Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 138 (2007) 461–487
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

On the efficient use of the informational content of estimating equations: Implied probabilities and Euclidean empirical likelihood ☆

Bertille Antoine[a,*], Hélène Bonnal[b], Eric Renault[c]

[a]*Université de Montréal and CIREQ, Canada*
[b]*GREMAQ, Université de Toulouse 1, France*
[c]*University of North Carolina at Chapel Hill, CIRANO and CIREQ, USA*

Available online 14 June 2006

Abstract

A number of information-theoretic alternatives to GMM have recently been proposed in the literature. For practical use and general interpretation, the main drawback of these alternatives, particularly in the case of conditional moment restrictions, is that they give up the computational and interpretational simplicity of quadratic optimization. The main contribution of this paper is to analyze the informational content of estimating equations within the unified framework of Chi-square distance. Improved inference by control variables, closed form formulae for implied probabilities and information-theoretic interpretations of continuously updated GMM are discussed in the two cases of unconditional and conditional moment restrictions.

JEL classification: C13; C14; C31

Keywords: Information-based inference; Generalized method of moments; Continuous updating; Empirical likelihood; Control variates

☆A previous version of this paper has circulated under the title "Minimum Chi-Square Estimation with Conditional Moment Restrictions" W.P. 2001.

*Corresponding author.
E-mail addresses: bertille.antoine@umontreal.ca (B. Antoine), hbonnal@gremaq.univ-tlse1.fr (H. Bonnal), renault@email.unc.edu (E. Renault).

0304-4076/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2006.05.005

1. Introduction

It has long been appreciated that in some circumstances likelihood functions may not be available and the focus of parametric inference is only on a limited number of structural parameters associated to the data generating process (DGP) by a structural econometric model. Hansen (1982) has fully settled the theory to use efficiently the informational content of such moment conditions about unknown structural parameters while Chamberlain (1987) showed that the semiparametric efficiency bound for conditional moment restriction models is attained by optimal GMM.

However, and somewhat surprisingly, the pre-1990 GMM literature seems to have forgotten that moment restrictions, when they overidentify the structural parameters of interest, may bring useful information about other characteristics of the DGP. To see this, let us consider n i.i.d. observations $(X_i, Z_i), i = 1, \ldots, n$ of a random vector (X, Z) on $\mathbb{R}^l \times \mathbb{R}^d$. In this paper we focus on the informational content of either q unconditional moment restrictions:

$$E[\Psi(X, \theta^0)] = 0 \tag{1.1}$$

or q conditional moment restrictions

$$E[\Psi(X, \theta^0)|Z] = 0 \tag{1.2}$$

which, in both cases, are assumed to define the true unknown value θ^0 of a vector $\theta \in \Theta \subset \mathbb{R}^p$ of p unknown parameters, while $\Psi : \mathbb{R}^l \times \Theta \to \mathbb{R}^q$ is a known function. When $q \geqslant p$ in case (1.1) or irrespective of the value of q in case (1.2), only one part of the informational content of these moment restrictions is actually used by traditional GMM approaches to estimate θ efficiently. The usefulness of residual information due to overidentification is overlooked.

Actually, following Hansen (1982), efficient estimation of θ^0 from (1.1) goes through a preliminary consistent estimation of the matrix $M(\theta^0)$ of optimal selection of estimating equations:

$$M(\theta^0) = E\left[\frac{\partial \Psi'}{\partial \theta}(X, \theta^0)\right] Var^{-1}[\Psi(X, \theta^0)]. \tag{1.3}$$

As surveyed by Newey (1993), efficient estimation of θ^0 from (1.2) rests upon a preliminary consistent estimation of the matrix $M(Z, \theta^0)$ of optimal instruments:

$$M(Z, \theta^0) = E\left[\frac{\partial \Psi'}{\partial \theta}(X, \theta^0)|Z\right] Var^{-1}[\Psi(X, \theta^0)|Z]. \tag{1.4}$$

The important idea that such over-identified moment restrictions should also lead us to revise our empirical views about the DGP has first been put forward by the empirical likelihood (EL) literature (Owen, 1990, 1991; Qin and Lawless, 1994) for a classical approach, and by Zellner's Bayesian method of moments (BMOM) for a Bayesian one (Zellner, 1991; Zellner and Tobias, 2001).

Typically, as clearly explained in Zellner (2003), the idea is to seek the least informative density function in terms of expected distance subject to the moment conditions. But, while Zellner considers expected distances with respect to priors, we are going to consider distances with respect to empirical probability distributions, that put weights $1/n$ on the n