Bayesian Inference for Econometric Models using Empirical Likelihood Functions: Extended Abstract

Hyungsik Roger Moon and Frank Schorfheide

Introduction

Estimators based on moment conditions of the form $\mathbb{E}[g(X_i, \theta)]$, where θ is a finite-dimensional parameter vector of interest, are a popular tool in applied econometrics. Unlike likelihood-based estimators, moment-based estimators do not require the researcher to specify the probability distribution of the random vector X_i in detail. While the use of inappropriate auxiliary assumptions about the distribution of X_i potentially leads to misspecification bias, reasonable distributional assumptions may improve the precision of the estimator substantially, in particular in small samples.

In the literature on moment-based estimation, information-theoretic estimators such as empirical likelihood (EL) estimators have emerged as an attractive alternative to generalized method of moments (GMM) estimators. For instance, Kitamura (2001) showed that the empirical likelihood ratio test for moment restrictions is asymptotically optimal under the Generalized Neyman-Pearson criterion. Newey and Smith (2001) find that the asymptotic bias of EL estimators does not grow with the number of moment conditions and that bias-corrected EL estimators have higher-order efficiency properties. A detailed discussion of empirical likelihood methods in econometrics and statistics is provided in the monograph by Owen (2001).

In this paper we propose a method to combine the empirical likelihood function with a prior distribution over the parameters θ and the probability measures for X_i. Rather than imposing beliefs about the distributional form of the X_i's dogmatically by specifying a fully parametric likelihood function, we only use these additional restrictions loosely.

We consider the following approach: in addition to the actual data we generate artificial draws from a version of the model in which we make a specific assumption about the distribution of X_i. We apply empirical likelihood-based estimation methods to the combined sample of actual and artificial data. Such mixed estimation has a long tradition in econometrics dating back to Theil and Goldberger (1961). From a Bayesian perspective, the artificial observations induce a prior distribution for the parameters that are estimated.

Some Background

Del Negro and Schorfheide (2002) use the notion of mixed estimation to specify a prior for a vector autoregression (VAR) that is based on a dynamic stochastic general equilibrium (DSGE) model. DSGE models impose strong cross-parameter restrictions on vector autoregressive representations that are, to some extent, misspecified. However, if these restrictions are only assumed to be approximately correct and the VAR estimates are shrunk toward them, one can obtain VAR estimates that lead to better predictive performance than either the unrestricted VAR or the DSGE model alone.

Let θ be the DSGE model parameters and ϕ be the VAR parameters. The VAR representation of the DSGE model is obtained by the mapping $\tilde{\phi} = f(\theta)$. The likelihood function of the data Y only depends on the VAR parameters and the posterior density is given by ($\propto$ denotes proportionality):

$$p(\theta, \phi|Y) \propto p(Y|\phi)p(\phi|\theta)p(\theta), \tag{1}$$

where $p(\phi|\theta)$ and $p(\theta)$ are prior densities. Heuristically, the prior $p(\phi|\theta)$ is constructed by simulating n^* artificial observations from the DSGE model and fitted a VAR to the artificial observations. This prior has the property that it is not restricted to the subset $\Phi^* = \{\phi : \phi = f(\theta), \theta \in \Theta\}$ of the VAR parameter space. However, it concentrates increasing mass in the neighborhood of Φ^* as $n^* \longrightarrow \infty$. Del Negro and Schorfheide (2002) establish the following results. As $n^* \longrightarrow \infty$ the inference becomes equivalent to inference based on the restricted likelihood function $p(Y|f(\theta))$. In large samples the posterior estimate of θ can be interpreted as projection of the estimate of ϕ onto the restricted subspace Φ^*. A Bayesian selection criterion can be used to choose the size n^* of the artificial sample based on the available data.

We will use the idea of mixed estimation based on the parametric completion of the moment-based model to conduct Bayesian inference.

Bayesian Limited Information Analysis

Typically, Bayesian inference methods are applied to models that provide a parametric likelihood function. However, in many applications, there are reasons to be skeptical about the auxiliary assumptions, e.g., the specific distribution of X_i, that are needed to obtain a parametric probability model for the endogenous variables. Unfortunately, there is no widely accepted Bayesian inference procedure (such as Generalized Method of Moments under the frequentist paradigm) for models that are specified in terms of a few moment conditions.

Recently, Kim (2002a, 2002b) proposed Bayesian inference methods based on limited information likelihood functions or posterior distributions. In the latter case, Kim restores the unknown posterior of the parameters of interest from some moment conditions. Within a set of candidate posteriors that satisfy the desired moment conditions he finds the one that is closest to the "true" yet unknown posterior in an information distance. Lazar (2000) on the other hand, suggests to use the empirical likelihood function directly to conduct Bayesian inference. Our approach follows this second route.

Prior Distributions.

The empirical likelihood

$$L_{EL}(\theta, p_1, \ldots, p_n) = \left\{ \prod_{i=1}^n p_i \middle| p_i > 0,\ \sum_{i=1}^n p_i = 1,\ \sum_{i=1}^n p_i g(X_i, \theta) = 0 \right\} \quad (2)$$

is a function of the parameter vector θ and the multinomial probabilities $p_1, \ldots, p_n$. The parameter of interest is θ, whereas the probability masses p_i are nuisance parameters in many applications. While Lazar (2000) focuses on the approach that concentrates the empirical likelihood function with respect to the p_i's and combines the profile likelihood function with a prior for θ, we plan to carefully construct a prior for the p_i's as well. In our moment-based framework, it is natural to factorize the prior as follows

$$p(\theta, p_1, \ldots, p_n) = p(\theta) p(p_1, \ldots, p_n | \theta). \quad (3)$$

A desirable property of $p(\theta)p(p_1, \ldots, p_n|\theta)$ is that it concentrates most of its mass on values of p_i for which the moment condition $\mathbb{E}[g(X_i, \theta)] = 0$ is at least approximately satisfied.

A common approach in non-parametric Bayesian analysis is to use a Dirichlet distribution as a prior for the p_i's (see, for instance, Ferguson (1973, 1974) and Rubin (1981)). We will use the following heuristic to obtain a prior $p(p_1, \ldots, p_n|\theta)$. Starting from an uninformative prior distribution, we generate a "posterior" for the probability masses based on n^* artificial observations from the parametric completion of the moment-based econometric model. Similar to the approach taken in Del Negro and Schorfheide (2002) the "posterior" distribution obtained from the simulated observations is used as a "prior" for the analysis of the actual data. This approach has the advantage of not dogmatically imposing a parametric form for the distribution of the endogenous variables, yet at the same time supplementing the sample information by model-consistent beliefs about likely values of the probability masses. The first step of our analysis will be to formalize the heuristic description of the prior distribution.

Posterior Analysis.

The proposed prior distribution is combined with the empirical likelihood function (2) to obtain a posterior distribution. We plan to address the following issues: (i) consistency of the Bayes estimate of θ. (ii) We will derive a large-sample approximation for the posterior distribution of θ and compare our results to other limit-information approaches, such as Kim (2002a, 2002b) and Lazar (2000). (iii) Develop a Markov-Chain-Monte-Carlo algorithm to generate draws from the posterior distribution of the p_i's and θ. (iv) Assess to what extent the parametric completion of the moment-based model is misspecified.

References.

Del Negro, Marco and Frank Schorfheide (2002): "Priors from Equilibrium Models for VARs," PIER Working Paper 02-024, University of Pennsylvania.

Ferguson, T. (1973): "A Bayesian Analysis of Some Nonparametric Problems," *The Annals of Statistics*, **1**, 209-230.

Ferguson, T. (1974): "Prior Distributions on Spaces of Probability Measures," *The Annals of Statistics*, **2**, 615-629.

Kim, Jae-Young (2002a): "Limited Information Likelihood and Bayesian Analysis," *Journal of Econometrics*, forthcoming.

Kim, Jae-Young (2002b): "Bayesian Limited Information Analysis in the GMM Framework," *Manuscript*, SUNY-Albany, Department of Economics.

Kitamura, Yuichi (2001): "Asymptotic Optimality of Empirical Likelihood for Testing Moment Restrictions," *Econometrica*, **69(6)**, 1661-1672.

Lazar, N.A. (2000): "Bayesian Empirical Likelihood," *Technical Report*, Carnegie Mellon University, Department of Statistics.

Newey, Whitney K. and Richard J. Smith (2001): "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Manuscript*, MIT, Department of Economics.

Owen, Art B. (2001): *"Empirical Likelihood,"* Chapman & Hall, New York.

Rubin, D. (1981): "The Bayesian Bootstrap," *The Annals of Statistics*, **9**, 130-134.

Theil, Henry and Arthur S. Goldberger (1961): "On Pure and Mixed Estimation in Economics," *International Economic Review*, **2**, 65-78.