Journal of Econometrics 172 (2013) 142–157

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Jackknife estimation of stationary autoregressive models

Marcus J. Chambers *

Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, England, United Kingdom

ARTICLE INFO

Article history:
Received 27 January 2010
Received in revised form
26 March 2012
Accepted 19 September 2012
Available online 26 September 2012

JEL classification:
C12
C13
C22

Keywords:
Jackknife
Bias
Autoregression

ABSTRACT

This paper explores the properties of jackknife methods of estimation in stationary autoregressive models. Some general results concerning the correct weights for bias reduction under various sampling schemes are provided and the asymptotic properties of a jackknife estimator based on non-overlapping sub-samples are derived for the case of a stationary autoregression of order p when the number of sub-samples is either fixed or increases with the sample size at an appropriate rate. The results of a detailed investigation into the finite sample properties of various jackknife and alternative estimators are reported and it is found that the jackknife can deliver substantial reductions in bias in autoregressive models. This finding is robust to departures from normality, ARCH effects and misspecification. The median-unbiasedness and mean squared error properties are also investigated and compared with alternative methods as are the coverage rates of jackknife-based confidence intervals.

1. Introduction

Jackknife techniques have a long history in statistics. The jackknife method of bias reduction was originally proposed by Quenouille (1956) with Tukey (1958) subsequently demonstrating how the method could also be used to construct a non-parametric estimator of variance. As a result it is often referred to as the Quenouille–Tukey jackknife; see, for example, Efron (1982, p. 1). According to Miller (1964, p. 1594) the procedure was named the jackknife by Tukey because "a boy scout's jackknife is symbolic of a rough-and-ready instrument capable of being utilized in all contingencies and emergencies." The applicability of the jackknife is certainly widespread but it has found fewer applications in econometrics than rival bootstrap methods. Indeed, Efron (1979) demonstrated that the jackknife is a linear approximation method for the bootstrap in the case of estimating the sampling distribution of a random variable based on a sample of i.i.d. (independently and identically distributed) data, a result that has perhaps been interpreted as favouring the bootstrap in a wider variety of situations than that to which this result relates. Moreover, as will be shown below, the standard formulation of the jackknife statistic is applicable only in the case of i.i.d. data, which may also help to explain its limited application in econometrics.

Notwithstanding the preceding comments and the proliferation of bootstrap methods in econometrics, there has recently been a realisation that jackknife methods can be effective in reducing the bias of estimators in models of interest in econometrics. In models with more instruments than endogenous variables Angrist et al. (1999) proposed the jackknife instrumental variables estimator and demonstrated its superior finite sample properties compared to the two-stage least squares estimator and its comparability to the limited information maximum likelihood estimator, although the performance of this estimator has subsequently been criticised by Davidson and MacKinnon (2006). Hahn et al. (2003) considered both bootstrap and jackknife bias corrections to maximum likelihood estimators based on an i.i.d. sample while applications to panel data models (including nonlinear and dynamic models) have been considered by Hahn and Newey (2004), Hahn and Moon (2006) and Dhaene et al. (2006). Jackknife methods have also been applied to maximum likelihood estimators of the parameters of continuous time models of the short-term interest rate by Phillips and Yu (2005) who also demonstrate the resulting gains that can be made by applying such techniques directly to the implied bond options prices. Based on the encouraging results obtained in the above situations this paper explores the properties of jackknife methods of estimation and inference in stationary autoregressive (AR) models. In the context of stationary time series Carlstein (1986) proposed an estimator of variance based on non-overlapping blocks while Künsch (1989) considered both jackknife and bootstrap methods of estimating standard errors by

* Tel.: +44 1206 872756; fax: +44 1206 872724.
E-mail address: mchamb@essex.ac.uk.

0304-4076/$ – see front matter
doi:10.1016/j.jeconom.2012.09.003

deleting whole blocks of observations. The focus here, however, is ultimately concerned more with issues of estimation of the parameters in AR models than it is with variance estimation, although the latter becomes important when using the jackknife estimator for inference.

Some general theoretical results on jackknife methods applied to a statistic of interest (such as an estimator of a parameter or a test statistic) are given in Section 2. The first result (Theorem 1) shows how the full-sample and sub-sample statistics should be combined in order to eliminate the first-order bias in a general setting before considering specific sampling situations such as i.i.d. data as well as non-overlapping and moving-block sub-samples which are of particular relevance in time series settings. A further refinement (Theorem 2) shows how statistics from different sub-sampling methods, or from the same sub-sampling method with different numbers of sub-samples, can be combined to eliminate both first- and second-order bias from the statistic of interest. Specific cases of sub-sampling are also considered, and a further general result (Theorem 3) shows how the jackknife weights need to be modified in cases where sub-samples of unequal lengths are encountered, this being potentially important in empirical applications.

Section 3 explores jackknife methods of estimation in stationary autoregressive models, the focus being on the p'th order model with an intercept. The motivation for employing jackknife estimators in this context is rooted in analytical work that provides Nagar-type expansions for the bias of the ordinary least squares (OLS) estimator of the AR parameter vector. Theorem 4 derives the limiting distribution of the jackknife estimator based on non-overlapping sub-samples of the type utilised by Phillips and Yu (2005) and shows that it has the same form as the OLS estimator irrespective of whether the number of sub-samples is fixed or increases with the sample size at an appropriate rate. Hence the jackknife estimator has the potential to reduce the finite sample bias without any loss of asymptotic efficiency, although the effect on the finite sample mean squared error (MSE) is unknown (but is explored in simulations in Section 5). Shao and Tu (1995, pp. 66–67) provide examples where the jackknife statistic can have either a larger or smaller MSE than the underlying statistic, concluding that, in general, "the relative performance … is indefinite and depends on the unknown population" (p. 67) and, furthermore, "we should keep in mind that the jackknife estimator … is designed to eliminate bias and, therefore, can be used when the bias is an important issue. We need to balance the advantage of unbiasedness against the drawback of a large mean squared error" (pp. 67–68). The limiting distribution in Theorem 4 can be used as the basis for inference provided an appropriate estimator of the asymptotic variance matrix can be obtained, and two possibilities are provided in Theorem 5.

Section 4 reports the results of an extensive simulation exercise (involving 100,000 replications) using the AR(1) model in an attempt to obtain evidence on a number of issues, including: which sub-sampling method produces the greatest bias reduction; the optimal number of sub-samples to employ; how the optimal number of sub-samples varies with sample size; and the extent of additional bias reduction that can be achieved by attempting to eliminate the second-order bias. The results cover a range of sample sizes and a range of positive values for the AR parameter that approaches the boundary of the stationarity region, these being of greatest empirical relevance in economics and finance. The analysis of bias reduction using the jackknife when a unit root is present can be found in Chambers and Kyriacou (2012). Comparisons of the jackknife estimators are also made with respect to the exact median unbiased (MU) estimator of Andrews (1993) and a recursive-design wild bootstrap estimator based on Gonçalves and Kilian (2004). The jackknife estimators are shown to result in the smallest bias in all cases considered. Section 4 also examines the robustness of the results to departures from normality, using the parameters of the Gamma distribution to control the degree of skewness and kurtosis, as well as to autoregressive conditional heteroskedasticity (ARCH) and to higher-order and misspecified autoregressions.

Additional considerations concerning the performance of the jackknife (and other) techniques are explored in Section 5. Although designed to reduce bias other distributional aspects are important to the usefulness of an estimator, and so the median-unbiasedness and mean squared error are examined first. Simulations reveal that it is possible to obtain an MSE less than the full-sample OLS estimator by using jackknife (and other) estimators, a feature of the jackknife estimators being that a larger number of sub-samples is required to minimise root MSE (RMSE) than to minimise bias. It is also shown that the distributions of the jackknife estimators are much closer to being median-unbiased than those of the OLS estimator, the latter being significantly negatively biased particularly for larger values of the AR parameter. Section 5 also looks at the coverage rates of jackknife confidence intervals based on the asymptotic distribution in Theorem 4 and compares them to those of OLS, MU and bootstrap methods. Proofs of all Theorems are contained in Appendix A, while Appendix B contains supplementary results that are used in the proofs and elsewhere. Section 6 concludes.

The following notation will be used throughout the paper. The symbol $\xrightarrow{p}$ denotes convergence in probability; $\xrightarrow{d}$ denotes convergence in distribution; and, for a $k \times 1$vector x, $\|x\| = \left(\sum_{i=1}^k x_i^2\right)^{1/2}$ denotes the Euclidean norm.

2. Jackknife methods: some general results

The idea behind the jackknife method of bias reduction is to combine a statistic based on a full sample of data with a set of statistics based on sub-samples in a way that eliminates the first-order bias. The statistic of interest, $\hat{\beta}_n$, is often an estimator of a parameter or parameter vector although functions of model parameters and test statistics, for example, can also be considered provided they satisfy (or are assumed to satisfy) certain properties. The following general result for the jackknife statistic will be used to deal with specific cases of interest.

Theorem 1. *Let $y = (y_1, \dots, y_n)'$ be a sample of n observations on a random variable and let $\hat{\beta}_n = \beta(y)$ denote the statistic of interest satisfying*

$$E(\hat{\beta}_n) = \beta + \frac{a_1}{n} + \frac{a_2}{n^2} + O\left(n^{-3}\right), \quad (1)$$

where a_1 and a_2 are constants. Let $Y_i (i = 1, \dots, m)$ denote a set of sub-samples of y, each of which has equal length $\ell = O(n)$, and let $\hat{\beta}_i = \beta(Y_i)$ $(i = 1, \dots, m)$ denote the corresponding sub-sample statistics. Then the jackknife statistic

$$\hat{\beta}_J = \left(\frac{n}{n-\ell}\right)\hat{\beta}_n - \left(\frac{\ell}{n-\ell}\right)\frac{1}{m}\sum_{i=1}^m \hat{\beta}_i \quad (2)$$

satisfies $E(\hat{\beta}_J) = \beta + O(n^{-2})$.

Theorem 1 is a general result that holds for both i.i.d. samples as well as dependent samples of the type arising in time series. The expression for bias in (1) can usually be justified by a Nagar-type expansion; see, for example, Bao and Ullah (2007) for some results in the general time series setting and Bao (2007) for the AR(1) model under general error distributions. Some specific cases will now be considered and Theorem 1 will be employed to determine