qquad i = 2,\dots,m , \end{aligned} \tag{12}$$

where $\breve{e}_n$ is the last estimated observation on e_t, $t = 1,\dots,n$ (see King and McAleer (1987)).

The third variation of (1) and (2) is to assume an extra regressor so that (1) is replaced by

$$y = X\beta + w\delta + u \tag{13}$$

where w is an $n\times1$ vector of observations on the extra exogenous regressor and δ is a scalar. The disturbances are assumed to follow (2) which implies that the BLUP of y_{n+i} is

$$\hat{y}^+_{n+i} = x'_{n+i}\hat{\beta}^+ + w_{n+i}\hat{\delta} \tag{14}$$

where $(\hat{\beta}^{+\prime},\hat{\delta})'$ denotes the OLS estimator of $(\beta',\delta)'$ from (13) and w_{n+i} is the $(n+i)^{th}$ observation on the extra exogenous regressor.

The fourth variation of (1) and (2) is to assume the dependent variable lagged one period is an extra regressor so that (1) is replaced by

$$y = y_{-1}\theta + X\beta + u \tag{15}$$

where $y_{-1} = (y_0, \ldots, y_{n-1})'$ and θ is an unknown scalar such that $|\theta| < 1$. The disturbances are assumed to follow (2) in which case the standard predictor of y_{n+i} is

$$\begin{aligned} \bar{y}_{n+1} &= y_n\bar{\theta} + x'_{n+1}\bar{\beta} , \\ \bar{y}_{n+i} &= \bar{y}_{n+i-1}\bar{\theta} + x'_{n+i}\bar{\beta}, \qquad i = 2,\ldots,m , \end{aligned} \tag{16}$$

where $(\bar{\theta}, \bar{\beta}')'$ is the OLS estimator of $(\theta, \beta')'$ from (15).

2.2 Tests

The DW test for positive autocorrelation rejects the null hypothesis of (2) when

$$d = \sum_{t=2}^{n}\left(\hat{u}_t - \hat{u}_{t-1}\right)^2 \Big/ \sum_{t=1}^{n}\hat{u}_t^2 < d_\alpha$$

where $\hat{u}_t$, $t = 1,\ldots,n$, are the OLS residuals from (1) and d_α is the critical value at the α level of significance. In the Monte Carlo study reported below, exact values of d_α were found as outlined by King (1987a, p27) using a modified version of Koerts and Abrahamse's (1969) FQUAD subroutine.

King's (1987b) point optimal invariant (POI) test of the null hypothesis that u_t is generated by the stationary AR(1) process (4) against the alternative hypothesis that u_t is generated by the MA(1) process (10), optimizes power at $\gamma = \gamma_1 > 0$ where γ_1 is a fixed value chosen by the investigator. The test is based on the critical region

$$r(\gamma_1,\rho_1) = \frac{\breve{u}'\Sigma^{-1}(\gamma_1)\breve{u}}{\tilde{u}'\Omega^{-1}(\rho_1)\tilde{u}} < c_\alpha \tag{17}$$

where $\breve{u}$ and $\tilde{u}$ are the GLS residual vectors from (1) assuming covariance matrices $\Sigma(\gamma_1)$ and $\Omega(\rho_1)$, respectively, and c_α is the critical value. For (17) to be the critical region of a POI test requires that

$$\Pr\left[r(\gamma_1,\rho_1) < c_\alpha \mid u \sim N(0,\Omega(\rho_1))\right] = \alpha$$

and

$$\Pr\left[r(\gamma_1,\rho_1) < c_\alpha \mid u \sim N(0,\Omega(\rho)), \quad 0 \le \rho < 1\right] \le \alpha$$

be solved jointly for ρ_1 and c_α where α is the desired level of significance. An algorithm for achieving this is outlined by King (1987b) who also presents tables of solved values of c_α and ρ_1 for X matrices which are considered representative in one sense and for the choices $\gamma_1 = 0.5$ and 0.75. He suggested such values can be used to apply an approximate test. This approximate POI test with $\gamma_1 = 0.5$ was used in the study reported below.

The standard OLS-based t-test was used to test $H_0 : \delta = 0$ against $H_a : \delta \neq 0$ in (13) and also to test $H_0 : \theta = 0$ against $H_a : \theta > 0$ in (15). In both cases, critical values from the Student's t distribution with n-k-1 degrees of freedom were used for sample sizes of 15 and 30 while standard normal critical values were used for n = 60. In the case of testing $H_0 : \theta = 0$ in the dynamic regression (15), the justification for the use of Student's t critical values comes from Nankervis and Savin (1987). They showed that the small disturbance distributions of t and F statistics from (15) are the same as those for (15) with y_{-1} replaced by its expected value.

3. The Alternative Pre-test Strategies

The various pre-test strategies investigated in the study reported below are outlined in this section. In order to keep the study manageable, we restricted our attention to nine different strategies. Clearly there are many more that we could have included. The nine strategies, which all begin with the application of the DW test to (1), are:

S1 : Ignore the DW test and always predict using (3) which assumes (1) and (2) is the true model.

S2 : If the DW test is insignificant, predict using (3), otherwise correct for AR(1) disturbances and predict using (9).

S3 : If the DW test is insignificant, predict using (3), otherwise conduct King's approximate POI test of AR(1) disturbances against MA(1) disturbances. If this test finds in favour of AR(1) disturbances, correct for AR(1) disturbances and predict using (9), otherwise correct for MA(1) disturbances and predict using (12).

S4 : If the DW test is insignificant, predict using (3), otherwise test for an omitted exogenous regressor as in (13). If the t-test of δ is insignificant, predict using (3), otherwise include the extra regressor and predict using (14).

S5 : If the DW test is insignificant, predict using (3), otherwise test for an omitted exogenous regressor as in (13). If the t-test of δ is insignificant, correct for AR(1) disturbances and predict using (9), otherwise include the extra regressor and predict using (14).

S6 : If the DW test is insignificant, predict using (3), otherwise test for an omitted exogenous regressor as in (13). If the t-test of δ is significant, include the extra regressor and predict using (14), otherwise conduct King's approximate POI test of AR(1) disturbances against MA(1) disturbances. If this test finds in favour of AR(1) disturbances, correct for AR(1) disturbances and predict using (9), otherwise correct for MA(1) disturbances and predict using (12).

S7 : If the DW test is insignificant, predict using (3), otherwise test for an omitted lagged dependent variable as in (15). If the t-test of θ is insignificant, predict using (3), otherwise include the lagged dependent variable and predict using (16).

S8 : If the DW test is insignificant, predict using (3), otherwise test for an omitted lagged dependent variable as in (15). If the t-test of θ is insignificant, correct for AR(1) disturbances and predict using (9), otherwise include the lagged dependent variable and predict using (16).

S9 : If the DW test is insignificant, predict using (3), otherwise test for an omitted lagged dependent variable as in (15). If the t-test of θ is significant, include the lagged dependent variable and predict using (16), otherwise conduct King's approximate POI test of AR(1) disturbances against MA(1) disturbances. If this test finds in favour of AR(1) disturbances, correct for AR(1) disturbances and predict using (9), otherwise correct for MA(1) disturbances and predict using (12).

In summary, strategies S1, S2 and S3 ignore the possibility of omitted regressors and successively respond to a significant DW test by ignoring it (S1), assuming AR(1) disturbances is the only possibility (S2) and assuming AR(1) and MA(1) disturbances are the only possibilities (S3). Strategies S4, S5 and S6 allow for the possibility of an omitted exogenous regressor whose identity is known. If the regressor is found to be significant, the new model is used for prediction. Otherwise we fall back to strategies S1, S2 and S3 to deal with the significant DW statistic. Strategies S7, S8 and S9 are just as S4, S5 and S6, respectively, but with the lagged dependent variable being the likely omitted regressor. In all cases we assume positive autocorrelation.

4. Design of the Monte Carlo Experiment

In order to study the effects of the various pre-test strategies on prediction accuracy, a Monte Carlo experiment was conducted using 1000 replications. Let $a_1,\dots,a_n$ denote the eigenvectors corresponding to the eigenvalues of the n×n DW A_1 matrix arranged in ascending order, i.e.

$$a_1 = n^{-1/2}(1,1,\ldots,1)',$$

$$a_i = (2/n)^{1/2}(\cos[\pi(i-1)/2n],\ \cos[3\pi(i-1)/2n],\ \ldots,\ \cos[(2n-1)\pi(i-1)/2n])', \quad i = 2,\ldots,n .$$

The following design matrices with sample sizes of n = 15, 30 and 60 were used:

X1 : The regressors are $a_1, a_2, \ldots, a_k$ where k = 3,5. The extra exogenous regressor in (13) is a_{k+1}.

X2 : The regressors are $a_1, (a_2+a_n)/\sqrt{2}, \ldots, (a_k+a_{n-k+2})/\sqrt{2}$ where k = 3,5. The extra exogenous regressor in (13) is $(a_{k+1}+a_{n-k+1})/\sqrt{2}$.

X3 : The regressors are a constant dummy and log of annual real income per capita from Durbin and Watson's (1951) consumption of spirits example (k=2). The extra exogenous regressor in (13) is the log of the relative price of spirits.

X4 : The regressors are a constant dummy, the Australian quarterly consumer price index (CPI) and quarterly seasonally adjusted Australian household disposable income commencing 1959(1) (k=3). The extra exogenous regressor in (13) is the income series lagged one quarter.

X5 : The regressors are a constant dummy, three quarterly seasonal dummies, quarterly Australian gross national expenditure (GNE) commencing 1966(3) (k=5). The extra exogenous regressor in (13) is the two-year Australian Government bond rate.

X1 and X2 are artificially generated data sets that represent two extremes. For X1, OLS is approximately the best linear unbiased estimator in the presence of either AR(1) or MA(1) disturbances and the DW test is approximately uniformly most powerful invariant. Watson (1955) found that for X2, OLS is most inefficient when the regression disturbances are (approximately) AR(1). The DW test is also known to perform poorly against

AR(1) disturbances when ρ is near one for X2 (see King (1985)). X3 is comprised of relatively smooth annual data while X4 and X5 are constructed from quarterly data with varying degrees of seasonality and collinearity.

Where required, independent pseudo-random normal disturbances were generated by applying the Box-Muller transformation to [0,1] uniform random variables produced by the RAN intrinsic on a VAX11-780 computer. The quality of the resultant string of N(0,1) pseudo-random variates has been examined by Giles and Beattie (1984).

After each set of disturbances was calculated, for each design matrix, the y_t's were generated by each of the following models:

M1 : The linear regression, (1), with AR(1) errors, (4), for ρ = 0.0, 0.3, 0.5, 0.9.

M2 : The linear regression, (1), with MA(1) errors, (10), for γ = 0.1, 0.3, 0.5, 0.9.

M3 : The linear regression with an extra exogenous regressor, (13), with independent errors, (2), for δ = 0.2, 0.5, 0.8, 1.0.

M4 : The linear regression with a lagged dependent variable as an additional regressor, (15), with independent errors, (2), for θ = 0.2, 0.4, 0.6, 0.8.

Observe that model M1 with $\rho = 0$ is the well behaved regression (1) and (2).

We conducted some preliminary experiments which showed that prediction errors were reasonably insensitive to the values taken by β. In fact because of invariance of the component tests and estimators, strategies S1 - S6 are invariant to the value taken by β for models M1 and M2 and are nearly invariant for S7 - S9. For models M3 and M4, large β values tend to push the power of the component t-tests towards one. In order to obtain t-test powers that are neither too high or too low, the coefficient of the CPI series in X4 was set to 10 and the coefficient of the GNE series in X5 was set to 1.5.

For each set of generated y_t values, nine sets of one-, two- and four-step ahead predictions were produced using each of the nine strategies in turn. Eight post estimation sample observations were used to assess the accuracy of these predictions. For each set of predictions, we calculated the mean prediction error (MPE), mean absolute prediction error (MAPE), and the root mean squared prediction error (RMSPE).

5. The Results

In order to analyse the results, we ranked each of the nine strategies for each different data generating process on the basis of the RMSPEs. Aggregates of the number of times different strategies were ranked first to third for M1 - M4 are presented in Tables 1 - 4, respectively. The rankings based on MAPEs were almost identical to those based on RMSPEs so we concentrated on RMSPEs and MPEs in the following analysis.

5.1 AR(1) Disturbances

Under M1 (AR(1) disturbances), the MPEs show the forecasts are typically biased upwards and this bias increases as ρ increases, m increases and n decreases, *ceteris paribus*. Also generally RMSPEs increase with ρ and m but decrease with n, *ceteris paribus*. The greatest differences between the best and worst strategy are observed when m = 1 and these differences become less evident as m increases.

As expected when $\rho = 0.0$, S1 is the optimal choice of strategy although when n = 15, the RMSPEs of all nine strategies are very similar. For non-zero values of ρ, S2 is frequently the optimal strategy in terms of RMSPE, particularly for m = 1. The next best overall strategy appears to be S3. Not surprisingly, for moderate and large values of ρ, the strategies that only use OLS (S1, S4 and S7) produce the worst predictions.

5.2 MA(1) Disturbances

The MPEs suggest a general upward bias in the predictions under M2 (MA(1) disturbances) with no systematic pattern as γ and n vary. The RMSPEs of all strategies under M2 increase as γ increases and as n declines, *ceteris paribus*. For $\gamma \leq 0.3$, there is a tendency for the RMSPEs of all nine strategies to merge as m increases, especially for large n. When m and γ are large, the results tend to cluster into 3 distinct groups, with S1, S2 and S3 having the smallest RMSPEs and S7, S8 and S9 typically having the largest RMSPEs.

In general, S1 is the best strategy when γ is small. S2 is best overall for large γ values when m = 1 which is somewhat surprising given that S3 takes account of the possibility of MA(1) disturbances. For larger m values, S1 continues to perform well although S3 is optimal for some design matrices. When m = 4, all strategies tend to merge. Overall S1, S2 and S3 are the best strategies and the small differences observed between their RMSPEs suggests that in the presence of MA(1) disturbances, any one of these three strategies is suitable.

5.3 An Extra Exogenous Regressor

In the case of M3 (an extra exogenous regressor), calculated MPEs indicate that the direction of bias in the predictions depends on the design matrix. It is generally a downward bias for X1, X2 and X3 and an upward bias for X4 and X5. Also the bias increases as δ, the coefficient of the extra exogenous regressor, increases and n increases, *ceteris paribus*. The bias of S2 and S3; S5 and S6; and S8 and S9 tend to be identical. When δ =

1.0, the bias of S1 is significantly larger than those of the other strategies.

The RMSPEs of all nine strategies typically first decline slightly as δ increases from 0.2 to 0.5 and then increase as δ increases beyond 0.5, *ceteris paribus*. As expected, the three appropriate strategies, S4, S5 and S6, frequently rank in the top three strategies in terms of RMSPEs, particularly when δ is large. The differences between these strategies are negligible when $\delta = 1.0$. For small to moderate δ values, S2 is often the best strategy followed by S3. This is particularly true for $\delta = 0.2$. In almost all cases, the performances of S2 and S3 are relatively close to those of S4, S5 and S6. Again we see that even though other strategies should be favoured on theoretical grounds, S2 is a competitive strategy.

5.4 The Lagged Dependent Variable as an Extra Regressor

The MPEs under M4 show a general upward bias in the predictions from S1 - S6 and these biases increase with θ. In contrast, the predictions from S7 - S9 show a general downward bias, especially when n is large. Groupings of strategies with very similar biases are S2 and S3; S5 and S6; and S7, S8 and S9. The RMSPEs of all strategies generally increase with θ, *ceteris paribus*.

One would expect the results to favour a strategy from S7, S8 and S9. However, the best overall strategy appears to be S2 for nearly all θ values. As m increases, the frequency with which S2 is best declines. In most cases, especially for large θ values, S8, S9 and S7 rank amongst the top four on the basis of RMSPE. The dominance of S2 is specific to design matrices, being favoured by X1, X2 and X3. For X4 and X5, S8 is the best strategy and for these design matrices as θ and n increase, the discrepancy in RMSPE between S8 and S2 gets significantly larger. A distinguishing feature of X4 and X5 is that they contain seasonal regressors. Although S2 appears to be the best strategy overall, some design matrices exist for which, when θ is large, a substantial loss in accuracy of prediction can occur if one proceeds to correct for AR(1) disturbances without first checking whether to include a lagged dependent variable.

6. Concluding Remarks

Naturally, care must be taken when generalizing from a limited Monte Carlo experiment such as the one reported above. However, in this case, all the results point in a similar direction. We find that strategy S2, which simply involves correcting for AR(1) disturbances if the DW statistic is significant, frequently provides the best predictions. When they are not the best predictions, their RMSPEs are almost always not far from the RMSPE of the best predictions. The only serious exceptions occur when the coefficient of the additional term, whether it be MA(1) errors, an extra exogenous regressor or the lagged dependent variable as an extra regressor, is large. What is somewhat surprising is the fact that the AR(1) disturbance model provides better forecasts than those that use knowledge of the true model when this coefficient is small. We therefore conclude that the familiar procedure of testing for autocorrelation using the DW test in a time-series regression and correcting for AR(1) disturbances when the test is significant, provides remarkably robust predictions for a range of misspecified models. If the misspecification is extreme then obviously better predictions will result from using a strategy which includes the correct model.

References

Beach, C.M. and MacKinnon, J.G., 'A maximum likelihood procedure for regression with autocorrelated errors', *Econometrica*, 46 (1978), 51-8.

Burke, S.P., Godfrey, L.G. and Tremayne, A.R., 'Testing AR(1) against MA(1) disturbances in the linear regression model: An alternative approach', *Review of Economic Studies*, 57 (1990), 135-45.

Carr, J., 'A suggestion for the treatment of serial correlation: A case in point', *Canadian Journal of Economics*, 5 (1972), 301-6.

Durbin, J. and Watson, G.S., 'Testing for serial correlation in least squares regression II', *Biometrika*, 38 (1951), 159-78.

Giles, D.E.A. and Beattie, M., 'Autocorrelation pre-test estimation in models with a lagged dependent variable', in King, M.L. and Giles, D.E.A. (eds) *Specification Analysis in the Linear Model*, London: Routledge and Kegan Paul, 1987, 99-116.

Giles, J.A. and Giles, D.E.A., 'Pre-test estimation and testing in econometrics: Recent developments', *Journal of Economic Surveys*, forthcoming (1993).

Goldberger, A.S., 'Best linear unbiased prediction in the generalised linear regression model', *Journal of the American Statistical Association*, 57 (1962), 369-75.

Griliches, Z., 'Specification analysis in econometrics', mimeo (1966).

Judge, G.G. (ed), 'Pre-test and Stein-rule estimators: Some new results', Special issue *Journal of Econometrics*, 25 (1984), Amsterdam: North Holland, 1-239.

Judge, G.G. and Bock, M.E., *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*, Amsterdam: North-Holland, 1978.

King, M.L., 'Testing for autoregressive against moving average errors in the linear regression model', *Journal of Econometrics*, 21 (1983a), 35-51.

King, M.L., 'Testing for moving average regression disturbances', *Australian Journal of Stataistics*, 25 (1983b), 23-34.

King, M.L., 'A point optimal test for autoregressive disturbances', *Journal of Econometrics*, 24 (1985), 269-77.

King, M.L., 'Testing for autocorrelation in linear regression models: A survey', in King, M.L. and Giles, D.E.A. (eds) *Specification Analysis in the Linear Model*, London: Routledge and Kegan Paul, 1987a, 19-73.

King, M.L., 'Towards a theory of point optimal testing', *Econometric Reviews*, 6 (1987b), 169-218.

King, M.L. and Giles, D.E.A., 'Autocorrelation and pre-testing in the linear model: Estimation, testing and prediction', *Journal of Econometrics*, 25 (1984), 35-48.

King, M.L. and McAleer, M., 'Further results on testing AR(1) against MA(1) disturbances in the linear regression model', *Review of Economic Studies*, 54 (1987), 649-63.

Koerts, J. and Abrahamse, A.P.J., *On the Theory and Application of the General Linear Model*, Rotterdam: Rotterdam University Press, 1969.

Nankervis, J.C. and Savin, N.E., 'Finite sample distributions of t and F statistics in an AR(1) model with an exogenous variable', *Econometric Theory*, 3 (1987), 387-408.

Nicholls, D.F., Pagan, A.R. and Terrell, R.D., 'The estimation and use of models with moving average disturbance terms: A survey', *International Economic Review*, 16 (1975), 113-34.

Pesaran, M.H., 'Exact maximum likelihood estimation of a regression equation with a first order moving average error', *Review of Economic Studies*, 40 (1973), 529-35.

Rowley, J.C.R. and Wilton, P.A., 'Quarterly models of wage determination: Some new efficient estimates', *American Economic Review*, 63 (1973), 380-9.

Silvapulle, P. and King, M.L., 'Testing moving average against autoregressive disturbances in the linear regression model', *Journal of Business and Economic Statistics*, 9 (1991), 329-35.

Sims, C.A., 'Distributed lags', in Intriligator, M.D. and Kendrick, D.A. (eds) *Frontiers of Quantitative Economics 2*, Amsterdam: North-Holland, 1974, 289-338.

Wallace, T.D., 'Pretest estimation in regression: A survey', *American Journal of Agricultural Economics*, 59 (1977), 431-43.

Watson, G.S., 'Serial correlation in regression analysis I', *Biometrika*, 42 (1955), 327-41.

Table 1: Number of times each strategy is ranked 1st, 2nd and 3rd when the regression disturbances follow an AR(1) process (M1).

		m = 1				m = 2				m = 4			
	ρ =	0.0	0.3	0.5	0.9	0.0	0.3	0.5	0.9	0.0	0.3	0.5	0.9
Strategy	Rank												
S1	1st	15	5	1	0	15	9	2	1	18	13	3	1
	2nd	5	0	1	1	3	3	0	0	0	5	2	0
	3rd	1	2	0	0	0	7	2	0	1	1	6	0
S2	1st	3	7	17	17	7	11	16	17	12	10	13	17
	2nd	3	4	2	1	5	4	4	3	1	6	4	1
	3rd	5	7	0	1	4	2	1	0	4	3	3	2
S3	1st	3	8	5	11	8	8	7	11	10	9	13	14
	2nd	2	5	11	7	5	6	12	8	1	3	2	4
	3rd	5	4	3	0	3	2	1	0	4	3	5	2
S4	1st	7	1	0	0	6	3	0	1	8	4	0	0
	2nd	8	2	0	1	1	1	0	0	1	1	1	0
	3rd	1	1	1	1	1	1	1	0	1	0	0	0
S5	1st	3	2	1	1	4	0	0	2	6	4	3	2
	2nd	1	0	3	0	1	2	1	0	0	1	1	0
	3rd	1	5	5	3	2	3	9	9	0	4	5	7
S6	1st	2	2	1	0	4	1	0	1	4	2	1	2
	2nd	2	1	0	1	1	0	2	1	0	1	1	1
	3rd	0	4	3	2	2	2	7	7	2	3	6	7
S7	1st	5	1	1	1	7	2	1	1	11	0	1	0
	2nd	6	0	0	0	5	0	0	0	1	0	0	0
	3rd	2	1	2	5	2	1	2	4	2	1	0	5
S8	1st	2	1	0	0	1	0	0	0	8	1	1	2
	2nd	1	2	1	1	0	1	3	0	0	0	0	0
	3rd	3	3	7	11	4	1	5	8	2	2	1	8
S9	1st	1	1	0	0	6	0	0	0	7	2	1	2
	2nd	2	3	3	2	1	0	0	1	0	0	0	0
	3rd	4	3	4	9	2	2	8	7	2	1	1	6

Table 2: Number of times each strategy is ranked 1st, 2nd and 3rd when the regression disturbances follow an MA(1) process (M2).

		m = 1				m = 2				m = 4			
	γ =	0.1	0.3	0.5	0.9	0.1	0.3	0.5	0.9	0.1	0.3	0.5	0.9
Strategy	Rank												
S1	1st	15	6	2	0	17	17	17	12	18	15	12	7
	2nd	4	1	0	0	2	3	1	5	2	2	3	5
	3rd	0	2	3	1	1	1	2	1	0	2	2	0
S2	1st	1	9	16	17	4	2	3	0	9	10	8	2
	2nd	4	4	4	3	7	6	3	3	5	6	5	6
	3rd	3	5	1	1	3	5	6	5	3	1	6	10
S3	1st	2	4	2	5	4	2	7	6	10	11	10	12
	2nd	4	5	6	2	6	8	4	11	5	5	5	4
	3rd	5	5	4	3	4	5	5	0	1	1	2	4
S4	1st	4	0	0	0	5	4	3	2	6	4	3	1
	2nd	7	3	1	0	5	5	6	3	2	2	1	2
	3rd	3	0	0	0	1	3	1	6	0	0	0	0
S5	1st	1	3	2	0	2	3	1	0	4	2	2	2
	2nd	0	1	5	8	0	0	0	1	0	3	0	0
	3rd	1	2	4	5	0	1	1	1	1	1	5	0
S6	1st	1	2	1	0	1	0	3	2	3	3	2	1
	2nd	1	1	1	1	0	1	0	2	0	1	0	4
	3rd	3	3	1	2	1	0	1	4	1	2	2	3
S7	1st	4	1	0	0	3	1	0	0	8	1	0	0
	2nd	3	0	0	1	5	0	0	0	3	0	0	0
	3rd	3	1	2	4	2	4	2	2	0	2	0	0
S8	1st	1	1	1	1	0	0	0	0	7	1	1	0
	2nd	0	2	4	4	2	0	0	0	2	0	0	1
	3rd	3	7	3	7	0	0	1	0	0	1	1	0
S9	1st	1	0	0	0	0	0	0	0	7	1	1	2
	2nd	1	2	0	1	2	0	0	0	1	0	0	0
	3rd	3	8	3	2	0	1	1	0	0	1	2	0

Table 3: **Number of times each strategy is ranked 1st, 2nd and 3rd when the regression contains an extra exogenous regressor (M3).**

		m = 1				m = 2				m = 4			
	δ =	0.2	0.5	0.8	1.0	0.2	0.5	0.8	1.0	0.2	0.5	0.8	1.0
Strategy	Rank												
S1	1st	0	4	5	6	2	5	7	7	1	3	6	5
	2nd	0	1	3	1	0	0	1	1	0	1	0	2
	3rd	1	1	0	1	2	1	1	1	6	4	0	0
S2	1st	17	13	0	0	14	5	3	2	19	7	3	2
	2nd	1	2	2	1	3	4	3	2	0	3	2	2
	3rd	1	1	2	1	2	3	0	2	0	2	0	2
S3	1st	2	3	0	0	9	3	4	1	15	6	2	1
	2nd	7	5	1	1	6	2	1	2	4	4	1	2
	3rd	7	0	0	0	2	3	1	1	0	1	0	1
S4	1st	2	3	8	13	1	4	8	13	0	4	15	17
	2nd	0	1	2	2	0	1	0	1	0	0	1	0
	3rd	0	0	8	4	0	5	7	3	0	2	2	1
S5	1st	3	2	12	12	9	9	14	14	0	11	12	13
	2nd	3	7	0	1	1	1	1	2	0	0	4	2
	3rd	3	2	3	2	2	2	1	1	8	1	1	3
S6	1st	2	2	8	10	2	7	13	12	0	9	10	13
	2nd	0	3	3	1	0	3	1	2	0	1	4	1
	3rd	3	2	2	2	7	1	1	1	6	0	2	3
S7	1st	0	1	2	2	0	0	2	5	2	2	1	1
	2nd	1	0	1	3	1	0	1	1	0	0	0	1
	3rd	0	1	3	1	3	2	0	0	1	0	0	0
S8	1st	0	2	0	0	1	4	2	2	2	3	1	2
	2nd	5	4	1	0	0	0	0	1	0	0	0	0
	3rd	8	2	1	1	5	2	1	0	1	0	0	0
S9	1st	0	1	0	0	0	2	2	2	2	2	0	1
	2nd	3	1	1	0	1	1	0	1	0	0	0	0
	3rd	2	4	0	0	4	2	1	0	1	0	0	0

Table 4: Number of times each strategy is ranked 1st, 2nd and 3rd when the regression contains a lagged dependent variable as a regressor (M4).

		m = 1				m = 2				m = 4			
	θ =	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
Strategy	Rank												
S1	1st	4	0	0	1	9	2	0	2	9	9	6	2
	2nd	0	0	1	0	1	1	1	0	2	0	1	0
	3rd	2	0	0	0	5	3	2	1	4	4	1	0
S2	1st	10	11	12	10	10	5	11	12	7	7	7	12
	2nd	2	2	2	1	5	3	1	0	5	1	4	0
	3rd	4	0	1	0	3	2	1	1	3	3	0	0
S3	1st	5	2	3	4	8	9	4	4	9	8	7	3
	2nd	5	6	5	4	7	2	7	4	3	3	3	8
	3rd	3	4	0	0	0	2	1	2	4	2	1	1
S4	1st	0	1	1	3	2	2	0	2	4	2	4	2
	2nd	1	0	1	0	1	0	1	0	3	4	2	1
	3rd	0	0	0	0	1	1	1	0	0	1	0	0
S5	1st	2	0	1	2	2	2	1	2	3	0	1	2
	2nd	2	3	2	1	1	1	2	1	1	0	1	1
	3rd	3	2	1	0	2	2	2	2	2	1	3	3
S6	1st	2	0	0	2	1	4	1	2	3	0	1	2
	2nd	1	0	2	1	0	0	2	0	1	0	0	0
	3rd	1	2	0	0	2	1	1	2	2	0	3	3
S7	1st	0	4	3	8	2	5	6	5	2	4	5	7
	2nd	0	0	5	3	0	0	2	3	1	0	0	0
	3rd	4	3	5	6	3	5	4	5	3	2	7	6
S8	1st	5	9	7	11	2	8	7	7	1	4	6	8
	2nd	5	3	5	3	1	1	2	5	0	1	0	0
	3rd	2	6	6	4	2	5	4	4	2	3	5	6
S9	1st	3	8	7	9	3	6	7	9	1	5	6	8
	2nd	3	1	5	4	1	2	3	3	0	0	0	0
	3rd	4	4	6	4	1	4	4	4	3	3	6	7