Economics Letters 171 (2018) 29–33

Contents lists available at ScienceDirect

Economics Letters

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

Moment redundancy test with application to efficiency-improving copulas[☆]

Bowen Hao[a], Artem Prokhorov[a,b,c,*], Hailong Qian[d]

[a] The University of Sydney Business School, Australia
[b] St. Petersburg State University, Russian Federation
[c] Innopolis University, Russian Federation
[d] Saint Louis University, Richard A. Chaifetz School of Business, United States

HIGHLIGHTS

- Moment redundancy is a testable hypothesis. We propose a redundancy test in the context of copula-based pseudo-maximum likelihood estimation.
- A robust and efficiency-improving parametric copula permits improvement in precision at no cost in terms of bias.
- The proposed test can be used to select such copulas.

ARTICLE INFO

Article history:
Received 6 May 2018
Received in revised form 21 June 2018
Accepted 3 July 2018
Available online xxxx

JEL classification:
C12
C13

Keywords:
GMM
Moment redundancy
Copulas

ABSTRACT

Moment redundancy as defined by Breusch et al. (1999) is a testable hypothesis. We propose a simple test of the hypothesis in the context of copula-based pseudo-maximum likelihood estimation considered by Prokhorov and Schmidt (2009b). A robust and efficiency-improving parametric copula permits sizable improvement in precision at no cost in terms of bias and the proposed test can be used to select such copulas.

1. Introduction

In a very well-cited paper, Breusch et al. (1999) define moment redundancy as follows. For a random sample $\{y_i\}_{i=1}^N$, let $g_1(y_i;\theta)$ and $g_2(y_i;\theta)$ be a k_1- and k_2-valued moment function, respectively, of a parameter vector $\theta : p \times 1$. Assume $k_1 \geq p$ so that just the first moment function identifies the true value θ_0. The Generalized Method of Moments (GMM) estimator of θ based on both moment conditions

$$Eg(y_i;\theta_0) \equiv E\begin{bmatrix} g_1(y_i;\theta_0) \\ g_2(y_i;\theta_0) \end{bmatrix} = 0 \tag{1}$$

is usually preferred to the GMM estimator based on only $Eg_1(y_i;\theta_0) = 0$ because the former uses more information (about θ) than the latter.

However, it is possible that $Eg_2(y_i;\theta_0) = 0$ is not informative about θ given $Eg_1(y_i;\theta_0) = 0$. Then, using the two moment conditions is no better than using just $Eg_1(y_i;\theta_0) = 0$, in terms of asymptotic efficiency. The moment function $Eg_2(y_i;\theta_0) = 0$ is *redundant* (for the estimation of θ) if the asymptotic variance matrix of the optimal GMM estimator of θ based on both moment conditions is equal to the asymptotic variance matrix of the optimal GMM estimator based on only $Eg_1(y_i;\theta_0) = 0$.

Breusch et al. (1999) provide the necessary and sufficient condition for moment redundancy and illustrate it using a linear regression. The condition has since received many applications including efficient estimation of panels with time-varying individual effects (Ahn et al., 2001), dynamic panels (Han and Kim, 2014; Sarafidis, 2016), various autoregressive models (Kim et al., 1999; West, 2002; Liu et al., 2010), comparisons of GMM and empirical likelihood based estimators (Shi, 2016; Andrews et al., 2017),

☆ Research for this paper was supported by a grant from the Russian Science Foundation (Project No. 16-18-10432).
* Corresponding author at: The University of Sydney Business School, Australia.
E-mail addresses: bhao9306@uni.sydney.edu.au (B. Hao), artem.prokhorov@sydney.edu.au (A. Prokhorov), hailong.qian@slu.edu (H. Qian).

https://doi.org/10.1016/j.econlet.2018.07.009
0165-1765/

studies of relevance of instruments (Anatolyev, 2007; Antoine and Renault, 2017) and selectivity models (Prokhorov and Schmidt, 2009a; Han and Kim, 2011).

In this paper we propose a simple test of the null of redundancy against the alternative of non-redundancy. Our test uses the condition of Breusch et al. (1999) and is in essence a conditional moment test of Newey (1985) and Tauchen (1985). A closely related paper is Larin (2016), which considers testing whether an extra set of moment conditions helps identification. His test for irrelevance to identification of an extra set of moment conditions (given in Definition 4) is generically similar to our test, except that the variance matrix of the moment conditions is re-estimated using the optimal GMM estimator of the parameters and the asymptotic distribution of the test statistics is therefore somewhat more complicated.

We apply our test to the problem of constructing a copula-based pseudo-maximum likelihood estimator (PMLE) proposed by Prokhorov and Schmidt (2009b). In the setting of the PMLE, a copula provides additional information if the moment conditions arising from using the copula score function are not redundant given the moment conditions implied by the marginal distributions. Prokhorov and Schmidt (2009b) show that there are non-trivial cases when copula-based moment conditions are valid and non-redundant. The new test helps identify such cases.

2. Moment redundancy test

In the standard GMM notation, define the following matrices

$$\Omega = \mathrm{E}\, g\left(y_i;\theta_0\right) g\left(y_i;\theta_0\right)',$$

$$D = \mathrm{E}\,\nabla_\theta\, g\left(y_i;\theta_0\right) = \mathrm{E}\left.\frac{\partial g\left(y_i;\theta_0\right)}{\partial\theta'}\right._{k\times p},$$

where θ_0 denotes the true value of θ and "∇_θ" denotes the gradient. It is well known that the asymptotic variance matrix of the efficient GMM of θ_0 based on $Eg(y_i;\theta_0)=0$ can be written as follows

$$AV = \left(D'\Omega^{-1}D\right)^{-1}.$$

This estimator uses both sets of moment conditions.

Now, consider the GMM estimator based only on $Eg_1(y_i;\theta_0)=0$. Partition the above matrices as follows

$$D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} \mathrm{E}\,\nabla_\theta\, g_1\left(y_i;\theta_0\right) \\ \mathrm{E}\,\nabla_\theta\, g_2\left(y_i;\theta_0\right) \end{bmatrix} \tag{2}$$

$$\begin{aligned}\Omega &= \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \\ &= \mathrm{E}\begin{bmatrix} g_1\left(y_i;\theta_0\right) g_1\left(y_i;\theta_0\right)' & g_1\left(y_i;\theta_0\right) g_2\left(y_i;\theta_0\right)' \\ g_2\left(y_i;\theta_0\right) g_1\left(y_i;\theta_0\right)' & g_2\left(y_i;\theta_0\right) g_2\left(y_i;\theta_0\right)' \end{bmatrix}.\end{aligned} \tag{3}$$

Then, the asymptotic variance of the efficient GMM based on $Eg_1(y_i;\theta_0)=0$ can be written as follows

$$AV_1 = \left(D_1'\Omega_{11}^{-1}D_1\right)^{-1}.$$

Breusch et al. (1999) show that $AV_1 > AV$ in the positive definite sense unless the following redundancy condition holds

$$D_2 = \Omega_{21}\Omega_{11}^{-1}D_1, \tag{4}$$

in which case the two matrices are equal. They also provide a linear projection interpretation of this redundancy condition. Specifically, let $r_2\left(y_i;\theta\right)$ represent the error of the linear projection of g_2 on g_1. That is,

$$r_2\left(y_i;\theta\right) = g_2\left(y_i;\theta\right) - \Omega_{21}\Omega_{11}^{-1}g_1\left(y_i;\theta\right).$$

Then, condition (4) is equivalent to the condition that the expected value of the derivative of r_2 with respect to θ, evaluated at θ_0, is equal to zero. We can write this condition as follows:

$$\mathrm{E}\left(\nabla_\theta\, g_2(y_i) - \Omega_{21}\Omega_{11}^{-1}\nabla_\theta\, g_1(y_i)\right) = 0, \tag{5}$$

where $\nabla_\theta\, g_j(y_i)$, $j=1,2$, is the shorthand notation for the gradient of $g_j(y_i;\theta)$ evaluated at θ_0.

The redundancy test we propose is a simple moment test which tests the validity of (5) assuming that the moment conditions $Eg(y_i;\theta_0)=0$ are valid.

We will need more notation. Let

$$h(y_i;\theta) = \nabla_\theta\, g_2(y_i;\theta) - \Omega_{21}\Omega_{11}^{-1}\nabla_\theta\, g_1(y_i;\theta) \tag{6}$$

and let $h_i = h(y_i;\theta_0)$. Then the moment redundancy condition (4) can be simply written as $\mathrm{E}\,h_i = 0$, where h_i is a random matrix of dimension $k_2\times p$.

When $p > 1$ it is easier to operate with a vectorized version of h_i. It is not difficult to see that it can be obtained from the vectorized versions of $\nabla_\theta\, g_j(y_i)$ using the following equations

$$h_i^v = \mathrm{vec}(\nabla_\theta\, g_2(y_i)) - \mathrm{vec}\left(\Omega_{21}\Omega_{11}^{-1}\nabla_\theta\, g_1(y_i)\right) \tag{7}$$

$$= \mathrm{vec}(\nabla_\theta\, g_2(y_i)) - \left(I_p\otimes\left(\Omega_{21}\Omega_{11}^{-1}\right)\right)\mathrm{vec}(\nabla_\theta\, g_1(y_i)), \tag{8}$$

where h_i^v is a vector with dimension $k_2 p\times 1$. For simplicity, we will assume that $p=1$ in what follows.

Given the valid moment conditions in (1) and a sample of observations $\{y_i\}_{i=1}^N$, it is natural to replace θ_0 in (5) with a GMM estimator based on (1) and to use a sample mean over i in constructing the test statistic for the null that $Eh_i=0$. We now derive the asymptotic distribution of this test statistic.

Let $\hat\theta$ denote the GMM estimator of θ_0 based on $\mathrm{E}\,g\left(\theta_0\right)=0$. It is a standard GMM asymptotic result that $\hat\theta$ satisfies the following equation

$$\sqrt{N}\left(\hat\theta-\theta_0\right) = -\left[D'\Omega^{-1}D\right]^{-1}D'\Omega^{-1}\sqrt{N}\bar g\left(\theta_0\right) + o_p(1), \tag{9}$$

where $\bar g(\theta_0)$ is the sample average of $g(y_i;\theta_0)$.

Define

$$\bar h\left(\hat\theta\right) \equiv \frac{1}{N}\sum_{i=1}^{N} h\left(y_i;\hat\theta\right). \tag{10}$$

Using a Taylor expansion at θ_0, it is easy to show that

$$\sqrt{N}\bar h\left(\hat\theta\right) = \sqrt{N}\bar h\left(\theta_0\right) + D_h\sqrt{N}\left(\hat\theta-\theta_0\right) + o_p(1) \tag{11}$$

where $D_h = \mathrm{E}\,\nabla_\theta\, h\left(\theta_0\right)$ is the expected value of the gradient of $h(y_i,\theta)$, evaluated at θ_0.

Substituting Eq. (9) into Eq. (11) gives:

$$\sqrt{N}\bar h\left(\hat\theta\right) = \sqrt{N}\bar h\left(\theta_0\right) - D_h\left[D'\Omega^{-1}D\right]^{-1}D'\Omega^{-1}\sqrt{N}\bar g\left(\theta_0\right) + o_p(1) \tag{12}$$

$$= \sqrt{N}M\begin{bmatrix}\bar g\left(\theta_0\right) \\ \bar h\left(\theta_0\right)\end{bmatrix} + o_p(1), \tag{13}$$

where $M = \left[\ -D_h\left[D'\Omega^{-1}D\right]^{-1}D'\Omega^{-1},\ I_{\dim h}\ \right]$.

Assuming that $\bar g\left(\theta_0\right)$ and $\bar h\left(\theta_0\right)$ obey a central limit theorem,

$$\sqrt{N}\begin{bmatrix}\bar g\left(\theta_0\right) \\ \bar h\left(\theta_0\right)\end{bmatrix} \overset{a}{\sim} N\left(0,C\right), \quad \text{where} \quad C = \mathrm{E}\begin{bmatrix} g_i g_i' & g_i h_i' \\ h_i g_i' & h_i h_i' \end{bmatrix}, \tag{14}$$

it is no surprise that

$$\sqrt{N}\bar h\left(\hat\theta\right) \overset{a}{\sim} M\cdot N\left(0,C\right) = N\left(0,MCM'\right). \tag{15}$$