Journal of Econometrics 178 (2014) 86–100

Contents lists available at ScienceDirect

Journal of Econometrics

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

Neglected heterogeneity in moment condition models[☆]

Jinyong Hahn[a], Whitney K. Newey[b], Richard J. Smith[c,d,e,*]

[a] Department of Economics, U.C.L.A., 8283 Bunche Hall, Mail Stop: 147703, Los Angeles, CA 90095, USA

[b] Department of Economics, M.I.T., 50 Memorial Drive, E52-391, Cambridge, MA 02142, USA

[c] Centre for Microdata Methods and Practice, Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE, UK

[d] Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, UK

[e] Economics and Finance Department, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

ARTICLE INFO

Article history:
Available online 12 August 2013

JEL classification:
C13
C30

Keywords:
GMM
GEL
Unconditional moments
Conditional moments
Score and LM tests
Information matrix equality

ABSTRACT

The central concern of this paper is parameter heterogeneity in models specified by a number of unconditional or conditional moment conditions and thereby the provision of a framework for the development of apposite optimal m-tests against its potential presence. We initially consider the unconditional moment restrictions framework. Optimal m-tests against moment condition parameter heterogeneity are derived with the relevant Jacobian matrix obtained in terms of the second order own derivatives of the moment indicator in a leading case. GMM and GEL tests of specification based on generalized information matrix equalities appropriate for moment-based models are described and their relation to optimal m-tests against moment condition parameter heterogeneity examined. A fundamental and important difference is noted between GMM and GEL constructions. The paper is concluded by a generalization of these tests to the conditional moment context and the provision of a limited set of simulation experiments to illustrate the efficacy of the proposed tests.

1. Introduction

For econometric estimation with cross-section and panel data the possibility of individual economic agent heterogeneity is a major concern. In particular, when parameters represent agent preferences investigators may wish to entertain the possibility that parameter values might vary across observational economic units. Although it may in practice be difficult to control for such parameter heterogeneity, the formulation and conduct of tests for parameter heterogeneity are often relatively straightforward. Indeed, in the classical parametric likelihood context, Chesher (1984) demonstrates that the well-known information matrix (IM) test due to White (1982) can be interpreted as a test against random parameter variation. In particular, the White (1980) test for heteroskedasticity in the classical linear regression model is a test for random variation in the regression coefficients. Such tests often provide useful ways of checking for unobserved individual heterogeneity.

The central concern of this paper is parameter heterogeneity in models specified by moment conditions and thereby the provision of a framework for the development of apposite optimal m-tests against its potential presence. We consider both unconditional and conditional model settings. Based on the results in Newey (1985a), to formulate an optimal m-test we find the linear combination of moment functions with maximal noncentrality parameter in the limiting noncentral chi-square distribution of a class of m-statistics under a local random parameter alternative. In a leading case, the optimal linear combination has a simple form, being expressed in terms of the second order own derivatives of the moments with respect to those parameters that are considered possibly to be random, multiplied by the optimal weighting matrix. Thus, the moment conditions themselves provide all that is needed for the construction of test statistics for parameter heterogeneity.

We also consider generalized IM equalities associated with efficient two-step (2S) generalized method of moments (GMM)

☆ This paper is dedicated to the memory of the victims of the February 22 2011 Christchurch NZ earthquake. We are grateful for the constructive comments from the Editor and two referees that have helped to improve this paper. Earlier versions were presented at Nuffield College, Oxford, the Universities of Cambridge, Melbourne and Nottingham, Victoria University of Wellington, the Conference in Honour of P.C.B. Phillips, University of York, and as a Keynote address by the third author at TEST 2010 International Symposium on Econometrics of Specification Tests in 30 Years held at WISE, Xiamen University. Part of the research for the preparation of the initial version of the paper was undertaken while the third author was a Visiting Cambridge Fellow at the University of Canterbury and during research visits at the University of Otago and the Reserve Bank of New Zealand. He gratefully acknowledges the hospitality and support of these institutions. The authors would also like to thank C.E. Hernandez and J. Kang who provided excellent research assistance.

* Corresponding author at: Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, UK.

E-mail address: rjs27@econ.cam.ac.uk (R.J. Smith).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.08.009

(Hansen, 1982) and generalized empirical likelihood (GEL) (Newey and Smith (2004), henceforth NS, and Smith (1997, 2011)) estimation. The 2SGMM-based version of the generalized IM test statistic employs all second derivatives including cross-derivatives of the moments. The GEL form is associated with a more general form of parameter heterogeneity test involving additional components that may be interpreted in terms of a particular correlation structure linking the sample Jacobian and the random variates driving potential parameter heterogeneity.

To provide a background for the subsequent discussion Section 2 reconsiders the IM test of White (1982) and its interpretation as a test against parameter heterogeneity in Chesher (1984). We then consider the effect of parameter heterogeneity on moment conditions in Section 3 and derive the optimal linear combination to be used in constructing the tests in a leading case when the sample Jacobian is uncorrelated with the random heterogeneity variate. We give alternative Lagrange multiplier and score forms of the optimal m-statistic that, using the results of Newey (1985a), maximize asymptotic local power. Section 4 of the paper provides moment specification tests obtained by consideration of generalized forms of the IM equality appropriate for efficient 2SGMM and GEL estimation. These statistics are then compared with those against moment condition parameter heterogeneity developed in Section 3. Components of the 2SGMM form coincide with those of Section 3 whereas the GEL statistic incorporates additional terms that implicitly allow for a particular form of correlation between the sample Jacobian and the random variates potentially driving parameter heterogeneity. These results are illustrated by consideration of empirical likelihood, a special case of GEL that allows a direct application of the classical likelihood-based approach to IM test construction discussed in Section 2. The results of earlier sections are then extended in Section 5 to deal with models specified in terms of conditional moment conditions. Section 6 provides a set of simulation experiments to illustrate the potential efficacy for empirical research of the tests proposed in the paper. The paper is concluded in Section 7. The Appendices contain relevant assumptions and proofs of results and assertions made in the main text.

Throughout the text (x_i, z_i), $(i = 1, \ldots, n)$, will denote i.i.d. observations on the observable d_x-dimensional covariate or instrument vector x and the d_z-dimensional vector z that may include a sub-vector of x. The vector β denotes the parameters of interest with $\mathcal{B}$ the relevant parameter space. Positive (semi-) definite is denoted as p.(s.)d. and f.c.r. is full column rank. Superscripted vectors denote the requisite element, e.g., a^j is the jth element of vector a. UWL will denote a uniform weak law of large numbers such as Lemma 2.4 of Newey and McFadden (1994), and CLT will refer to the Lindeberg–Lévy central limit theorem. "$\xrightarrow{p}$" and "$\xrightarrow{d}$" are respectively convergence in probability and distribution.

2. The classical information matrix test

We first consider the classical fully parametric likelihood context and briefly review the information matrix (IM) test initially proposed in the seminal paper White (1982). See, in particular, White (1982, Section 4, pp. 9–12). The interpretation presented in Chesher (1984) of the IM test as a Lagrange multiplier (LM) or score test for neglected (parameter) heterogeneity is then discussed.

For the purposes of this section it is assumed that z has (conditional) distribution function $F(\cdot, \beta)$ given covariates x known up to the $p \times 1$ parameter vector $\beta \in \mathcal{B}$. We omit the covariates x from the exposition where there is no possibility of confusion. Suppose also that $F(\cdot, \beta)$ possesses Radon–Nikodým conditional density $f(z, \beta) = \partial F(z, \beta)/\partial v$ and that the density $f(z, \beta)$ is twice continuously differentiable in $\beta \in \mathcal{B}$.

2.1. ML estimation

The ML estimator $\hat{\beta}_{ML}$ is defined by

$$\hat{\beta}_{ML} = \arg\max_{\beta \in \mathcal{B}} \frac{1}{n} \sum_{i=1}^{n} \log f(z_i, \beta).$$

Let $\beta_0 \in \mathcal{B}$ denote the true value of β and $E_0[\cdot]$ denote expectation taken with respect to $f(z, \beta_0)$. The IM $\mathcal{I}(\beta_0)$ is then defined by $\mathcal{I}(\beta_0) = -E_0[\partial^2 \log f(z, \beta_0)/\partial\beta\partial\beta']$, its inverse defining the classical Cramér–Rao efficiency lower bound. Under standard regularity conditions, see, e.g., Newey and McFadden (1994), $\hat{\beta}_{ML}$ is a root-n consistent estimator of β_0 with limiting representation

$$n^{1/2}(\hat{\beta}_{ML} - \beta_0) = -\mathcal{I}(\beta_0)^{-1} n^{-1/2} \sum_{i=1}^{n} \partial \log f(z_i, \beta_0)/\partial\beta + O_p(n^{-1/2}). \tag{2.1}$$

Consequently the ML estimator $\hat{\beta}_{ML}$ has an asymptotic normal distribution described by

$$n^{1/2}(\hat{\beta}_{ML} - \beta_0) \xrightarrow{d} N(0, \mathcal{I}(\beta_0)^{-1}).$$

2.2. IM equality and IM specification test

With $E[\cdot]$ as expectation taken with respect to $f(z, \beta)$, twice differentiation of the identity $E[1] = 1$ with respect to β demonstrates that the density function $f(z, \cdot)$ obeys the familiar IM *equality*

$$E\left[\frac{1}{f(z, \beta)} \frac{\partial^2 f(z, \beta)}{\partial\beta\partial\beta'}\right] = E\left[\frac{\partial^2 \log f(z, \beta)}{\partial\beta\partial\beta'}\right] + E\left[\frac{\partial \log f(z, \beta)}{\partial\beta} \frac{\partial \log f(z, \beta)}{\partial\beta'}\right] = 0.$$

Therefore, under correct specification, i.e., z distributed with density function $f(z, \beta_0)$, and given the consistency of $\hat{\beta}_{ML}$ for β_0, by an i.i.d. UWL, the contrast with zero

$$\frac{1}{n} \sum_{i=1}^{n} \frac{1}{f(z_i, \hat{\beta}_{ML})} \frac{\partial^2 f(z_i, \hat{\beta}_{ML})}{\partial\beta\partial\beta'} = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\partial^2 \log f(z_i, \hat{\beta}_{ML})}{\partial\beta\partial\beta'} + \frac{\partial \log f(z_i, \hat{\beta}_{ML})}{\partial\beta} \frac{\partial \log f(z_i, \hat{\beta}_{ML})}{\partial\beta'}\right]$$

consistently estimates a $p \times p$ matrix of zeros. The IM test of White (1982) is a (conditional) moment test (Newey, 1985b) for correct specification based on selected elements of the rescaled moment vector[1,2]

$$n^{1/2} \sum_{i=1}^{n} \frac{1}{f(z_i, \hat{\beta}_{ML})} vec\left(\frac{\partial^2 f(z_i, \hat{\beta}_{ML})}{\partial\beta\partial\beta'}\right)/n. \tag{2.2}$$

2.3. Neglected heterogeneity

The IM test may also be interpreted as a test for neglected heterogeneity; see Chesher (1984). To see this we now regard β as a random vector and the density $f(z, \beta)$ as the *conditional* density of z given β. Absence of parameter heterogeneity corresponds to $\beta = \beta_0$ almost surely.

1 Apart from symmetry, in some cases there may be a linear dependence and, thus, a redundancy between the elements of $\partial^2 f(z, \beta)/\partial\beta\partial\beta'$, in particular, those associated with parametric models based on the normal distribution, e.g., linear regression, Probit and Tobit models.

2 Chesher and Smith (1997) provide a likelihood ratio form of (conditional) moment specification test. An attractive feature of this test is that it admits a "Bartlett correction" by division by a scale factor that creates a statistic with higher order accuracy as compared to conventional moment-based tests.