Journal of Econometrics 148 (2009) 114–123

Contents lists available at ScienceDirect

Journal of Econometrics

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

Simulation based selection of competing structural econometric models[☆]

Tong Li [*]

Department of Economics, Vanderbilt University, VU Station B #351819, Nashville, TN 37235-1819, United States

ARTICLE INFO

Article history:
Available online 11 October 2008

Keywords:
Model selection
Non-nested structural models
Simulated mean squared error of predictions

ABSTRACT

This paper proposes a formal model selection test for choosing between two competing structural econometric models. The procedure is based on a novel lack-of-fit criterion, namely, the simulated mean squared error of predictions (SMSEP), taking into account the complexity of structural econometric models. It is asymptotically valid for any fixed number of simulations, and allows for any estimator which has a $\sqrt{n}$ asymptotic normality or is n^{α}-consistent for $\alpha > 1/2$. The test is bi-directional and applicable to non-nested models which are both possibly misspecified. The asymptotic distribution of the test statistic is derived. The proposed test is general, regardless of whether the optimization criteria for estimation of competing models are the same as the SMSEP criterion used for model selection. A Monte Carlo study demonstrates good power and size properties of the test. An empirical application using timber auction data from Oregon illustrates the usefulness and generality of the proposed testing procedure.

1. Introduction

Model selection is an important component of statistical inference. It involves comparing competing models based on some appropriately defined goodness-of-fit or selection criterion. For the competing models that can be estimated by (conditional) maximum likelihood estimation (MLE), there has been a vast literature on model selection procedures, such as the Akaike (1973, 1974) information criterion (AIC), the Cox test (1961) and the Vuong (1989) likelihood ratio test, to name only a few. Another important development is the use of the encompassing principle in testing non-nested models assuming that one of them is correctly specified. See, e.g., Mizon and Richard (1986), and Wooldridge (1990), among others. For a comprehensive review of the literature, see Gourieroux and Monfort (1994) and Pesaran and Weeks (2001). In light of the development of new estimation methods in econometrics such as the generalized method of moments (GMM) and empirical likelihood estimation methods, which offer robust alternatives to the conventional MLE, recent work in model selection has attempted to develop procedures that can be used for models estimated by other methods than the MLE. For example, see Smith (1992) for extensions of the Cox test and the encompassing test to non-nested regression models that are both estimated by instrumental variables, Rivers and Vuong (2002) for the extension of Vuong's (1989) test to dynamic models, Kitamura (2002) for using empirical likelihood ratio-type statistics for testing non-nested conditional models, and Chen et al. (2003) for likelihood ratio tests between parametric and (unconditional) moment condition models.

These model selection tests have been found useful in some of structural microeconometric models, which have been developed in the last two decades and applied in such fields of modern economics as labor and industrial organization.[1] For example Vuong's (1989) likelihood ratio test has been used to select structural models both of which are estimated by MLE. See, e.g., Gasmi et al. (1992) for testing collusive behavior, Wolak (1994) for testing asymmetric information, and Li (2005) for testing binding reservation prices in first-price auctions, to name only a few. Also, Chen et al. (2003) develop a test to distinguish between a parametric model which can be estimated by the MLE and an unconditional moment model which can be estimated by the empirical likelihood method, and then apply their procedure to choose between a sequential search model and a non-sequential model. Despite these interesting applications of the aforementioned model selection tests, there are many other situations in which these model selection tests may not be applicable.[2] Such a gap can be mainly attributed to the

☆ I am grateful to Co-Editor Takeshi Amemiya, an associate editor, and three referees for their constructive comments that greatly improved the paper. I also thank D. Andrews, F. Diebold, Y. Kitamura, H. Pesaran, and seminar participants at Indiana University, University of Cambridge, University of Pennsylvania, and Yale University for helpful comments and discussions, and Bingyu Zhang for excellent assistance in Monte Carlo experiments conducted in this paper.

* Tel.: +1 615 322 3582; fax: +1 615 343 8495.
E-mail address: tong.li@vanderbilt.edu.

[1] Heckman (2001) gives an insightful discussion on the development and the issues on identification and inference of structural microeconometric models.

[2] For instance, Laffont et al. (1995) develop a simulated nonlinear least squares estimator to estimate a structural model of first-price auctions. They encounter

0304-4076/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2008.10.001

complexity associated with the nature of structural econometric models. Model selection criteria are formulated in such ways that they are calculated using sample information and compared between competing models. Most of the structural econometric models, however, are constructed based on economic theory which defines maps between the latent variable of interest or/and its distribution and the observables. For instance, in structural auction models, it is assumed that the observed bids are Nash–Bayesian equilibrium strategies which are strictly increasing functions of bidders' private valuations whereas identifying and estimating the private values distribution is one of the main objectives of the structural approach. The presence of latent variables and the complex relationship between the latent and observed variables defined by structural models make the formulation of a well-defined model selection criterion more involved. Moreover, in many cases, structural econometric models are constructed through moment conditions, meaning that they are estimated not by MLE but by GMM or method of simulated moments (MSM). Therefore, to accommodate these specific features arising from the nature of structural models and the estimation methods, new model selection tests need to be developed.

Developing model selection procedures suitable in distinguishing between competing structural models is especially relevant in using the structural approach to analyze economic data and make policy evaluations. In the structural approach, policy analysis and the resulting recommendations are based on a structural model that is closely derived from economic theory assuming that the involved economic agents are in the environment described by the theory and behave according to the theory. As a result, it is pivotal to validate the structural model under consideration. For example, when analyzing auction data using the structural approach, an econometrician faces choices among different paradigms such as a private value model or a common value model. Even within a chosen paradigm, the econometrician may also need to determine an appropriate parametric functional form for the latent distribution. Furthermore, the researcher sometimes needs to choose between different equilibria if multi-equilibria exist, as is the case for models of two-stage dynamic games which yield a large number of Bayesian perfect equilibria (Laffont and Maskin, 1990).

The goal of this paper is thus to propose a new model selection test in discriminating between competing structural econometric models. Our test is based on a comparison of the predictability of competing structural models. In time series literature, there has been a rich set of papers since Diebold and Mariano (1995) and West (1996) in using predictability for model evaluation. More recently, a general model selection framework based on predictability is developed in Rivers and Vuong (2002). Our test falls within this framework, and uses a similar MSEP criterion. On the other hand, given that structural econometric models usually contain some latent variables that are unobserved, we propose to simulate these latent variables in order to make the predictions on the equilibrium outcomes. Also, since the simulation is used, when formulating the sample analog to those population quantities, we need to correct for the asymptotic bias term caused by the simulation, and hence propose a simulated MSEP (SMSEP) as a consistent sample analog to the population predictability criterion.[3] As a result, while those using simulation based prediction for model evaluations in time series framework usually require that the number of simulations tend to infinity, ours works for any fixed number of simulations. Moreover, our model selection test allows for any estimators that are $\sqrt{n}$ asymptotically normally distributed, or are n-consistent that can arise from some structural microeconometric models such as auction models and job search models (Donald and Paarsch, 1993, 1996, 2002; Hong, 1998; Chernozhukov and Hong, 2004; Hirano and Porter, 2003). Lastly, in a similar spirit to that of Vuong (1989) and Rivers and Vuong (2002), the test is bi-directional and applicable to non-nested structural models which are both possibly misspecified. This adds a considerable advantage to the proposed test because in real applications, structural econometric models can be best considered an approximation, but not exact modeling of the true data generating process. Nevertheless, with two possibly misspecified models, our model selection procedure enables one to tell which one is closer to the truth.

While some empirical work has used predictions from structural models to validate a particular choice of the model, because of the lack of a formal test, it has been based on an ad-hoc comparison of the closeness between the predictions and the observed outcomes. The statistical significance of such a closeness is not assessed. In contrast, our testing procedure provides a formal framework in which the statistical significance of the difference in predictability of competing structural models can be assessed. The asymptotic distribution of the test statistic is derived. The proposed test is general regardless of whether the optimization criteria for estimation of competing models are the same as the SMSEP criterion used for model selection. We conduct Monte Carlo experiments to study size and power properties of the test. An empirical application using timber auction data from Oregon is used to illustrate the usefulness and generality of the proposed testing procedure.

It is worth noting that most of the recent work in model selection tests has been based on comparing the Kullback–Leibler Information Criterion (KLIC) between two competing models. See, e.g., Kitamura (2000, 2002), and Chen et al. (2003). Our approach is different, as it is based on the simulated mean squared errors of predictions, a lack-of-fit criterion. This is motivated by the fact that many structural econometric models are estimated by GMM or MSM other than the MLE, thus the KLIC cannot be used as a model selection criterion.[4] Our model selection criterion, on the other hand, can be used for any estimation methods that yield estimators with root-n asymptotic normality, or with n^{α}-consistency for $\alpha > 1/2$, and hence has an appealing generality.

This paper is organized as follows. Section 2 describes the general model selection framework for structural econometric models using the SMSEP criterion. The hypotheses for model selection are formulated. The asymptotic properties of the proposed test statistic are established. Section 3 is devoted to Monte Carlo experiments in investigating finite sample properties of the tests, and Section 4 considers an empirical application of the proposed test to structural auction models. Section 5 concludes.

2. An SMSEP criterion and the resulting model selection test

Two models $\mathcal{M}_1$ and $\mathcal{M}_2$ are estimated using data $\{y_i, \mathbf{x}_i\}$, $i = 1, \ldots, n$, where y is a dependent variable and $\mathbf{x}$ is a $1 \times K$

a problem of determining between 11 and 18 potential bidders. This problem, significant from an economic viewpoint as having 11 bidders could imply the existence of a large trader and hence asymmetric bidding, calls for a formal test of non-nested models, as the structural models with different numbers of potential bidders are non-nested. While this issue was not further pursued in Laffont et al. (1995) (see footnote 21 in Laffont et al. (1995)), and cannot be addressed using the existing model selection methods, it can be resolved using our proposed procedure, as illustrated in the empirical application.

3 The bias correction we use here in constructing the SMSEP adopts the one introduced by Laffont et al. (1995) for simulated nonlinear least squares estimation.

4 On the other hand, if the structural models considered here are estimated using empirical likelihood or other KLIC based methods, then one can apply the recent model selection tests such as Kitamura (2002).