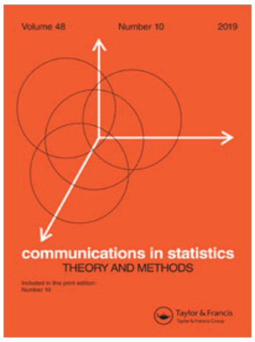

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

New test statistics for hypothesis testing of parameters in conditional moment restriction models

Ziqi Chen & Yan Zhou

To cite this article: Ziqi Chen & Yan Zhou (2019) New test statistics for hypothesis testing of parameters in conditional moment restriction models, Communications in Statistics - Theory and Methods, 48:10, 2521-2528, DOI: 10.1080/03610926.2018.1472775

To link to this article: https://doi.org/10.1080/03610926.2018.1472775

Published online: 25 Mar 2019.

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2019, VOL. 48, NO. 10, 2521–2528
https://doi.org/10.1080/03610926.2018.1472775

New test statistics for hypothesis testing of parameters in conditional moment restriction models

Ziqi Chen[a] and Yan Zhou[b]

[a]School of Mathematics and Statistics, Central South University, Changsha, China; [b]College of Mathematics and Statistics, Institute of Statistical Sciences, Shenzhen University, ShenZhen, China

ABSTRACT

Based on the difference in the objective function proposed by Dominguez and Lobato between the unconstrained and constrained estimators, a simply test is proposed for hypothesis testing of parameters in conditional moment restriction models. This test is guaranteed to be consistent. The asymptotic distribution of the proposed test statistic is proved to be a linear combination of independent χ_1^2 random variables under the null hypothesis. In the simulation study, the power of the proposed test is larger than that of the GMM based test under the alternative hypothesis.

ARTICLE HISTORY
Received 11 October 2016
Accepted 5 March 2018

KEYWORDS
Conditional moment restrictions; test statistic; hypothesis testing

1. Introduction

Models defined in terms of conditional moment restriction arise in various areas of econometrics (Chamberlain 1987; Dominguez and Lobato 2004; Donald, Imbens and Newey 2003), including panel data, instrumental variable settings and rational expectations. Bravo (2012) considers specification tests for this model based on the generalized empirical likelihood (Smith 1997; Newey and Smith 2004) and the references therein include the literature on this test. Statistical hypothesis testing of parameters in conditional moment restriction is important (e.g., Newey and McFadden 1994; Romano, Shaikh, and Wolf 2010). This paper proposes a consistent test based on the objective function in Dominguez and Lobato (2004).

For hypothesis testing of parameter values in the conditional moment restriction models, a distance metric test statistic based on the difference of the GMM criterion between the unconstrained and constrained estimators (DM-GMM) is proposed by Newey and McFadden (1994), which is also mentioned by Romano, Shaikh, and Wolf (2010). Under the null, they show that the DM-GMM converges in distribution to a Chi-square. Motivated by the objective function which the consistent estimator proposed by Dominguez and Lobato (2004) minimizes, we propose a new distance metric test statistic (DM-DL). The DM-DL is simple since it requires no additional user-chosen objects (such as smoothing parameters). Moreover, the limiting distribution of the proposed DM-DL is a weighted sum of independent χ_1^2 random variables under the null. Unlike DM-GMM, the proposed test statistic is always consistent.

The rest of the paper is organized as follows. In Section 2, we propose a test statistic for hypothesis testing of parameters in conditional moment restriction models. Subsequently, we

CONTACT Yan Zhou zhouy1016@szu.edu.cn College of Mathematics and Statistics, Institute of Statistical Sciences, Shenzhen University, ShenZhen 518061, China.

present the asymptotical results of the proposed test as well as the Assumptions on which the theoretical results are based on Section 3. In Section 4, simulation studies are taken to evaluate the performance of the proposed test and its competitor. A brief discussion is presented in Section 6. The proofs of Lemma 1 and Lemma 2 are presented in the Appendix.

2. The test statistics

Let $Y_i = (Z_i, X_i)$ be a random sample in $R^s \times R^d$, which could have common coordinates, for $i = 1, 2, \ldots, n$. We consider the conditional moment restriction

$$E\{g(Z_i, \theta_0)|X_i\} = 0, \quad a.s. \tag{1}$$

for a unique value $\theta_0 \in \Theta$, where $\Theta \subset R^q$. The test problem is $H_0 : \theta_0 = \theta_0^{'}$. Sometimes, we are interested in testing part of the parameter θ. Therefore, we divide θ into (θ_1, θ_2) and suppose θ_1 is of interest with dimension p $(q \geq p > 0)$ and θ_2 is a nuisance parameter with dimension $q - p$. Then, we consider the follow hypothesis test $\dot{H}_0 : \theta_{10} = \theta_{10}^{'}$. The conditional moment restriction model (Equation 1) implies an infinite number of unconditional moment restrictions. The generalized method of moments (GMM), a commonly used method for estimating model (Equation 1), can produce inconsistent estimators since the number of arbitrarily chosen instruments is finite (Dominguez and Lobato 2004). In order to achieve consistent estimators, GMM typically requires an additional assumption that the selected unconditional restrictions identify globally the parameters of interest (Newey and McFadden 1994, Lemma 2.3). In fact, the marginal distribution of the conditioning variables are restricted by this additional assumption, which results in a contradiction, i.e., the consistent estimation of the conditional moment restriction models should be uncorrelated to the distribution of the conditioning variables, it leads to that this distribution plays an important role for GMM estimators since it guarantees global identification of the parameters of interest. Dominguez and Lobato (2004) proposed a new approach to consistently estimate models defined by conditional moment restrictions. Shin (2008) adopted this estimation procedure for the Box-Cox transformation model estimation problems. This method substitutes the conditional restrictions by an infinite number of unconditional moment restrictions which characterize the conditional restriction fully. These infinite unconditional restrictions here come from the expectation of the function of interest times a class of indicators functions. Unlike GMM, the identification problem does not arise for this method, since it is based on the conditional moment restrictions directly. Specifically, we have

$$E\{g(Z_i, \theta_0)|X_i\} = 0 \quad a.s. \quad \Leftrightarrow \quad H(\theta_0, x) = 0 \quad \text{for almost all } x \in R^d \tag{2}$$

where $H(\theta, x) = E\{g(Z_i, \theta)I(X_i \leq x)\}$ (Billingsley 1995). Based on this equivalence, from (1), we have that $L(\theta)$ is minimized at $\theta = \theta_0$ uniquely, where $L(\theta) := E\{H(\theta, X_i)^2\}$. Based on the sample analogue principle, the estimator of θ_0 in Equation (1) is proposed as

$$\hat{\theta} = \text{argmin}_\theta L_n(\theta)$$

where

$$L_n(\theta) := \frac{\sum_{i=1}^n H_n(\theta, X_i)^2}{n} \tag{3}$$

with $H_n(\theta, x) := \frac{\sum_{i=1}^n g(Z_i, \theta)I(X_i \leq x)}{n}$.

Motivated by the above estimation procedure, we propose a so-called DM-DL statistic for testing $H_0 : \theta_0 = \theta_0'$ as

$$W(\theta_0') = -2n\{L_n(\hat{\theta}) - L_n(\theta_0')\}$$

Similarly, the distance metric statistic for testing the composite hypothesis $\dot{H}_0 : \theta_{10} = \theta_{10}'$ is proposed to be

$$\dot{W}(\theta_{10}') = -2n\{L_n(\hat{\theta}) - L_n(\theta_{10}', \tilde{\theta}_2)\}$$

where $\tilde{\theta}_2 = \text{argmin}_{\theta_2} L_n(\theta_{10}', \theta_2)$. We note that the test proposed is based on the difference between minimum of objective function $L_n(\theta)$ for the case when the null hypothesis of θ is satisfied and that for the case when no restriction is imposed for θ. This is essentially a quasi LR test. Intuitively, the DM-DL type test is consistent since $L(\theta)$ is uniquely minimized at the true parameter point, whereas, the DM-GMM test can be inconsistent, because the GMM objective function may has more than one global minimum, that shown in the Example 1 and Example 2 of Dominguez and Lobato (2004). Moreover, the proposed test statistics is simple, since we do not need to introduce any user chose objects such as the order of a lag or a bandwidth number.

3. Assumptions and theorems

Define

$$\psi(\theta, y) := S(\theta, y, P, P) + S(\theta, P, y, P) + S(\theta, P, P, y)$$

where $S(\theta, Y_i, Y_j, Y_l) := g(Z_i, \theta)g(Z_j, \theta)I(X_i \leq X_l)I(X_j \leq X_l)$. Let ∇_m be the mth partial derivative operator with respect to θ,

$$|\nabla_m|\sigma(\theta) := \sum_{i_1,\ldots,i_m} \left| \frac{\partial^m}{\partial\theta_{i_1}\cdots\partial\theta_{i_m}} \sigma(\theta) \right|$$

and the symbol $\|\cdot\|$ denotes the matrix norm:$\| (a_{ij}) \| = (\sum_{i,j} a_{ij}^2)^{1/2}$. In order to achieve the lemmas and theorems below, we state the following assumptions, which are mild and can be found in Sherman (1993, 1994) and Shin (2008):

Assumption 1: $E\{g(Z_i, \theta)|X_i\} = 0$ a.s. if and only if $\theta = \theta_0$.

Assumption 2: Θ is a compact subset of R^q.

Assumption 3: The sample $(Z_i, X_i)_{i=1}^n$ is *i.i.d.*

Assumption 4: $\{g(\cdot, \theta), \theta \in \Theta\}$ is Euclidean for an envelope F with $EF^2 < \infty$.

Assumption 5: $L(\theta)$ is continuous on Θ.

Assumption 5': $g(z, \cdot)$ is continuous in Θ for each $z \in R^s$.

Assumption 6: Let $\mathcal{N}$ be a neighborhood of θ_0.

(1) For each y, all mixed second partial derivatives of $\psi(\cdot, y)$ exist on $\mathcal{N}$.

(2) There is an integral function $M(y)$ such that, for all y and $\theta \in \mathcal{N}$,

$$\| \nabla_2\psi(\theta, y) - \nabla_2\psi(\theta_0, y) \| \leq M(y)|\theta - \theta_0|$$

(3) $E|\nabla_1\psi(\theta_0, \cdot)|^2 < \infty$

(4) $E|\nabla_2|\psi(\theta_0, \cdot) < \infty$
(5) The matrix $E\nabla_2\psi(\theta_0, \cdot)$ is positive definite.

Assumption 7: $\theta_0 \in \mathrm{int}(\Theta)$.

Remark 1. By Assumption 1 and Equation (2), we have that $L(\theta)$ is uniquely minimized at θ_0 and $L(\theta_0) = 0$. Therefore, We assume that $E\nabla_2\psi(\theta_0, \cdot)$ is positive definite.

Lemma 1. *If Assumptions 1–5 hold, then $\hat{\theta} \to \theta_0$ in probability.*

Remark 2. A dominated convergency arguments, combining Assumption 4, shows that Assumption $5^{'}$ is stronger than Assumption 5. The consistency of $\hat{\theta}$ only requires the continuity of $L(\theta)$, needs not the continuity of $g(z, \cdot)$. Notice that even though $g(z, \cdot)$ is discontinuous, $L(\theta)$ could be continuous. Thus, this condition for consistency is weaker than that of Dominguez and Lobato (2004).

Lemma 2. *Under Assumptions 1–4, Assumption $5^{'}$, and Assumption 6,*

$$L_n(\theta) - L_n(\theta_0) = \frac{1}{2}(\theta - \theta_0)^{'} V(\theta - \theta_0) + \frac{1}{\sqrt{n}}(\theta - \theta_0)^{'} W_n + o_p(|\theta - \theta_0|^2) + o_p(\frac{1}{n})$$

uniformly in $o_p(1)$ neighborhoods of θ_0, where $3V = E\nabla_2\psi(\theta_0, Y_i)$ and $W_n = \frac{1}{\sqrt{n}} \sum_i \nabla_1\psi$ (θ_0, Y_i)

Applying Theorems 1 and 2 of Sherman (1993), by Lemmas 1, 2 and Equation (A4), we have:

Lemma 3. *Under Assumptions 1–4, Assumption $5^{'}$, and Assumption 6 and 7,*

$$\sqrt{n}(\hat{\theta} - \theta_0) = -V^{-1} W_n + o_p(1)$$

and thus,

$$\sqrt{n}(\hat{\theta} - \theta_0) \longrightarrow N(0, V^{-1} \triangle V^{-1})$$

in distribution, as $n \to \infty$, where $\triangle = E\nabla_1\psi(\theta_0, Y_i)\nabla_1\psi(\theta_0, Y_i)^{'}$.

The following theorem is straightforward by Lemma 2 and Lemma 3.

Theorem 4. *Under Assumptions 1–4, Assumption $5^{'}$, and Assumption 6–7, when H_0 is true,*

$$W(\theta_0^{'}) \longrightarrow \sum_{k=1}^{q} c_k \chi_{1k}^2$$

in distribution, as $n \to \infty$, where c_k are the eigenvalues of $V^{-1/2} \triangle V^{-1/2}$ and χ_{1k}^2 are independent χ_1^2 random variables.

We divide θ into (θ_1, θ_2), and suppose θ_1 is of interest with dimension p, with $q \geq p > 0$, and θ_2 is a nuisance parameter with dimension $q - p$. V is partitioned similarly into submatrices as

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

Theorem 5. *Under Assumptions 1–4, Assumption $5^{'}$, and Assumption 6–7, when $\dot{H}_0$ is true,*

$$\dot{W}(\theta_{10}^{'}) \longrightarrow \sum_{k=1}^{p} c_k \chi_{1k}^2$$

in distribution, as $n \to \infty$, where c_k are the p positive eigenvalues of $G^{1/2} \bigtriangleup G^{1/2}$ with $G = V^{-1} - \mathrm{diag}(\mathbf{0}, V_{22}^{-1})$ and χ_{1k}^2 are independent χ_1^2 random variables.

Remark 3. The matrix $G^{1/2} \bigtriangleup G^{1/2}$ is non-negative definite and $\mathrm{rank}(G^{1/2} \bigtriangleup G^{1/2}) = p$.

Remark 4. When $p = q$, $G = V^{-1}$, Theorem 5 is reduced to Theorem 4.

Remark 5. By Theorem 4 and Theorem 5, we can use $W(\theta_0^{'})$ and $\dot{W}(\theta_{10}^{'})$ to construct asymptotically valid rejection regions for the parameters of interest.

4. Simulation study

To investigate the performance of the proposed method, we consider the model in Example 2 of Dominguez and Lobato (2004), i.e., $E(Y|X) = \theta_0^2 X + \theta_0 X^2$ with $\theta_0 = 1.25$ and $\mathrm{Var}(Y|X)$ being constant. Here, $X \sim N(1, 1)$. For the GMM, we adopt the optimal instrument $W = 2\theta X + X^2$ as proposed in Chamberlain (1987). The parameter θ_0 could not be identified for the GMM, because $\theta = 1.25$ and $\theta = -3$ both satisfy the equation $E\{(Y - \theta^2 X - \theta X^2) W\} = 0$. The null hypothesis is $H_0 \colon \theta = \theta_0$. We could get the reject region for DM-DL test by Theorem 4 for given type I error γ. The reject region for DM-GMM test could also be derived for given type I error γ.

We consider two sample sizes, i.e., $n = 50$ and 100. γ is set to 0.01. We generate datasets when θ is set to be 1.25, -2.9, -1.2, 0.45 and 2.05, respectively. For all experiments, the number of replications is 1000 for each sample size. Table 1 reports the empirical rejection probabilities under the null ($\theta = \theta_0$) and under the alternatives ($\theta = -2.9$, $\theta = -1.2$, $\theta = 0.45$ and $\theta = 2.05$) for $n = 50$ and $n = 100$, respectively. We observe that the empirical type I errors of the two methods both agree well with the nominal value (i.e., 1%), whereas the empirical rejections probabilities under the alternative hypotheses based on DM-DL are larger than those of DM-GMM, respectively. Therefore, under controlling type I error, DM-DL test performs better than DM-GMM test in terms of power.

5. Conclusions

In this paper, for hypothesis testing of parameter values in the conditional moment restriction models, we proposed a DM-DL statistic for testing $H_0 : \theta_0 = \theta_0^{'}$ or test part of θ_0. Under the null hypothesis, we deduce the asymptotical distribution of test statistics as in Theorem 4

Table 1. Summary of empirical rejection probabilities over 1000 simulations.

	θ	1.25	−2.9	−1.2	0.45	2.05
n = 50	DM-GMM	0.006	0.035	0.899	0.965	0.995
	DM-DL	0.003	1.000	1.000	1.000	1.000
n = 100	DM-GMM	0.006	0.301	0.998	1.000	1.000
	DM-DL	0.004	1.000	1.000	1.000	1.000

(or Theorem 5), via which the rejection region could be derived for the given type I error. The analysis is routine in the framework of test. In simulation studies, we consider a model, $E(Y|X) = \theta_0^2 X + \theta_0 X^2$ with $\theta_0 = 1.25$ and $\mathrm{Var}(Y|X)$ being constant. The simulation shows that the power of the DM-DL is larger than that of DM-GMM under the alternative hypotheses. We could see that the DM-DL statistic is an effective method for hypothesis testing of parameter values in the conditional moment restriction models.

Acknowledgements

We are grateful to the anonymous referee, an associate editor and the co-editor for helpful and constructive comments on earlier versions.

Funding

Chen's research is supported in part by National Nature Science Foundation of China (Project Nos. 11871477 and 11401593) and Hunan Provincial Natural Science Foundation of China (Project No. 2016JJ3138). Yan Zhou's research was supported by the National Science Foundation of China (Grant No. 11701385), National Statistical Research Project (Grant No. 2017LY56) and Doctor start fund of Guangdong Province (Grant No. 2016A030310062).

References

Billingsley, P. 1995. *Probability and measure.* New York: Wiley and Sons.

Bravo, F. 2012. Generalized empirical likelihood testing in semiparametric conditional moment restrictions models. *Econometrics Journal* 15:1–31.

Chamberlain, G. 1987. Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics* 34:305–34.

Dominguez, M., and I. Lobato. 2004. Consistent estimation of models defined by conditional moment restrictions. *Econometrica* 72:1601–15.

Donald, S. G., G. W. Imbens, and W. K. Newey. 2003. Empirical likelihood estimation and consistent tests with conditional moment testrictions. *Journal of Econometrics* 117:55–93.

Newey, W. K., and D. McFadden. 1994. Large sample estimation and hypothesis testing. In R. F. Engle, and D. McFadden (Eds.), *Handbook of econometrics*, Vol. 4, 2111–2245. Amsterdam: Elsevier.

Newey, W. K., and R. J. Smith, 2004. Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica* 72:219–56.

Romano, J. P., A. M. Shaikh, and M. Wolf. 2010. Hypothesis testing in econometrics. *Annual Review of Economics* 2:75–104.

Sherman, R. P. 1993. The limiting distribution of the maximum rank correlation estimator. *Econometrica* 61:123–37.

Sherman, R. P. 1994. Maximal inequalities for degenerate U-processes with applications to optimization estimators. *Annals of Statistics* 22:439–59.

Shin, Y. 2008. Semiparametric estimation of the Box-Cox transformation model. *Econometrics Journal* 11:517–37.

Smith, R. J. 1997. Alternative semi-parametric likelihood approaches to generalised method of moments estimation. *Economic Journal* 107:503–19.

Appendix

Proof of Lemma 1. (a) By Assumption 1 and Equation (2), we know that $L(\theta)$ is uniquely minimized at θ_0; (b) Assumption 2; (c) Assumption 5.

Simple calculations show that

$$L(\theta) = ES(\theta, Y_i, Y_j, Y_l)$$

I define the 3rd order U-process

$$U_n^3 S(\theta, \cdot) = \{n(n-1)(n-2)\}^{-1} \sum_{i \neq j \neq l} S(\theta, Y_i, Y_j, Y_l)$$

By Assumption 4, we can get that

$$\sup_\theta |L_n(\theta) - U_n^3 S(\theta, \cdot)| = O_p\left(\frac{1}{n}\right)$$

Moreover, because of Assumption 4, by using Corollary 7 in Sherman (1994), we have

$$\sup_\theta |U_n^3 S(\theta, \cdot) - L(\theta)| = O_p\left(\frac{1}{\sqrt{n}}\right)$$

Then, we get easily that

$$\sup_\theta |L_n(\theta) - L(\theta)| = O_p\left(\frac{1}{\sqrt{n}}\right) \tag{A1}$$

which is stronger than the uniform convergence. Combining (a), (b), (c) and Equation (A1), we obtain the result by Theorem 2.1 of Newey and McFadden (1994). □

Proof of Lemma 2. From Assumption $5'$, we can deduce easily that

$$\sup_{\theta \in \mathcal{N}_0} |L_n(\theta) - L_n(\theta_0) - U_n^3 S^*(\theta, \cdot)| = o_p\left(\frac{1}{n}\right) \tag{A2}$$

where $\mathcal{N}_0$ denote the $o_p(1)$ neighborhoods of θ_0 and $S^*(\theta, Y_i, Y_j, Y_l) := S(\theta, Y_i, Y_j, Y_l) - S(\theta_0, Y_i, Y_j, Y_l)$. By Equation (6) of Sherman (1994), we have the decomposition of the U-process $U_n^3 S^*(\theta, \cdot)$

$$U_n^3 S^*(\theta, \cdot) = L(\theta) - L(\theta_0) + \frac{1}{n}\sum_i S_1^*(\theta, Y_i) + \frac{1}{n(n-1)}\sum_{i \neq j} S_2^*(\theta, Y_i, Y_j) + U_n^3 S_3^*(\theta, \cdot) \tag{A3}$$

where $S_i^*(\theta, \cdot)$ is P-degenerate for each θ in Θ, for $i = 1, 2, 3$. Note that, here, $S_1^*(\theta, y) = S^*(\theta, y, P, P) + S^*(\theta, P, y, P) + S^*(\theta, P, P, y) - 3\{L(\theta) - L(\theta_0)\}$.

Note that $E\{\psi(\theta, \cdot) - \psi(\theta_0, \cdot)\} = 3\{L(\theta) - L(\theta_0)\}$. We do the Taylor expansion of $\psi(\theta, \cdot)$ around θ_0, following the standard arguments (see, Sherman 1993) based on Assumption 6, we can get

$$L(\theta) - L(\theta_0) = \frac{1}{2}(\theta - \theta_0)' V(\theta - \theta_0) + o_p(|\theta - \theta_0|^2) \tag{A4}$$

and

$$\frac{1}{n}\sum_i S_1^*(\theta, Y_i) = \frac{1}{\sqrt{n}}(\theta - \theta_0)' W_n + o_p(|\theta - \theta_0|^2) \tag{A5}$$

uniformly over $o_p(1)$ neighborhoods of θ_0, where $W_n = \frac{1}{\sqrt{n}} \sum_i \nabla_1 \psi(\theta_0, Y_i)$. Moreover, Lemma 6 and Corollary 8 in Sherman (1994), plus dominated convergence arguments, show that

$$\frac{1}{n(n-1)} \sum_{i \neq j} S_2^*(\theta, Y_i, Y_j) + U_n^3 S_3^*(\theta, \cdot) = o_p\left(\frac{1}{n}\right) \tag{A6}$$

uniformly over $o_p(1)$ neighborhoods of θ_0. Then, the result is established by Equation (A2)–(A6). □