Available online at www.sciencedirect.com
ScienceDirect
Economics Letters 99 (2008) 375–378
economics letters
www.elsevier.com/locate/econbase

Specification testing under moment inequalities

Patrik Guggenberger [a,*], Jinyong Hahn [a], Kyooil Kim [b]

[a] *Department of Economics, UCLA, USA*
[b] *Department of Economics, University of Minnesota, USA*

Received 3 April 2006; received in revised form 7 August 2007; accepted 4 September 2007
Available online 11 September 2007

Abstract

We provide a specification test for moment inequalities based on a dual characterization of the moment inequalities. For linear moment inequalities, the test is the asymptotic version of the multi-dimensional linear one-sided tests. For nonlinear moment inequalities, the implementation of the test is not practical because the dual characterization takes the form of a multi-dimensional nonlinear one-sided hypothesis.

Keywords: Moment inequalities; Partial identification; Specification test

JEL classification: C12

1. Introduction

There has been a recent surge of interest in statistical inference in situations where parameters of interest are only partially identified. See Manski (2003) for an overview of this literature. In some applications, the parameter is real-valued and the identified set is an interval whose lower and upper bounds may be estimated from the sample. A confidence interval (CI) of the identified set may be constructed by taking account of the sampling variation of these estimates. The CI may be constructed to cover the entire identified or the true value of the parameter with a certain fixed probability. Imbens and Manski (2004), who proposed the latter, observed that the two CIs can be quite different. The difference of the widths of the two CIs can be related to the difference in critical values of one-sided and two-sided tests.

The purpose of this note is to extend Imbens and Manski's (2004) insight to a situation where the parameter of interest is multi-dimensional and can be characterized by moment inequalities.[1] We propose a specification test to test whether such moment inequalities can hold by providing a dual characterization of the moment inequalities. For a model characterized by linear moment inequalities, we find that such a test is the asymptotic version of the multi-dimensional linear one-sided tests as discussed by, e.g., Gourieroux et al. (1982). On the other hand, when the model is given by nonlinear moment inequalities, the test will be subject to practical problems of implementation because the dual characterization takes the form of multi-dimensional nonlinear one-sided hypothesis. Wolak (1991) noted that the main difficulties of the nonlinear one-sided hypothesis tests are (i) "the lack of an empirically implementable procedure for computing an asymptotically exact size critical value", (ii) "the absence of tight upper and lower bounds on the asymptotic distribution of the test statistics", and (iii) "the least favorable null asymptotic distribution may not occur at the unique parameter value satisfying all of the inequality constraints

* Corresponding author. Bunche Hall 8385, Department of Economics UCLA, Box 951477, Los Angeles, CA 90095-1477, USA. Tel.: +1 310 825 0849; fax: +1 310 825 9528.
E-mail address: guggenbe@econ.ucla.edu (P. Guggenberger).

0165-1765/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2007.09.002

[1] Models characterized by moment inequalities have been recently considered by Moon and Schorfheide (2004) and Pakes et al. (2005). Other related papers include Manski and Tamer (2002), Andrews et al. (2004), Andrews and Guggenberger (2005), Beresteanu and Molinari (2005), Rosen (2005), Romano and Shaikh (2006), and Chernozhukov et al. (2007).

with equality". Our dual characterization of the models with nonlinear moment inequalities suggests that these problems in the nonlinear one-sided hypothesis tests will be carried over to our specification test. We suspect that specification tests of nonlinear models with inequalities will remain elusive unless there is progress in testing of nonlinear one-sided hypothesis.

2. Dual characterization of moment inequalities

We first consider a linear model given by a restriction of the form $C\theta\leq\mu$ for $C\in R^{m\times d}$, $\mu\in R^m$, and $\theta\in R^d$.[2] We are interested in the specification test

$$H_0:\exists\theta\quad\text{such that }C\theta\leq\mu. \tag{1}$$

Example. Consider the linear regression model

$$y_i=x_i'\theta+\varepsilon_i\text{ and }E[x_i\varepsilon_i]=0 \tag{2}$$

for $\theta\in R^d$. Assume the range space of y_i is partitioned into a certain number of disjoint intervals. For example, income of an individual is often times reported as an interval rather than a specific value. We do not actually observe y_i but instead only observe the lower and upper interval bounds, denoted by $y_{i\mathrm{L}}$ and $y_{i\mathrm{U}}$ respectively, of the interval that y_i is part of, $y_i\in[y_{i\mathrm{L}}, y_{i\mathrm{U}})$. Assume x_i has bounded support. Then w.l.o.g. we can assume that $x_i\geq 0$. It follows that

$$E[x_iy_{i\mathrm{L}}]\leq E[x_iy_i]\leq E[x_iy_{i\mathrm{U}}].$$

Therefore, writing $A=E[x_ix_i']$, $\mu^{\mathrm{L}}=E[x_iy_{i\mathrm{L}}]$, and $\mu^{\mathrm{U}}=E[x_iy_{i\mathrm{U}}]$, we obtain the restriction, $\mu^{\mathrm{L}}\leq A\theta\leq\mu^{\mathrm{U}}$ or $C\theta\leq\mu$ for $C=(A',-A')'$ and $\mu=(\mu^{\mathrm{U}\prime},-\mu^{\mathrm{L}\prime})'$. The specification test in Eq. (1) tests whether there exist a $\theta\in R^d$ such that Eq. (2) holds.

We show that the null hypothesis (1) can be given a dual characterization of the form

$$H_0:B\mu\geq 0. \tag{3}$$

More precisely, we show that there exists a matrix $B=B(C)$ such that there is a θ satisfying $C\theta\leq\mu$ if and only if $B\mu\geq 0$. See the Appendix for an algorithm that generates such a matrix B. Note that the hypothesis (3) is the multi-dimensional one-sided test discussed, e.g., by Gourieroux et al. (1982) and Wolak (1991). If C is known, then B is known, and the test can be based on the Wald-type test statistic of the form

$$W_n=\inf_{t\in R^p}\{n(B\widehat{\mu}-t)'\widehat{\mathcal{J}}^{-1}(B\widehat{\mu}-t)\quad\text{subject to }t\geq 0\}, \tag{4}$$

where $\widehat{\mu}$ is a $\sqrt{n}$-consistent asymptotically normal estimator of μ and $\widehat{\mathcal{J}}$ is a consistent estimator for the asymptotic variance matrix of $B\widehat{\mu}$.[3] The asymptotic distribution of W_n is a mixture of χ^2-distributions, see Kudo (1963).

[2] If x and y are both m-vectors, $x\leq y$ means $x_i\leq y_i$ for $i=1,\ldots, m$.

[3] If B is unknown but a $\sqrt{n}$-consistent asymptotically normal estimator $\hat{B}$ of B is available, then $\widehat{\mathcal{J}}$ in W_n need to be replaced by a consistent estimator of the asymptotic covariance matrix of $\widehat{B}\widehat{\mu}$.

Assume now that the model is given by $E[\varphi(w;\theta)]\geq\mu$ where φ is a nonlinear function with values in R^m. Letting $\Phi(\theta)\equiv(\Phi_1(\theta),\ldots,\Phi_m(\theta))'\equiv E[\varphi(w;\theta)]$, we can write the null hypothesis as

$$H_0:\exists\theta\quad\text{such that }\Phi(\theta)\geq\mu. \tag{5}$$

Assume $\theta\in\Theta\subset R^d$ for some set Θ. To derive the dual version of Eq. (5) define functions $\overline{\Phi}_j$, $j=1,\ldots, m$:

1. Let $\overline{\Phi}_1\equiv\max_{\theta\in\Theta}\Phi_1(\theta)$.
2. For $j\geq 2$, let $\overline{\Phi}_j\equiv\overline{\Phi}_j(\mu_1,\ldots,\mu_{j-1})\equiv\max_{\theta\in\Theta_j}\Phi_j(\theta)$ if $\Theta_j\neq\emptyset$, where $\Theta_j\equiv\{\theta\in\Theta:\Phi_{j'}(\theta)\geq\mu_{j'}$ for all $j'<j\}$.[4]

Let $\overline{\Phi}\equiv(\overline{\Phi}_1,\ldots,\overline{\Phi}_m)'$. Clearly, there exists a θ such that $\Phi(\theta)\geq\mu$ if and only if $\Psi(\mu)\equiv\overline{\Phi}(\mu)-\mu\geq 0$ and thus Eq. (5) holds if and only if

$$\Psi(\mu)\geq 0. \tag{6}$$

If we have a $\sqrt{n}$-consistent, asymptotically normal estimator $\hat{\mu}$ of μ, testing Eq. (5) is equivalent to testing the one-sided hypothesis on μ. The latter has been studied in Wolak (1991) using a Wald-type statistic as in Eq. (4). In order to determine the asymptotic critical value, one has to maximize the probability of rejection over all μ vectors that satisfy the restriction (6). However, as discussed in Wolak (1991), such a maximization and therefore determination of a critical value typically is computationally intractable.[5]

3. Discussion: CI for a scalar component of $\boldsymbol{\theta}$

The discussion in the preceding section has natural implications for the construction of a CI for a scalar component of a vector-valued parameter. Suppose for simplicity that the model is given by a set of linear restrictions $C\theta\leq\mu$. We are interested in testing whether the first component θ_1 of θ is equal to θ_1^*. We can write the null hypothesis as

$$H_0:\exists\theta_{-1}\quad\text{such that }C_{-1}\theta_{-1}\leq\mu-c_1\theta_1^*,$$

where c_1 is the first column of C, C_{-1} is a submatrix consisting of the remaining columns, and $\theta_{-1}=(\theta_2,\ldots,\theta_d)'$. The CI for θ_1 can in principle be obtained as a set of all θ_1^* which are not rejected by the test of the above hypothesis. The same intuition implies that the confidence interval of a scalar component can be very difficult to construct when a model is given by nonlinear moment restrictions.

[4] We assume the maximum exists. We also assume that $\overline{\Phi}_j$, thus defined, is differentiable on the set $\{\mu\in R^m;\Theta_j\neq\emptyset\}$. On the set $\{\mu\in R^m;\Theta_j=\emptyset\}$ we define $\overline{\Phi}_j(\mu)$ such that the extended function $\overline{\Phi}_j$ is differentiable on R^m.

[5] Wolak's (1991) Lemma 1 (3) establishes that the maximum probability of rejection is achieved in some particular set $\mathbb{B}\equiv C^b-\{\mu:\mu\in C^b\text{ and }\Psi_j(\mu)=0\text{ for only one }j=1,\ldots,m\}$, where $C^b\equiv\{\mu:\Psi_j(\mu)\geq 0,\ j=1,\ldots, m\}-\{\mu:\Psi_j(\mu)>0,\ j=1,\ldots, m\}$. The nonlinearity implies that we cannot further reduce the set of potential maximizers. This implies that in practice we need to simulate the distribution of Wolak's (1991) test statistic over $\mathbb{B}$, which is generally an impossible computational task.