



Testing linear regression model with AR(1) errors against a first-order dynamic linear regression model with white noise errors: A point optimal testing approach[☆]



Sivagowry Srianthakumar^{*}

School of Economics, Finance and Marketing, RMIT University, GPO Box 2476, Melbourne 3001, Australia

ARTICLE INFO

Article history:

Accepted 14 March 2013

JEL classification:

C1
C2
C4
C5
C8

Keywords:

Approximate point optimal test
Marginal likelihood based tests
Simulated annealing algorithm
Monte Carlo simulations
Generalized Neyman–Pearson lemma
Nuisance parameters

ABSTRACT

We know very little about the performance of point optimal (PO) and approximate point optimal (APO) tests in the presence of unavoidable nuisance parameters. Because marginal likelihood based tests are said to perform well in the presence of unavoidable nuisance parameters, this paper compares the performance of marginal likelihood based APO tests and classical tests using a testing problem which has been largely overlooked by econometric practitioners, namely testing for a static linear regression model with AR(1) errors against a dynamic linear regression model with white noise errors. It is well known that the classical tests are specifically designed for nested testing, they are applied to test for the significance of the dynamic coefficient of a dynamic linear regression model with AR(1) errors.

A testing procedure is proposed, where the size and power comparisons used are based on near-exact non-similar critical values of tests obtained using the simulated annealing (SA) algorithm, as the near-exact non-similar critical values control the sizes of the tests well overall.

Among marginal likelihood based classical tests, the likelihood ratio (LR) test and Lagrange multiplier (LM) test seem to perform well under the alternative hypothesis, particularly when the dynamic parameter is large and the sample size is reasonably big. The Wald (W) test is the worst performer overall. This concurs with previous observations that the W test performs poorly in small samples. Compared to the classical approach, APO tests appear to have good power properties, particularly in the neighborhood of the chosen parameter point under the alternative hypothesis. This finding may advance the use of PO and APO tests.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

In the context of a linear regression model, researchers have often tried to correct for autocorrelation by including a first-order autoregressive (AR(1)) or first-order moving average (MA(1)) error term, when the Durbin and Watson (1950, 1951) (DW) statistic for autocorrelation is significant. They have then estimated the model with either generalized least squares (GLS) or the maximum likelihood (ML) method. However, a number of researchers have warned that if the true model is dynamic, a substantial bias in estimation can occur, and vice-versa (De Boef and Keele, 2008; Keele and Kelly, 2006; McGuirk and Spanos, 2002).

As McGuirk and Spanos (2002) and Spanos (1998, 1999) point out, rejecting the null hypothesis in a misspecification test such as the DW test does not indicate that the alternative model is true; only that the original model may be misspecified. That is, in a linear regression

model context, a significant DW test indicates that the data possess some kind of dependence. The type of dependence present in the data can be captured with several alternative statistical models which allow for such dependence. One such alternative is the dynamic linear regression model. An advantage of this model is that it (almost) nests the autocorrelation corrected linear model (McGuirk and Spanos, 2002). King and Rankin (1993) find that when the DW test in a linear model is significant, and the true model is the dynamic linear model with a large autoregressive parameter, a substantial loss in accuracy of prediction can occur, if one proceeds to correct for AR(1) disturbances without checking for the possibility of a dynamic linear model with white noise errors. In particular they show that when the true model is the dynamic linear model, correcting for serial correlation (or vice-versa) can result in substantially biased estimates. Similar findings are reported by Keele and Kelly (2006) and McGuirk and Spanos (2002).

The AR(1) model versus first-order dynamic linear model are widely used in political science (Beck and Katz, 2009; Kono, 2006). The dynamic model is often used in this discipline to remove autocorrelation (Blaydes, 2011) and is often considered theoretically more appropriate because past behavior/performance is believed to influence current political decision making. Arbitrarily using the dynamic linear model when the true model is AR(1) (or vice-versa) may produce biased

[☆] An earlier version of this paper was presented at the Econometric Society Australasian Meeting (July 2012) in Melbourne, Australia. I would like to thank the participants at this conference for their valuable comments.

^{*} Corresponding author. Tel.: +61 39925 1456; fax: +61 3 9925 5624.
E-mail address: Sivagowry.Srianthakumar@rmit.edu.au.

estimates, particularly in finite samples. While not used as extensively in econometrics, a number of researchers have indicated that there are different ways to model the data when significant autocorrelation is detected (Heaney and Srianthakumar, 2012; Jones, 1993a,b; Srianthakumar and King, 2006). Jones (1993a,b)¹ uses US fuel consumption and passenger vehicle use data to conduct a detailed study involving several possible models to illustrate this point. His models include those considered here.² Jones (1993a) stresses the importance of selecting a model which is consistent with the data and recommends the general-to-simple modeling approach of Hendry (1979) to achieve this. He also warns that parameter estimates from data-rejectable models may give very misleading indications of the dynamic nature of the behavioral relationships being modeled. Jones observes that researchers often tend to use some familiar specification they have been trained to use, or find easy to implement, and some might appeal to other researchers' success with a popular specification. Sadly this trend continues even now. Jones (1993a) insists that untested acceptance of a popular model specification may lead to erroneous inferences. For example, in his study, a popular dynamic specification results in reasonable and significant estimates. However, close scrutiny reveals that the model is not data-acceptable because it either over-estimated or under-estimated important economic parameters, such as the long run price elasticity and the speed of the adjustment process.

Jones (1993b) double-checked Greene's (1992) selection of model for the US passenger vehicle use data using different modeling strategies. This motivates one to question an outcome (even if it is suggested by a well-known researcher) rather than accepting it at face value. However, finding a data acceptable model is not that simple. It is worth noting that if the true model which generated the given sample is linear with AR(1) errors, and one fits the dynamic linear model with white noise errors to the sample by the ordinary least squares (OLS) method, then it is likely to explain the data rather well (Giles, 1975; Griliches, 1967). Therefore usual significance testing and a goodness of fit measure are not going to be helpful in this context. Because correct model specification is important for forecasting, and also for the purpose of further inference, it is desirable to have a powerful testing procedure to distinguish between plausible alternative models.

For this type of problem nuisance parameters cannot be avoided. This may make the test non-similar in the sense that the test's size³ varies with the value of these nuisance parameters. Different approaches when testing in the presence of nuisance parameters are suggested in the literature. The classical approach to non-similar tests is to find exact non-similar critical values, for which sizes are never greater than the nominal significance level for all possible values of the nuisance parameters. Such critical values are typically obtained using the Monte Carlo method (Inder, 1986; King and McAleer, 1987; Silvapulle, 1991). Other popular approaches include using bounds type tests and confidence intervals, as suggested in Dufour (1990) and Pesaran et al. (2001), and replacing unknown nuisance parameters with consistent estimates and then relying on asymptotic theory (Moreira, 2009). However, exact bounds tests can be less powerful than non-similar tests. Forchini (2005) shows that any test, with size bounded from above by a known constant, has potentially very low power and a large type II error.

Kiviet and Dufour (2003) and Dufour (2006) suggest a maximized Monte Carlo (MMC) test that maximizes a simulated p value over the nuisance parameter space using SA. Dufour (2006) also suggests (and proves) asymptotically valid MMC (AMMC) tests that use consistent set estimators of nuisance parameters.⁴ The maximum p value of an

AMMC test can be obtained by maximizing the p value over a subset of the nuisance parameter space (for example a confidence set of the nuisance parameters) instead of the entire nuisance parameter space.⁵ Thus, compared to the MMC approach, the AMMC approach will be less time-consuming (Phipps and Byron, 2007). Although the MMC approach is gaining popularity (Beaulieu, Dufour and Khalaf, in press; Dufour and Tarek, 2006; Dufour and Valéry, 2009; Frederic and Olivier, 2006; Thomas et al., 2007), it is criticized for the following reasons: (1) it can be computationally demanding, (2) MMC-based actual rejection frequency may be very much less than the level of the test and may in consequence be severely lacking power and (3) there is a possibility of getting a much larger p value for nuisance parameter values far away from the ones that actually generated the data (MacKinnon, 2009).

This paper adopts the classical approach. In particular, near-exact non-similar critical value-based size and power properties are analyzed, because obtaining exact non-similar critical values involves a large amount of computation (Inder, 1986).

Studies, such as those of Tunnicliffe Wilson (1989), Rahman and King (1997) and Francke and de Vos (2007), indicate that in the presence of nuisance parameters marginal likelihood-based tests perform better than conventional likelihood-based tests in finite samples. On the other hand, if the null hypothesis and alternative hypothesis spaces are tightly restricted, King's (1987) PO tests can be most powerful at a chosen parameter point under the alternative hypothesis.

However, PO tests cannot always be constructed when testing a composite null hypothesis. Srianthakumar and King (2006) propose an APO test, called the g test, for this instance. The performance of PO and APO tests in the presence of unavoidable nuisance parameters is largely unknown. Theoretically, PO tests can be expected to be most powerful in the absence of nuisance parameters. If PO tests are found to work well in the presence of unavoidable nuisance parameters, this will advance their use in practice.

The main contributions of this paper are the study of the finite sample performance of the APO test and marginal likelihood-based classical tests in the presence of unavoidable nuisance parameters; the construction of tests of an interesting, but largely overlooked, issue: testing for a static linear regression model with AR(1) errors against a dynamic linear regression model with white noise errors⁶; the comparison of different information measures in the context of marginal likelihood estimation in order to decide which is more accurate; and the application of near-exact non-similar critical values of the tests obtained using SA.

The models and assumptions are discussed next; then descriptions of the APO test and marginal likelihood based classical tests are provided. Section 4 briefly introduces SA and discusses how exact and near-exact non-similar critical values can be obtained. Section 5 presents the details of the Monte Carlo experiment and summarizes the main findings. An illustrative application of the g test to a specific data set is presented in Section 6. Finally, concluding remarks reflect on the use of the APO test and marginal likelihood-based classical tests in the presence of unavoidable nuisance parameters.

2. Model and assumptions

Suppose we wish to test a linear regression model with AR(1) errors

$$H_0 : y_t = x_t' \beta + u_t, \quad u_t = \rho u_{t-1} + e_t, \quad t = 1, \dots, n, \quad (1)$$

against a dynamic linear regression model with white noise errors

$$H_a : y_t = \mu y_{t-1} + x_t' \beta + e_t, \quad t = 2, \dots, n, \quad (2)$$

¹ Jones' study is worth mentioning because, unlike other studies, Jones uses a number of different seemingly acceptable models for the same data and explains the importance of correct modeling. His study is a valuable guide for learning model building.

² However Jones did not consider the possibility of moving average errors. This is also a possibility and can be easily tested (Srianthakumar and King, 2006).

³ Throughout this paper the term size is used to denote probability of a type I error.

⁴ Always feasible as long as consistent point estimates of the nuisance parameters are available.

⁵ This idea was originally suggested and investigated by Berger and Boos (1994).

⁶ This testing problem can be easily generalized to include any competing models.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات