Journal of Econometrics 178 (2014) 132–145

Contents lists available at ScienceDirect

Journal of Econometrics

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

An alternative quasi likelihood approach, Bayesian analysis and data-based inference for model specification

Jae-Young Kim*

Department of Economics, Seoul National University, Seoul, 151-746, Republic of Korea

ARTICLE INFO

Article history:
Available online 12 August 2013

JEL classification:
C11
C14
C2
C3
C5

Keywords:
Quasi likelihood
Model misspecification
Limited information
Bayesian methods
Sandwich covariance
Large sample correspondence
Model selection

ABSTRACT

This paper studies an alternative quasi likelihood approach under possible model misspecification. We derive a filtered likelihood from a given quasi likelihood (QL), called *a limited information quasi likelihood (LI-QL)*, that contains relevant but limited information on the data generation process. Our LI-QL approach, in one hand, extends robustness of the QL approach to inference problems for which the existing approach does not apply. Our study in this paper, on the other hand, builds a bridge between the classical and Bayesian approaches for statistical inference under possible model misspecification. We can establish a large sample correspondence between the classical QL approach and our LI-QL based Bayesian approach. An interesting finding is that the asymptotic distribution of an LI-QL based posterior and that of the corresponding quasi maximum likelihood estimator share the same "sandwich"-type second moment. Based on the LI-QL we can develop inference methods that are useful for practical applications under possible model misspecification. In particular, we can develop the Bayesian counterparts of classical QL methods that carry all the nice features of the latter studied in White (1982). In addition, we can develop a Bayesian method for analyzing model specification based on an LI-QL.

1. Introduction

In many cases of practical work in economics and social sciences models are approximates of the true data generation process (DGP). Such an approximate model is often misspecified or only partially correct. A useful approach in this case is the quasi maximum likelihood (QML) method studied by Huber (1967) and White (1982), among others, that generalizes the traditional maximum likelihood (ML) method to the case of possible model misspecification. In this paper, we study an alternative quasi likelihood approach to statistical inference under possible model misspecification. Our study extends robustness of the existing quasi likelihood (QL) approach to more general situations. Also, our study in this paper develops a Bayesian counterpart of the classical QL approach, which constructs a bridge between the classical and Bayesian approaches for statistical inference under possible model misspecification.

The ML estimation under model misspecification and sampling properties of the resulting estimator have been studied by Godambe (1960), Huber (1967), Wedderburn (1974), Desmond (1997) and Heyde (1997). White (1982, 1994) has built an architecture for classical inference in the QL environment, extending inference procedures in the ML method to the case of possible model misspecification. In a different angle, the QL method is seen as an approach of weakening or minimizing assumptions of a model and its probabilistic nature. Thus, the QL approach is a nice option for practical applications when the full information on DGP is not available.

One can find, however, limitations of the QL approach based on a given "naive" QL. First, in the presence of possible model misspecification some inference methods based on a given naive QL do not work. For example, the likelihood ratio (LR) test based on a naive QL does not apply for statistical inference in many cases of possible model misspecification. Second, the standard Bayesian approach has fundamental problems in the naive QL environment: (1) it does not have an inferential apparatus that takes care of possible misspecification. As a result, there is no difference in Bayesian inference methods between the case of model misspecification and that of correct specification. See Blackwell (1985), Chen (1985), Bunke and Milhaud (1998), Gelman et al. (2004), Geweke (2005), Kleijn and van der Vaart (2008), and Müller (forthcoming); (2) the classical QML method does not have a Bayesian counterpart in the naive QL environment, which implies that the Bayesian approach in the naive QL environment does not have proper interpretations in the repeated sampling context. This is a sharp contrast to the

---

* Tel.: +82 2880 6390.
E-mail address: jykim017@snu.ac.kr.

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.08.012

standard case of correct specification when the Bayesian and classical approaches have large sample correspondence, as is studied in Andrews (1994).

In this paper, we study an approach that provides solutions to the above explained inference problems involved in the naive QL approach. We first derive a likelihood from a set of statistics implied in a given naive QL but at the same time reflects some key aspects of the true DGP. The derived likelihood is a "filtered" version of a given QL that contains relevant, though limited, information on DGP. We call such a derived likelihood a limited information quasi likelihood (LI-QL). The limited information on DGP used for deriving such a filtered QL consists of a set of statistics implied in a given QL. Our approach of using some limited information on DGP contained in a QL follows the note of Wedderburn (1974), Godambe (1960), Heyde (1997) and Desmond (1997). Given a set of limited information, an LI-QL is derived based on the information theory (Csiszar, 1975, Jaynes, 1982, Jones, 1989, among others).

Based on the LI-QL we study the inference methods and provide solutions to the problems in the naive QL environment explained above. First, different from the naive classical QL approach under possible model misspecification the LI-QL allows us to derive the LR, LM and the Wald tests for testing the parametric hypotheses. It follows that these three tests based on the LI-QL are asymptotically equivalent under possible model misspecification. Second, we develop a Bayesian framework based on an LI-QL that carries all the nice features of classical QL methods studied in White (1982). As such, this Bayesian approach of ours solves problems of the standard Bayesian approach in a naive QL environment explained above or in Remark 2, Section 2. In particular, we have large sample correspondence between our alternative Bayesian formation and the classical approach under possible model misspecification.

An important feature of the posterior formed from our LI-QL is that its asymptotic second moment is of a sandwich type. This "sandwich posterior" constructs a bridge between the classical and Bayesian approaches for statistical inference under possible model misspecification. Müller (forthcoming) analyzes the risk of Bayesian inference in misspecified models.[1] He shows that, under misspecification, inference based on an "artificial" sandwich posterior results in decisions of lower frequentist risk than inference based on the corresponding original (non-sandwich) posterior. Our analysis in this paper provides a formal approach to the derivation of such a sandwich posterior.

In this paper, we also develop data-based methods for analyzing model specification based on the LI-QL. We derive a model selection rule by applying the Bayesian decision rule to the LI-QL. Our approach based on the LI-QL explicitly incorporates the possibility of misspecification for models under consideration, which is an important feature of our approach. Consistency of the model selection method is proved. We also derive a decision rule for determining whether or not a given model is correctly specified. This method is a Bayesian version of the information matrix equality test for determining validity of a given specification.

The LI-QL based Bayesian procedure studied in this paper is a kind of semi-parametric, limited information, Bayesian procedure whose literature is relatively small. There are earlier pieces of work by Zellner (1996, 1997, 1998) who developed a Bayesian limited information procedure based on the principle of maximum entropy. Boos and Monahan (1986) considered the limited information Bayesian analysis based on the translation equivalent estimator, and Doksum and Lo (1990) studied the limited information Bayesian analysis based on a robust estimator. Kim (2002) discusses a limited information approach in the generalized method of moments (GMM) framework but studies different problems from ours in this paper. Chernozhukov and Hong (2003) defined Laplace type quasi-posteriors as transformations of general (non-likelihood based) statistical criterion functions such as those in GMM and nonlinear IV methods. Chernozhukov and Hong (2003)'s work is important in that it enables practitioners to use Markov Chain Monte Carlo methods for many classical inference problems. There remains, however, an important issue of how to deal with the model misspecification in a GMM based approach. See Section 3.2, Remark 3. Our study in this paper addresses the issue based on an information theoretic approach.

The discussion of the paper goes as follows. Section 2 explains the quasi likelihood approach and its limitations for statistical inference. In Section 3 we discuss how to get a limited information quasi likelihood. Properties of the LI-QL are also explored in Section 3. Section 4 studies a Bayesian framework based on the LI-QL. A posterior based on the LI-QL is derived, and its properties are studied. We also discuss numerical methods for evaluating the LI-QL posterior. In Section 5 we discuss inference methods based on the LI-QL and its posterior. In Section 6 a model selection procedure based on the LI-QL is studied. Section 7 provides Monte Carlo results, and Section 8 concludes the paper.

2. Quasi likelihood and inference problems

Under limited information on DGP a natural approach is to set up a quasi likelihood that is supposed to be close, if not equal, to the true likelihood. There is a well documented literature on the QL approach in the classical framework. See White (1982, 1994), for example. The QL approach is a nice option for practical analysis under limited information on DGP. However, the existing QL method has limitations for econometric practice, which motivates our work in this paper. We lay out subjects of our study in this section by explaining the existing QL method and by highlighting its limitations.

2.1. Model misspecification and quasi likelihood

Let x_t be a vector of stationary and ergodic processes defined on a probability space $(\Omega, \mathcal{F}, P)$. Denote by $\mathbf{x}_n(\bar{\omega}) = (x_1(\bar{\omega}), \ldots, x_n(\bar{\omega}))$, for $\bar{\omega} \in \Omega$, an n-segment of a particular realization of $\{x_t : t \geq 1\}$. Denote by G and g, respectively, the true distribution function of x_t and its density, which are unknown a priori. For an integrable function H of $x(\omega)$ we denote the expectation with respect to the probability measure P by $E_P[H(x)] = \int_\Omega H(x(\omega)) P(d\omega) = \int_{\mathbb{R}^k} H(x) dG(x)$.

Since G and g are unknown, a researcher would consider a family of distribution functions F or probability densities f, which may or may not contain the true structure. Assume that F and f are of parametric families characterized by a $k \times 1$ vector of parameters $\theta \in \Theta$, where Θ is a convex subset of $\mathbb{R}^k$. Let $\mathcal{G}$ be the Borel σ-algebra of Θ, so that $(\Theta, \mathcal{G})$ is a measurable space.

Let $f_n(\mathbf{x}_n, \theta)$ be a quasi likelihood based on a postulated pdf of the random variable x such that

$$\ln f_n(\mathbf{x}_n, \theta) = \sum_{t=1}^{n} \ln f(x_t, \theta).$$

Note that, we consider a rather simple form of QL with an additivity structure of the log-QL for the convenience of analysis but without much loss of generality. In many cases of econometric practice x_t and thus $f(x_t, \theta)$ may be dependent and heteroscedastic processes. Dependency and heteroscedasticity in $f(x_t, \theta)$ or a function of it can be taken care of by a long-run variance. See Assumption

---

[1] After this paper was submitted for publication, a referee brought the author the attention of an independent work by Müller (forthcoming). The author appreciates the referee for it.