Statistics and Probability Letters 118 (2016) 117–123

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Improving confidence set estimation when parameters are weakly identified

Heather Battey[a,b,*], Qiang Feng[c], Richard J. Smith[d]

[a] *Imperial College London, United Kingdom*
[b] *Princeton University, United States*
[c] *UIBE, Beijing, China*
[d] *University of Cambridge, United Kingdom*

ARTICLE INFO

Article history:
Received 24 November 2015
Accepted 20 June 2016
Available online 29 June 2016

Keywords:
Weak instruments
Moment inequality conditions

ABSTRACT

We consider inference in weakly identified moment condition models when additional partially identifying moment inequality constraints are available. We detail the limiting distribution of the estimation criterion function and consequently propose a confidence set estimator for the true parameter.

1. Introduction

This paper considers the estimation of a d_θ-vector of parameters θ_0 which is the solution to the set of moment equality restrictions

$$\mathbb{E}\left[g(Z,\theta)\right]=0 \quad \text{at } \theta=\theta_0 \tag{1.1}$$

where $g(z,\theta)$ is a d_g-vector of known functions of the observation vector z and $\theta \in \Theta$ with Θ the parameter space. Estimators based on estimating equations of the form (1.1) are referred to as Z-estimators (e.g. van der Vaart, 1998) and have found application in numerous fields, e.g., survival modelling with incomplete covariate data (Lipsitz and Ibrahim, 1998) and causal inference with instrumental variables (Shpitser, 2014; Imbens, 2014).

A challenging problem arises when the identifying strength of the moment conditions (1.1) for θ_0 is weak, e.g., when instrumental variables used to construct the moment indicator $g(Z,\theta)$ are only weakly correlated with endogenous covariates. Existing inferential procedures robust to weak identification, see *inter alia* Stock and Wright (2000), Kleibergen (2005), Guggenberger and Smith (2005) and Andrews and Cheng (2012), share the shortcoming that confidence set estimators for θ_0 are frequently too large to be of practical use. In many applications where weak identification is a problem, however, moment inequality conditions of the form

$$\mathbb{E}\left[m(Z,\theta)\right]\geq 0 \quad \text{at } \theta=\theta_0 \tag{1.2}$$

are often available, where $m(z,\theta)$ is a d_m-vector of functions known up to θ. This is especially so when instruments are used to overcome estimator bias induced by confounding variables, that is, when latent variables causally affect both response and covariates.

* Corresponding author at: Imperial College London, United Kingdom.
E-mail addresses: h.battey@imperial.ac.uk (H. Battey), q.feng@uibe.edu.cn (Q. Feng), rjs27@econ.cam.ac.uk (R.J. Smith).

http://dx.doi.org/10.1016/j.spl.2016.06.015

Consider the effect of smoking on health. It has been postulated (Leigh and Schembri, 2004) that smoking is related to health through the unobservable confounding variable, risk aversion. Thus cigarette price, being weakly correlated with cigarette consumption and uncorrelated with risk aversion, is a possible but weak instrument. Additional moment inequality information is available here since cigarette consumption and unobserved risk aversion are known to be negatively correlated. A similar scenario arises in the returns to education example of Angrist and Krueger (1991), where quarter of birth is proposed as a (weak) instrument for years of schooling, and schooling and unobserved ability are known to be positively correlated, giving rise to an additional moment inequality condition.

From a technical point of view, progress is still possible in this weak instrument setting provided the strength of the correlation between the instrument and the endogenous regressor is not smaller than $\mu/\sqrt{n}$ for some $\mu \neq 0$ where n is the sample size. For this reason, data are viewed as realisations of the triangular array $\{Z_{in}, \ (i = 1, \ldots, n), \ (n = 1, 2, \ldots)\}$, and any row of the triangular array is endowed with the corresponding expectation operator $\mathbb{E}_n$, cf. Example 1.

Although moment inequalities taken in isolation typically only have partial or set identifying power, taken together both forms of information can result in a smaller confidence set estimator for θ_0 than that based solely on the moment equality constraints. See Manski (2003) and more recently Chernozhukov et al. (2007) and Rosen (2008) for discussions of partial identification. The concern, therefore, of this paper is the construction of a confidence set estimator for θ_0 in weakly identified models defined by (1.1) in the presence of additional partially identifying moment inequality (1.2) constraints.

To illustrate the similarities and differences between this paper and the existing literature consider the following example.

Example 1.

$$Y_i = \theta_0 X_i + \varepsilon_{1i}, \quad X_i = \gamma_{0,n} W_i + \vartheta_{0,n}\varepsilon_{1i} + \varepsilon_{2i}, \ i = 1, \ldots, n,$$

where ε_{1i}, ε_{2i} and W_i are mutually uncorrelated. The parameter θ_0 is weakly identified if $\gamma_{0,n} = \mathbb{E}_n[X_i W_i]/\mathbb{E}_n[W_i^2] = \mu/n^{1/2}$ for $\mu \neq 0$, and $\vartheta_{0,n} = \mathbb{E}_n[X_i\varepsilon_{1i}]/\mathbb{E}_n[\varepsilon_{1i}^2] = \vartheta_0 \neq 0$, and partially identified if $\gamma_{0,n} = 0$ and $\vartheta_{0,n} = \vartheta_0 \geq 0$. Andrews and Soares (2010) consider both non-weak moment equalities and moment inequalities, i.e., $\gamma_{0,n} = \gamma_0 \neq 0$ and $\vartheta_{0,n} = \vartheta_0 \geq 0$, while Moon and Schorfheide (2009) consider $\gamma_{0,n} = \gamma_0 \neq 0$ and $\vartheta_{0,n} = c/n^{1/2} \geq 0$, where c is a constant. This paper addresses the case $\gamma_{0,n} = \mu/n^{1/2}$ and $\vartheta_{0,n} = \vartheta_0 \geq 0$.

To aid clarity, the paper focuses on the special case in which no nuisance parameters are present. For recent contributions that discuss inference in partially identified models with nuisance parameters, see Bugni et al. (2015) and Canay and Azeem (2016).

The rest of the paper is organised as follows. Section 2 defines the confidence set estimator for θ_0 and establishes its properties. Section 3 discusses its implementation with Section 4 providing an examination of the finite sample performance of the confidence set estimator.

2. Inferential procedure

Given the sample of observations $\{Z_{in}, \ (i = 1, \ldots, n)\}$ the interest of the paper is a nominal α-level confidence set estimator $\{\widehat{C}_n(\alpha)\}$ for θ_0 based on the continuous updating (CUE) generalised method of moments (GMM) estimation criterion (Hansen et al., 1996); cf. Rosen (2008).

Let $\widehat{g}^n(\theta) = n^{-1}\sum_{i=1}^n g_{in}(\theta)$ and $\widehat{m}^n(\theta) = n^{-1}\sum_{i=1}^n m_{in}(\theta)$ where $g_{in}(\theta) = g(Z_{in}, \theta)$ and $m_{in}(\theta) = m(Z_{in}, \theta)$, $(i = 1, \ldots, n)$. The CUE GMM criterion is defined as

$$\widehat{Q}_n(\theta, t) = \begin{pmatrix} \widehat{g}^n(\theta) \\ \widehat{m}^n(\theta) - t \end{pmatrix}' \widehat{V}^n(\theta)^{-1} \begin{pmatrix} \widehat{g}^n(\theta) \\ \widehat{m}^n(\theta) - t \end{pmatrix},$$

where

$$\widehat{V}^n(\theta) = n^{-1}\sum_{i=1}^n \begin{pmatrix} g_{in}(\theta) - \widehat{g}^n(\theta) \\ m_{in}(\theta) - \widehat{m}^n(\theta) \end{pmatrix} \left((g_{in}(\theta) - \widehat{g}^n(\theta))', \ (m_{in}(\theta) - \widehat{m}^n(\theta))' \right),$$

and $t \in \mathbb{R}_+^{d_m}$ is a d_m-vector of slackness parameters reflecting the inequality moment constraints (1.2). Minimisation with respect to t yields the profile CUE GMM criterion,

$$\widehat{Q}_n(\theta) = \widehat{Q}_n(\theta, \widehat{t}_n(\theta)) \quad \text{where } \widehat{t}_n(\theta) = \operatorname*{arginf}_{t \in \mathbb{R}_+^{d_m}} \widehat{Q}_n(\theta, t). \tag{2.1}$$

The α-level confidence set estimator $\{\widehat{C}_n(\alpha)\}$ based on (2.1) is then defined as

$$\widehat{C}_n(\alpha) = \left\{\theta \in \Theta : n\widehat{Q}_n(\theta) \leq q\right\}, \tag{2.2}$$

where q is a critical value chosen to ensure that $\lim_{n\to\infty} \mathbb{P}\mathrm{r}_n(\theta_0 \in \widehat{C}_n(\alpha)) \geq 1-\alpha$ and $\mathbb{P}\mathrm{r}_n(\cdot)$ is probability taken with respect to the joint distribution of $\{Z_{in}\}_{i=1}^n$.