Journal of Econometrics 157 (2010) 120–128

Contents lists available at ScienceDirect

Journal of Econometrics

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

Indirect inference in structural econometric models

Tong Li

Department of Economics, Vanderbilt University, United States

ARTICLE INFO

Article history:
Available online 29 October 2009

JEL classification:
C13
C15

Keywords:
Auxiliary model
OLS
Parameter calibration
Parameter-dependent support
Simulation

ABSTRACT

This paper considers parametric inference in a wide range of structural econometric models. It illustrates how the indirect inference principle can be used in the inference of these models. Specifically, we show that an ordinary least squares (OLS) estimation can be used as an auxiliary model, which leads to a method that is similar in spirit to a two-stage least squares (2SLS) estimator. Monte Carlo studies and an empirical analysis of timber sale auctions held in Oregon illustrate the usefulness and feasibility of our approach.

1. Introduction

This paper considers parametric inference in a wide range of structural econometric models. These models include the job search models studied in Flinn and Heckman (1982), Wolpin (1987) among many others and empirical game-theoretic models such as structural auction models pioneered by Paarsch (1992) and subsequently further developed by Donald and Paarsch (1993, 1996), Laffont et al. (1995), Guerre et al. (2000), Li et al. (2000, 2002), and Hong and Shum (2003), to name a few. A common feature of these models is that while one's primary interest is the parameters that characterize the distribution of some latent variables such as the wage distribution in a job search model or the distribution of bidders' private values in an auction model, one only observes some variables that are related to but not identical to these latent variables. For example, in the job search case, the observed wages are those above the (unobserved) reservation wage, and in the case of auctions, we usually observe bids, which are generally different from private values. Thus, a researcher needs to rely on a structural model that defines a map between these latent and observed variables. This paper is motivated by two major complications arising from the structural inference. First, a consequence of the structural approach is that the support of the distribution of these observed dependent variables usually depends on the parameters of interest, and hence violates the regularity conditions for the consistency and asymptotic normality of the maximum likelihood (ML) estimator.[1] Second, since the relationship between the observed (equilibrium) outcomes and the latent variables defined by a structural model is complicated, it is usually difficult to derive explicit moment conditions and hence to implement Generalized Method of Moments (GMM) or Empirical Likelihood (EL) estimation.

To address the issues arising from the nonregular ML problem, Flinn and Heckman (1982) were the first to suggest employing the ML method subject to the restrictions imposed from economic theory including the condition for the support of the observed wage distribution by using the observed minimum wage as a superconsistent estimate of the (unobserved) reservation wage. Since then, progress has been made in the last two decades. For example, in considering a prototypical job search model with homogeneous samples, Christensen and Kiefer (1991) study the exact likelihood estimation. In empirical auction models, Donald and Paarsch (1993, 1996) first initiate the method of maximum likelihood estimation by considering the estimation as a constrained optimization problem with some constraints becoming binding as the sample size becomes large. Hong (1998) studies asymptotic properties of the nonregular maximum likelihood estimator. In particular, the estimator usually has a convergence rate at n, faster than the usual $\sqrt{n}$ rate, but with a nonstandard asymptotic distribution complicated and dependent

[1] Such a problem in implementing the ML method in structural econometric models is first recognized in Flinn and Heckman (1982), while there is a longer history in the statistical literature studying a somewhat related but different problem that is the estimation of the lower (or upper) boundary parameter of a distribution (see, e.g., Woodroofe, 1972, 1974).

E-mail address: tong.li@vanderbilt.edu.

0304-4076/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2009.10.027

on nuisance parameters. More recently, Donald and Paarsch (2002) propose an estimation method using the extreme order statistics, a method that is simpler to implement in practice. They show that, with discrete covariates, their estimator is also superconsistent with convergence rate n.

Although the aforementioned estimators have in general a convergence rate n, they are difficult to calculate and have limiting distributions specific to whether the covariates are discrete or continuous, and also related to the exponential distribution. This makes the inference more involved. It is thus not surprising that the focus from the recent literature has been mainly on proposing easy-to-implement estimation procedures that are consistent but may not be efficient, because at issue is how to make the estimation of structural econometric models tractable.[2] See, e.g., Laffont et al. (1995), Bajari (1998), Hong and Shum (2003), among others. This paper follows this line of literature in aiming to provide a computationally convenient procedure for estimating the structural models discussed above.[3] Specifically, we show that the indirect inference principle recently developed by Smith (1993), Gourieroux et al. (1993), and Gallant and Tauchen (1996) can be used in proposing a computationally convenient estimator that has the standard root-n asymptotic normality. In particular, the estimation involves two steps. In a first step, an auxiliary model that is easy to estimate is estimated using the real data to get estimates for those auxiliary parameters, which are functions of the structural parameters of interest. In the second step, the structural model is simulated and the simulated data are used to estimate the auxiliary parameters. The estimates for the structural parameters are obtained when some criterion that measures the difference between these estimates for the auxiliary parameters obtained from the actual data at the first step and from the simulated data at the second step is minimized. While the principle of indirect inference is proposed by Smith (1993), Gourieroux et al. (1993), and Gallant and Tauchen (1996) as a general estimation principle in estimating otherwise intractable econometric models, at issue is how to choose an appropriate auxiliary model in a particular problem. Although in principle a pseudo maximum likelihood estimation (PMLE) can be used for the first-stage auxiliary model, where a possibly misspecified likelihood function can be maximized, we offer a particular choice for the auxiliary models for its intuitive appeal and attractive computational convenience. Specifically, we demonstrate that an ordinary least squares (OLS) estimation can be used as an auxiliary model. This is an interesting approach for several reasons. First, using OLS as an auxiliary model significantly enhances the computational efficiency. Second, nonlinear functions of the covariates that the structural model conditions on can be used as additional instruments in the OLS estimation; the resulting indirect inference estimator can be more efficient because the use of the additional instruments makes the structural model of interest over-identified. Third, the auxiliary parameters have the interpretation of reduced-form parameters, and hence our method makes the connection between the modern structural econometric models and the conventional Cowles Commission structural simultaneous equations models.[4] Lastly, noting that many empirical studies have used the reduced-form approach to conduct analysis given the complications of a structural analysis, our approach offers an effective way to bridge the gap between the reduced-form and structural approaches through the indirect inference principle.

Besides the advantage of computational convenience, our approach makes inference in structural models feasible. For example, model specification and hypothesis testing can be made readily within the framework of the indirect inference. In addition, since the encompassing method for selecting between two non-nested models have been extended to the case where an indirect inference estimator is used for either/both alternative models (see, e.g., Dhaene et al., 1998), our approach can be adapted to distinguishing between non-nested structural models such as two different equilibria, or two competing primitive models. Another advantage of our approach is that it can be less sensitive to the outliers than the extreme order statistics-based estimator and the constrained MLE, which both heavily rely on the extreme values of the observations.

We conduct Monte Carlo studies to demonstrate the usefulness and feasibility of our methods. Moreover, we apply our method to analyze timber sale auctions that were held in Oregon and organized by the Oregon Department of Forestry. We use the method proposed in the paper to provide a structural analysis to this new data set by estimating the distribution of private values accounting for the heterogeneity of the auctioned timber lots.

This paper is organized as follows. Section 2 is devoted to proposing a unified approach for using the indirect inference principle in structural econometric models. Section 3 presents Monte Carlo results that demonstrate the usefulness and feasibility of our approach. Section 4 is devoted to the empirical application. Section 5 concludes.

2. Indirect inference in structural econometric models

2.1. The canonical model

We present in this subsection the basic or canonical model that will be studied. In structural modeling, we are interested in estimating the distribution of a latent variable, say, v, with a known support on $[a, b]$. Suppose that the distribution of v conditioning on some covariates is parameterized by a vector of K parameters $\theta \in \Theta$, where Θ is a compact set in $\mathbb{R}^K$. Denote this distribution by $F(\cdot|\mathbf{x}, \theta)$, where $\mathbf{x} \in \Xi \subset \mathbb{R}^q$ is a vector of q variables that are used to control for the observed heterogeneity. The corresponding density is denoted by $f(\cdot|\mathbf{x}, \theta)$. The true value of the structural parameters is denoted by θ^0. While v is partially or fully unobserved, we observe a (dependent) variable y, which, as a result of the structural modeling, has a density also dependent on θ with a support on $[h_1(\mathbf{x}, \theta), h_2(\mathbf{x}, \theta)]$. We denote the distribution of y as $G(\cdot|\mathbf{x}, \theta)$ and the corresponding density as $g(\cdot|\mathbf{x}, \theta)$. Also, assume that n independent observations y_i and $\mathbf{x}_i$, $i = 1, \ldots, n$ are available for both y and $\mathbf{x}$. Note that, if v were observed, then estimation of θ could be done as a straightforward application of the ML method. However, we do not observe v. Therefore a main objective of the structural approach is to estimate the structural parameters θ^0 using the observations on both y and $\mathbf{x}$. We have the following assumption on the latent density $f(\cdot|\mathbf{x}, \theta)$.

A1: For each $\theta \in \Theta$, $f(\cdot|\cdot, \theta)$ is a Borel measurable function on $[a, b] \times \Xi$; and for each $(v, \mathbf{x}) \in [a, b] \times \Xi$, $f(v|\mathbf{x}, \cdot)$ is a continuous function on Θ.

2 In fact, Chernozhukov and Hong (2004) and Hirano and Porter (2003) show that for structural models with parameter-dependent support, the (constrained) ML estimator is generally inefficient, but the Bayes estimator is efficient according to the local asymptotic minmax criterion for standard loss functions such as the squared error loss.

3 Some structural models such as the equilibrium search models (Eckstein and Wolpin, 1990; Hong and Shum, 2001) do not have the parameter-dependent support problem for the MLE. But the implementation of the MLE in these models can also be computationally demanding. The method proposed in this paper can be applied to these models as well.

4 See Heckman (2001) for an insightful overview on the development of structural econometrics.