Economics Bulletin

Volume 30, Issue 2

Testing an autoregressive structure in binary time series models

Henri Nyberg
University of Helsinki and HECER

Abstract

This paper introduces a Lagrange Multiplier (LM) test for testing an autoregressive structure in a binary time series model proposed by Kauppi and Saikkonen (2008). Simulation results indicate that the two versions of the proposed LM test have reasonable size and power properties when the sample size is large. A parametric bootstrap method is suggested to obtain approximately correct sizes also in small samples. The use of the test is illustrated by an application to recession forecasting models using monthly U.S. data.

The author would like to thank Markku Lanne and Pentti Saikkonen as well as participants at the Econometric Society Australasian Meeting in Canberra, July 2009, for constructive comments. The author is responsible for remaining errors. Financial support from the Academy of Finland, the Research Foundation of the OP Group and the Finnish Foundation for Advancement of Securities Markets is gratefully acknowledged.

Citation: Henri Nyberg, (2010) "Testing an autoregressive structure in binary time series models", *Economics Bulletin*, Vol. 30 no.2 pp. 1460-1473.

Submitted: Apr 14 2010. **Published:** May 18, 2010.

1. Introduction

Recently, Rydberg and Shephard (2003), Chauvet and Potter (2005) and Startz (2008), among others, have introduced new time series models for binary dependent variables. In this paper, the "dynamic autoregressive" probit model suggested by Kauppi and Saikkonen (2008) is considered. We develop Lagrange Multiplier (LM) test which can be used to test the adequacy of a restricted model in which the autoregressive structure is excluded. The proposed LM test is attractive because it only requires estimates from the restricted model, which can be obtained by using standard econometric software packages. According to our simulations, the two versions of the LM test considered have reasonable size and high power, especially in large samples. In small samples, a parametric bootstrap method is proposed to obtain critical values which are more reliable than the asymptotic ones. In an empirical application, the LM tests are used to assess recession forecasting models for the U.S.

The paper is organized as follows. The probit model is introduced in Section 2 and the LM tests are developed in Section 3. Results of the simulation and bootstrap experiments are provided in Section 4 and the empirical example is presented in Section 5. Finally, Section 6 concludes.

2. Model

Consider the binary valued stochastic process y_t, $t = 1, 2, ..., T$, and let $E_{t-1}(\cdot)$ and $P_{t-1}(\cdot)$, respectively, signify the conditional expectation and conditional probability given the information set Ω_{t-1}. Conditional on Ω_{t-1}, y_t has a Bernoulli distribution, that is,

$$y_t|\Omega_{t-1} \sim B(p_t). \tag{1}$$

In the probit model

$$p_t = E_{t-1}(y_t) = P_{t-1}(y_t = 1) = \Phi(\pi_t(\boldsymbol{\theta})), \tag{2}$$

where $\Phi(\cdot)$ is a standard normal cumulative distribution function and $\pi_t(\boldsymbol{\theta})$ is a linear function of variables in the information set Ω_{t-1} and the parameter vector $\boldsymbol{\theta}$.

The previous literature is mainly considered the "static" model

$$\pi_t(\boldsymbol{\theta}) = \omega + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{3}$$

where $\boldsymbol{x}_{t-1}^{'}$ is a vector of explanatory variables. An extension of this model (see, e.g., Cox 1981) is a dynamic model

$$\pi_t(\boldsymbol{\theta}) = \omega + \delta_1 y_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{4}$$

which also contains a lagged value of the dependent variable. As an extension of the dynamic model (4) Kauppi and Saikkonen (2008) propose a model with an autoregressive structure

$$\pi_t(\boldsymbol{\theta}) = \omega + \alpha_1 \pi_{t-1}(\boldsymbol{\theta}) + \delta_1 y_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{5}$$

where $|\alpha_1| < 1$. Note that alternative, but very similar, models have been proposed by Rydberg and Shephard (2003) and Kauppi (2008). The LM tests developed in the next section can straightforwardly be extended to these models as well.

The parameters of the models (3)–(5) can conveniently be estimated by the method of maximum likelihood (ML). Conditional on initial values, the log-likelihood function is

$$l(\boldsymbol{\theta}) = \sum_{t=1}^{T} l_t(\boldsymbol{\theta}) = \sum_{t=1}^{T} \Big(y_t \log(\Phi(\pi_t(\boldsymbol{\theta}))) + (1-y_t)\log(1-\Phi(\pi_t(\boldsymbol{\theta}))) \Big), \tag{6}$$

where $l_t(\boldsymbol{\theta})$ is the log-likelihood for t:th observation. The score function is

$$s(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{t=1}^{T} s_t(\boldsymbol{\theta}) = \sum_{t=1}^{T} \Big(\frac{y_t - \Phi(\pi_t(\boldsymbol{\theta}))}{\Phi(\pi_t(\boldsymbol{\theta}))(1-\Phi(\pi_t(\boldsymbol{\theta})))} \phi(\pi_t(\boldsymbol{\theta})) \frac{\partial \pi_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big), \tag{7}$$

where $\phi(\cdot)$ signifies the probability density function of the standard normal distribution and an explicit expression of the derivative term $\partial\pi_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ will be given in the next section.

3. LM Tests

In applications, model (5) may be a superior to its restricted version (4) but, on the other hand, its ML estimation is more complicated and no estimation procedures are readily available in standard econometric software packages. Thus, it is of interest to start with the simpler model (4) and check for its adequacy by testing whether the autoregressive coefficient α_1 in (5) is zero. The null hypothesis of interest is therefore

$$H_0 : \alpha_1 = 0. \tag{8}$$

In this context, the LM test is attractive because it only requires the estimation of the parameters of model (4). Following Davidson and MacKinnon (1984) we can construct two LM test statistics for the null hypothesis (8). The first one is

$$LM_1 = \boldsymbol{\iota}' S(\tilde{\boldsymbol{\theta}}) \Big(S(\tilde{\boldsymbol{\theta}})' S(\tilde{\boldsymbol{\theta}}) \Big)^{-1} S(\tilde{\boldsymbol{\theta}})' \boldsymbol{\iota}, \tag{9}$$

where $\boldsymbol{\iota}$ is a vector of ones and the matrix $S(\tilde{\boldsymbol{\theta}})$ is given by

$$S(\tilde{\boldsymbol{\theta}}) = \Big(s_1(\tilde{\boldsymbol{\theta}}) \quad s_2(\tilde{\boldsymbol{\theta}}) \ldots s_T(\tilde{\boldsymbol{\theta}}) \Big)'.$$

Expression (9) can also be seen as the regression sum of squares from the artificial linear regression

$$\boldsymbol{\iota} = S(\tilde{\boldsymbol{\theta}}) a + error.$$

Using the symbols $\tilde{\Phi}_t = \Phi(\pi_t(\tilde{\boldsymbol{\theta}}))$ and $\tilde{\phi}_t = \phi(\pi_t(\tilde{\boldsymbol{\theta}}))$, a second LM test statistic can be based on the artificial regression

$$r(\tilde{\boldsymbol{\theta}}) = R(\tilde{\boldsymbol{\theta}}) b + error, \tag{10}$$

where

$$R(\tilde{\boldsymbol{\theta}}) = \Big(R_1(\tilde{\boldsymbol{\theta}})' \quad R_2(\tilde{\boldsymbol{\theta}})' \ldots R_T(\tilde{\boldsymbol{\theta}})' \Big)'$$

with

$$R_t(\tilde{\boldsymbol{\theta}}) = \Big(\tilde{\Phi}_t (1 - \tilde{\Phi}_t) \Big)^{-1/2} \tilde{\phi}_t \frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}$$

and

$$r(\tilde{\boldsymbol{\theta}}) = \Big(r_1(\tilde{\boldsymbol{\theta}}) \quad r_2(\tilde{\boldsymbol{\theta}}) \ldots r_T(\tilde{\boldsymbol{\theta}}) \Big)'$$

with

$$\begin{aligned} r_t(\tilde{\boldsymbol{\theta}}) &= y_t \Big(\frac{1 - \tilde{\Phi}_t}{\tilde{\Phi}_t} \Big)^{1/2} + (y_t - 1) \Big(\frac{\tilde{\Phi}_t}{1 - \tilde{\Phi}_t} \Big)^{1/2} \\ &= \Big((1 - \tilde{\Phi}_t) \tilde{\Phi}_t \Big)^{-1/2} \Big(y_t - \tilde{\Phi}_t \Big). \end{aligned}$$

Running the artificial regression (10) and computing the regression sum of squares yields the test statistic

$$LM_2 = r(\tilde{\boldsymbol{\theta}})' R(\tilde{\boldsymbol{\theta}}) \Big(R(\tilde{\boldsymbol{\theta}})' R(\tilde{\boldsymbol{\theta}}) \Big)^{-1} R(\tilde{\boldsymbol{\theta}})' r(\tilde{\boldsymbol{\theta}}). \tag{11}$$

Because $R(\tilde{\boldsymbol{\theta}})' r(\tilde{\boldsymbol{\theta}}) = s(\tilde{\boldsymbol{\theta}}) = S(\tilde{\boldsymbol{\theta}})'\boldsymbol{\iota}$, it can be seen that the test statistics LM_1 and LM_2 only differ in the way the information matrix estimate $\mathcal{I}(\tilde{\boldsymbol{\theta}})$ is constructed.

Note that LM_1 and LM_2 can also be expressed as

$$LM_1 = \sum_{t=1}^{T} \tilde{d}_t \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big)' \left(\sum_{t=1}^{T} \tilde{d}_t^2 \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big) \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big)' \right)^{-1} \sum_{t=1}^{T} \tilde{d}_t \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big),$$

and

$$LM_2 = \sum_{t=1}^{T} \tilde{d}_t \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big)' \left(\sum_{t=1}^{T} \frac{\tilde{\phi}_t^2}{\tilde{\Phi}_t(1-\tilde{\Phi}_t)} \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big) \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big)' \right)^{-1} \sum_{t=1}^{T} \tilde{d}_t \Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \Big),$$

where

$$\tilde{d}_t = \frac{y_t - \tilde{\Phi}_t}{\tilde{\Phi}_t(1-\tilde{\Phi}_t)} \tilde{\phi}_t.$$

This shows that the derivative term $\partial \pi_t(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ evaluated at $\tilde{\boldsymbol{\theta}}$ is central for the test statistics. The derivative in model (5) is

$$\frac{\partial \pi_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial \pi_t(\boldsymbol{\theta})}{\partial \omega} \\ \frac{\partial \pi_t(\boldsymbol{\theta})}{\partial \alpha_1} \\ \frac{\partial \pi_t(\boldsymbol{\theta})}{\partial \delta_1} \\ \frac{\partial \pi_t(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} 1 + \alpha_1 \frac{\partial \pi_{t-1}(\boldsymbol{\theta})}{\partial \omega} \\ \pi_{t-1}(\boldsymbol{\theta}) + \alpha_1 \frac{\partial \pi_{t-1}(\boldsymbol{\theta})}{\partial \alpha_1} \\ y_{t-1} + \alpha_1 \frac{\partial \pi_{t-1}(\boldsymbol{\theta})}{\partial \delta_1} \\ \boldsymbol{x}_{t-1} + \alpha_1 \frac{\partial \pi_{t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \end{pmatrix}$$

and under H_0, it is

$$\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \omega} \\ \frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \alpha_1} \\ \frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \delta_1} \\ \frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} 1 \\ \pi_{t-1}(\tilde{\boldsymbol{\theta}}) \\ y_{t-1} \\ \boldsymbol{x}_{t-1} \end{pmatrix}.$$

4. Simulation Results

The two LM tests described in the previous section are asymptotically equivalent. In this section, the small-sample properties of the LM tests are studied by simulation.[1] We simulated realizations from the Bernoulli distribution (1) using two different models[2]

$$\pi_t(\boldsymbol{\theta}) = -0.30 + \alpha_1 \pi_{t-1}(\boldsymbol{\theta}) + 0.50\, y_{t-1} \tag{12}$$

and

$$\pi_t(\boldsymbol{\theta}) = -0.30 + \alpha_1 \pi_{t-1}(\boldsymbol{\theta}) + 1.00\, y_{t-1} - 0.20\, x_{t-1}, \tag{13}$$

where

$$x_t = 0.1 + 0.90 x_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{NID}(0,1).$$

Positive coefficients for the lagged value of y_t, y_{t-1}, in (12) and (13) indicate that the realized values of y_t, i.e. zeros and ones, tend to cluster in the same way as, for example, recession periods of the economy (see the empirical application in Section 5).

We provide simulation evidence for sample sizes 150, 300, 500, 1000 and 2000. For all generated series, 200 extra observations were simulated and discarded from the beginning of every sample to avoid initialization effects. We report empirical sizes of the models at 10%, 5% and 1% significance levels. All results are based on 2000 replications. However, in some cases a little more than 2000 replications (about 20–30 replications) are needed because of numerical difficulties in the optimization of the log-likelihood function (6).

Empirical sizes of the LM tests in models (12) and (13) are presented in Tables 1 and 2. Both tests seem to be rather severely oversized in small samples, but for larger samples, the empirical sizes are rather close to the nominal levels.

---

[1] Matlab version 7.5.0 and the BFGS algorithm in the Optimization Toolbox is used in simulation and estimation. Eviews code for computing LM tests (9) and (11) is also available upon request.

[2] The initial value $\pi_0(\boldsymbol{\theta})$ in (5) is set to in a similar way as suggested by Kauppi and Saikkonen (2008) where $\pi_0(\boldsymbol{\theta}) = (\omega + \delta_1 \bar{y} + \bar{\boldsymbol{x}}_{t-k}\boldsymbol{\beta})/(1-\alpha_1)$ with the parameter values used in (12) and (13). A bar is used to denote the sample mean of the considered variables.

Table 1: Empirical size of the LM_1 and LM_2 tests in the model (12).

T	LM_1			LM_2		
	10%	5%	1%	10%	5%	1%
150	28.5	15.3	3.1	28.9	14.7	3.2
300	19.6	9.0	1.5	19.3	8.9	1.4
500	17.0	8.5	1.4	16.8	8.4	1.3
1000	14.3	6.6	1.1	14.3	6.6	1.1
2000	10.3	5.3	1.2	10.3	5.4	1.2

Notes: In size simulations, $\alpha_1 = 0$. The results are based on the 2000 replications.

Table 2: Empirical size of the LM_1 and LM_2 tests in the model (13).

T	LM_1			LM_2		
	10%	5%	1%	10%	5%	1%
150	42.8	26.3	7.1	41.6	23.0	5.0
300	30.1	15.2	3.3	28.4	14.5	2.3
500	22.0	10.8	2.2	21.0	10.3	2.0
1000	14.0	7.6	1.5	13.7	7.3	1.3
2000	11.4	5.7	0.9	11.4	5.3	0.9

Notes: See notes to Table 1.

Rejection rates presented in Tables 1 and 2 are based on the critical values from the asymptotic χ_1^2 distribution. However, one can use a parametric bootstrap method to obtain alternative, potentially more accurate, critical values. The employed procedure is the following. The ML estimates $\tilde{\boldsymbol{\theta}} = (\tilde{\omega} \quad \tilde{\delta} \quad \tilde{\boldsymbol{\beta}})'$ are computed under H_0 and bootstrap samples y_τ^b, and the values of test statistics LM_1^b and LM_2^b, $b = 1, 2, ..., B$, are obtained by using

$$y_\tau^b \sim B(\Phi(\pi_\tau^b(\tilde{\boldsymbol{\theta}}))), \tag{14}$$

where $\tau = 1, 2, ..., T$, and

$$\pi_\tau^b(\tilde{\boldsymbol{\theta}}) = \tilde{\omega} + y_{\tau-1}^b \tilde{\delta} + \boldsymbol{x}_{\tau-1}' \tilde{\boldsymbol{\beta}}.$$

Finally, bootstrap critical values at different significance levels are obtained from the empirical distribution of the test statistics LM_1^b and LM_2^b. The number of bootstrap replications B is set to 500 and the simulation is carried out for 500 replications.

As an illustration for the usefulness of the proposed bootstrap method, Table 3 presents the rejection rates based on the bootstrap critical values for the three smallest sample sizes. Compared with the results shown in Tables 1 and 2, the empirical sizes of the LM tests are now much closer to the nominal values.

Table 3: Empirical size of the LM_1 and LM_2 tests using the model (13) and bootstrap critical values.

	LM_1			LM_2		
T	10%	5%	1%	10%	5%	1%
150	9.0	4.2	1.0	9.6	6.2	1.0
300	9.6	5.4	1.6	9.0	6.2	1.0
500	12.2	5.0	1.2	11.2	6.4	0.4

Size-adjusted empirical power functions with different sample sizes at the 5% level are depicted in Figures 1 and 2. In many applications, the parameter α_1 is expected to be non-negative and, therefore, we concentrate on values from $\alpha_1 = 0.00$ up to $\alpha_1 = 0.80$.[3] The power seems to increase rather quickly when the value of α_1 increases, in particular when the explanatory variable x_t is employed in the model. The power of LM_2 is typically slightly higher than that of LM_1 although the differences are minor.

5. Application: U.S. Recession Forecasting Models

Forecasting recession periods has been one of the most common empirical applications of binary time series models. In this application, the dependent variable is a binary recession indicator y_t which takes the value 1 when the economy is in a recession and 0 otherwise. Predicting the direction-of-change in stock market returns is an example of another potential application (see, e.g., Leung, Daouk and Chen 2000, Rydberg and Shephard 2003, and Nyberg 2010b).

[3] The evidence appears to be rather the same with the negative values of α_1, especially in large samples.

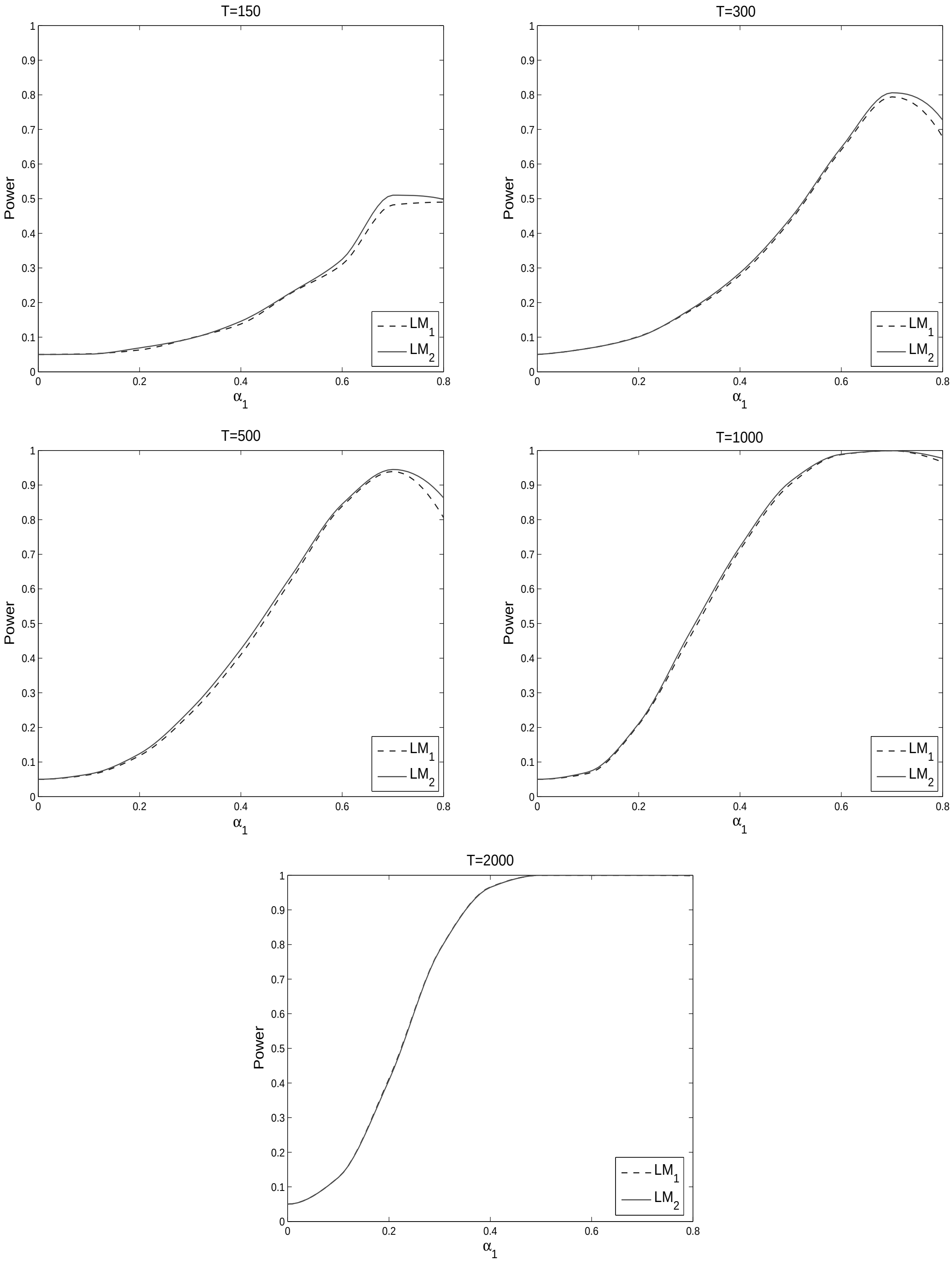

Figure 1: Empirical power in the case of model (12).

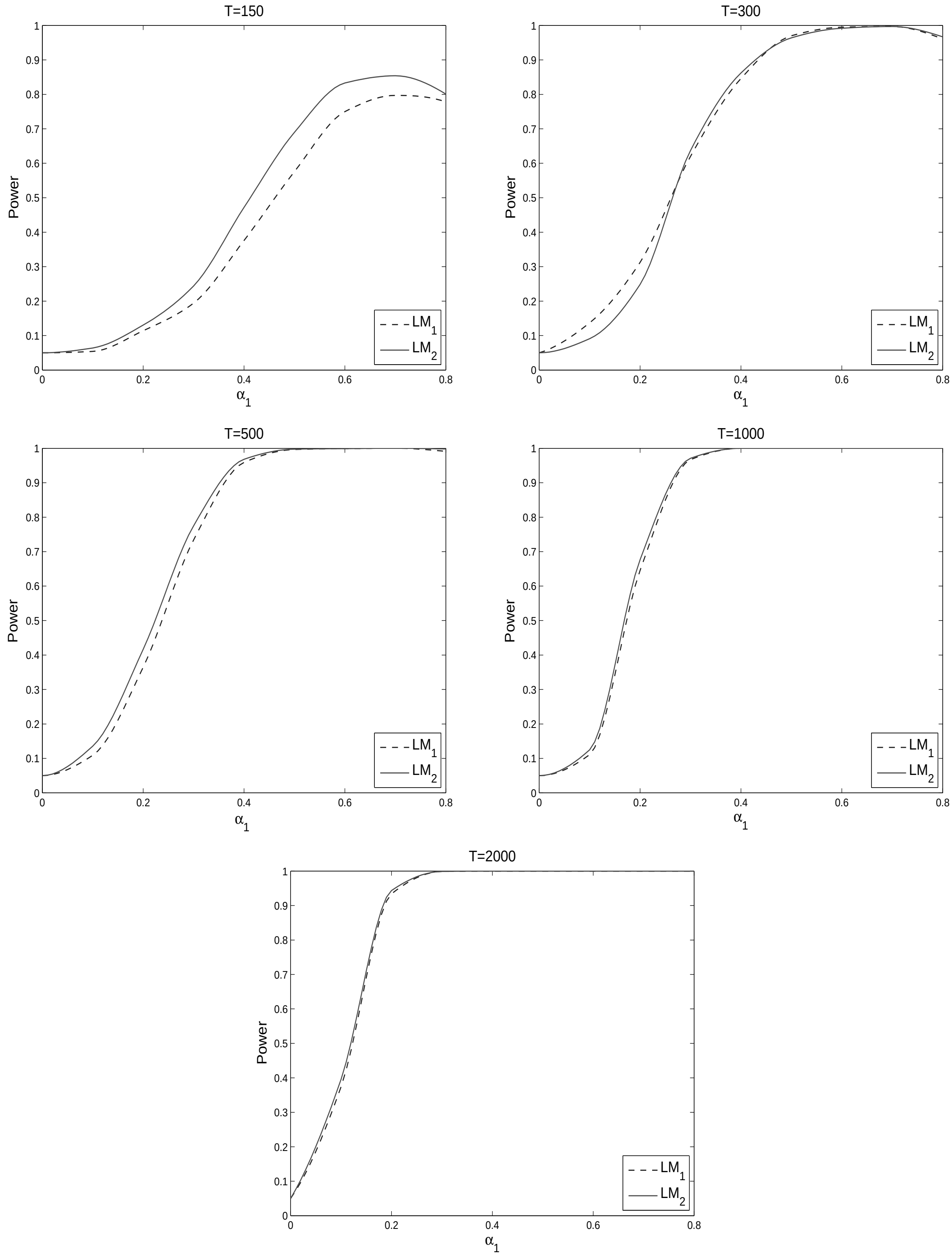

Figure 2: Empirical power in the case of model (13).

Although in this study we are not interested in out-of-sample forecasting, we consider forecasting models behind the "direct" (using a forecast horizon-specific predictor y_{t-15}) and "iterative" (y_{t-1}) multi-step forecasts for the recession indicator (for details, see Kauppi and Saikkonen 2008). The difference between "direct" and "iterative" forecasts is similar to that in time series models for traditional continuous variables (see, e.g., Marcellino, Stock, and Watson 2006).

Table 5 presents the estimated predictive models using the U.S. data described in more detail in Table 4. The forecast horizon is assumed to be six months. The fact the NBER business cycle turning points are announced with a delay is also taken into account in "direct" forecasting models.[4]

Table 4: U.S. dataset.

y_t	NBER recession indicator ($y_t = 1$ denotes a recession)
R_t	10-year Treasury bond yield rate, constant maturity
i_t	Three-month Treasury bill rate, secondary market
SP_t	Term spread, $R_t - i_t$
r_t	Monthly stock market return, log-difference of the S&P500 index

Notes: Recession ($y_t = 1$) and expansion ($y_t = 0$) periods are obtained from the business cycle chronology provided by the National Bureau of Economic Research (see details at http://www.nber.org/cycles/cyclesmain.html). Interest rates are from http://www.federalreserve.gov/releases/h15/data.htm. S&P500 stock index is taken from http://finance.yahoo.com and http://www.econstats.com.

In Table 5, outcomes of the two LM tests are in accordance with the Wald and likelihood ratio test at a 5% significance level. The recommendation is that an autoregressive model structure is worth considering as an alternative to a standard recession prediction model (see, e.g., Estrella and Mishkin 1998, and Bernard and Gerlach 1998), possibly augmented by the forecast horizon-specific lagged value y_{t-15}, as presented in Models 1 and 2. However, in Model 4 used in iterative out-of-sample forecasting approach (see Kauppi and Saikkonen 2008), the estimated coefficient for $\pi_{t-1}(\boldsymbol{\theta})$ is statistically insignificant and the lagged state of the economy, y_{t-1}, seems to be the main predictor.

[4] We assume that this "publication lag" is nine months. For further details see Kauppi and Saikkonen (2008), Kauppi (2008), and Nyberg (2010a).

Table 5: Estimation results for the recession prediction models.

Model		1	2	3	4
constant		-0.50	-0.02	-1.71	-2.00
		(0.16)	(0.04)	(0.17)	(0.24)
SP_{t-6}		-0.61	-0.21	-0.58	-0.67
		(0.13)	(0.05)	(0.14)	(0.16)
r_{t-6}		-0.05	-0.08	-0.01	-0.01
		(0.02)	(0.02)	(0.03)	(0.03)
$\pi_{t-1}(\boldsymbol{\theta})$			0.81		-0.17
			(0.03)		(0.09)
y_{t-1}				3.38	3.95
				(0.25)	(0.35)
y_{t-15}		-0.41	-0.07		
		(0.36)	(0.16)		
log-L		-185.94	-136.60	-55.63	-55.29
$\text{pseudo} - R^2$		0.192	0.367	0.689	0.691
LM_1		25.25		2.75	
p-value		0.000		0.097	
LM_2		36.18		0.65	
p-value		0.000		0.420	
Bootstrap					
critical values					
LM_1	10 %	3.06		5.10	
	5 %	3.93		6.88	
	1 %	5.90		9.61	
LM_2	10 %	2.71		2.30	
	5 %	3.75		3.07	
	1 %	5.37		6.45	

Notes: Models are estimated using the U.S. data from 1954 M01 to 2006 M12 ($T = 636$). First 21 months are used as initial values. Robust standard errors (see Kauppi and Saikkonen 2008) are reported in parentheses. The estimated value of the log-likelihood function (6) and pseudo-R^2 measure (Estrella 1998), which is a counterpart to the coefficient of determination in models for continuous dependent variables, are also provided as well as the values of the LM_1 and LM_2 test statistics, their p-values based on the asymptotic χ^2_1 distribution, and critical values obtained by bootstrap.

6. Conclusions

We have proposed LM tests for testing an autoregressive model structure in binary time series models. Based on a limited simulation study, the tests appear to have reasonable empirical size, especially in large samples, and high power. For small samples, the proposed bootstrap simulation method provides improved empirical sizes. An empirical example of recession forecasting models for the U.S. illustrates the use of the LM test and provides evidence that the inclusion of an autoregressive model structure may be a useful addition to the recession prediction model.

References

Bernard, H and S. Gerlach (1998) "Does the term structure predict recessions? The international evidence" *International Journal of Finance and Economics* **3**, 195–215.

Chauvet, M and S. Potter (2005) "Forecasting recession using the yield curve" *Journal of Forecasting* **24**, 77–103.

Cox, D.R. (1981) "Statistical analysis of time series: Some recent developments." *Scandinavian Journal of Statistics* **8**, 93–115.

Davidson, R and J.G. MacKinnon (1984) "Convenient specification tests for logit and probit models" *Journal of Econometrics* **25**, 241–262.

Estrella, A. (1998) "A new measure of fit for equations with dichotomous dependent variables" *Journal of Business and Economic Statistics* **16**, 198–205.

Estrella, A and F.S. Mishkin (1998) "Predicting U.S. recessions: Financial variables as leading indicators" *Review of Economics and Statistics*, **80**, 45–61.

Kauppi, H. (2008) “Yield-curve based probit models for forecasting U.S. recessions: Stability and dynamics” HECER Discussion Paper number 221. Helsinki Center of Economic Research.

Kauppi, H and P. Saikkonen (2008) “Predicting U.S. recessions with dynamic binary response models” *Review of Economics and Statistics* **90**, 777–791.

Leung, M.T, Daouk, H and A.S. Chen (2000) “Forecasting stock indices: A comparison of classification and level estimation models” *International Journal of Forecasting* **16**, 173–190.

Marcellino, M, Stock J.H and M.W. Watson (2006) “A comparison of direct and iterated AR methods for forecasting macroeconomic time series” *Journal of Econometrics* **135**, 499–526.

Nyberg, H. (2010a) “Dynamic probit models and financial variables in recession forecasting” *Journal of Forecasting* **29**, 215–230.

Nyberg, H. (2010b) “Forecasting the direction of the U.S. stock market with dynamic binary probit models” *International Journal of Forecasting*, forthcoming.

Rydberg, T and N. Shephard (2003) “Dynamics of trade-by-trade price movements: Decomposition and models” *Journal of Financial Econometrics* **1**, 2–25.

Startz, R. (2008) “Binomial autoregressive moving average models with an application to U.S. Recessions” *Journal of Business and Economic Statistics* **26**, 1–8.