Report on "Using Invalid Instruments on Purpose: Focused Moment Selection and Averaging for GMM"

by

Francis J. DiTraglia

This paper studies the moment selection problem in moment based econometrics models. The author emphasizes that one should balance the validity and the relevance of the moments by minimizing mean squared error (MSE) of the GMM estimator for the parameter of interest. Based on this conviction, the author provides the focused moment selection criterion (FMSC) as an asymptotically unbiased estimator for the MSE of GMM estimator. The subset of moments which minimizes FMSC is suggested to be selected. This paper also explores the effect of moment selection on inference. A Bonferroni-type of size corrected inference procedure is proposed to ensure that the inference after the moment selection/averaging has valid asymptotic size. A simple linear instrument selection simulation example is used to investigate the finite sample performance of the proposed moment selection method and post moment selection inference procedure. The author employs an empirical example from the development literature to demonstrate his moment selection procedure and the algorithm to construct the confidence interval. This example reassures the result in the development literature that malaria transmission directly influences income, even after controlling for institutions.

1 General Comments

This paper assumes that one can divide his/her moment conditions into two groups: valid moments and locally misspecified moments. A moment function is called locally misspecified if its expectation evaluated at the true parameter does not equal to zero in finite samples but goes to zero at the $n^{-\frac{1}{2}}$ rate. This assumption enables the author to derive a non-degenerated mean squared error for the GMM estimator based on the first order asymptotic theory. However, the reader might wonder why a set of valid moments is needed for selection of the locally misspecified moments. Will the moment selection method based on FMSC still work if all moments are locally misspecified? I guess the answer is no, because intuitively, the method proposed in this paper has to treat the expectations of the locally misspecified moments as unknown parameters, and valid statistical inference has to be conducted on the local values of these parameters. For this purpose, there must be a preliminary GMM estimator (based on the valid moments) which is free of first order bias. The author may need to make this point clear in the paper.

Several related references are missing in the paper. Newey (1985) studies the GMM estimator with locally misspecified moment conditions. The asymptotic distribution of the GMM estimator and specification tests are investigated in this paper. More recently, Guggenberger (2012), and Guggenberger and Kumar (2012) investigate the effect of locally misspecified moment conditions on accuracy of inference on structural parameters.

The FMSC constructed in this paper is shown to be an asymptotically unbiased estimator of the MSE of GMM estimator. As a moment selection criterion, the asymptotic unbiasedness does not seem to be a very attractive property. First, from the empirical application perspective, this property has no implication about how the moment selection based on FMSC behaves in any individual experiment, no matter how large the sample size is. Second, this property does not take the variance of the FMSC into account. Intuitively, the larger variance the FMSC has, the less reliable the selected moments will be. Again, the negative effect of large variance on FMSC's performance in moment selection cannot be alleviated even with larger sample size.

The Bonferroni-type of size corrected inference procedure does have the desired valid asymptotic size, as pointed out in Loh (1985) and more recently in Leeb and Pötscher (2013). However, I have two concerns about this approach in the framework of this paper. First, implementation of the proposed inference procedure is computationally intensive. The computational burden induced by Algorithm 4.1 seems so high that the author has to use reduced grid of parameter values and reduced number of simulation rounds to implement the algorithm to the simple example in the paper. Therefore, when it comes to a more involved model, the computational burden might be too formidable for any practitioner to implement it. Second, the author does not investigate whether the confidence interval constructed according to Algorithm 4.1 is sharp enough so make any meaningful empirical inference. If one only cares about the size of the inference, why does not he/she simply use the confidence interval constructed using the valid moment conditions?

In the empirical example (Section 5), the author does not follow the algorithm (Algorithm 4.1) he proposes to construct the confidence interval. According to Algorithm 4.1, one should try all values (empirically, one may try a grid of values) in the $100(1-\delta)\%$ confidence region $T(\widehat{\tau})$ of τ, and take the widest interval for μ as the conservative confidence interval. However, in the empirical example he only constructs and compares the traditional confidence interval for four different values of ($\widehat{\tau}$, 0, $\widehat{\tau}_{\max}$ and $\widehat{\tau}_{\min}$), and draws the conclusion that the confidence interval is insensitive to the value of τ. The reason given by the author for not doing so is that the minimization and maximization problems are badly behaved. The author does not show how they behave nor possible reason or solution for this problem. As pointed out in Loh (1985) and more recently in Leeb and Pötscher (2013), the confidence interval based on single or random critical value typically have the size distortion even in the asymptotic sense. Moreover, if Algorithm 4.1 is indeed used, I am wondering whether the resulting confidence interval would be too big for one to draw meaningful empirical conclusion.

2 Minor Comments

1. In line 25 of page 13, it might be better to use $\widetilde{\Omega}_{gg}$ instead of $\widetilde{\Omega}_{11}$, in accordance with the notation elsewhere in the paper.

2. Page 20, to implement CC-MSC-BIC, what moments should be chosen if GMM-BIC and CCIC-BIC give rise to different sets of moments?

3. Page 28, the third line, Corollary 12 should be Corollary 4.1; in Equation (15), Ω should be $\widehat{\Omega}$.

4. Page 29, in Algorithm 4.1, it is better to point out that B is the number of simulation rounds.

5. Page 43, proof of Theorem 3.1, the third line, $f_n(\theta_0)$ should be $g_n(\theta_0)$.

6. Page 43, proof of Theorem 4.1, the second line, $\sqrt{n}[\Xi_S f_n(\theta)]$ should be $\sqrt{n}[\Xi_S f_n(\widehat{\theta}_S)]$.

7. Page 44, the third line, Assumption 2.2 (viii) should be Assumption 2.2 (h); and the fourth line, $f_n(\theta_0)$ should be $f_n(\widehat{\theta}_S)$.

References

Guggenberger, P. (2012), "On the Asymptotic Size Distortion of Tests When Instruments Locally Violate the Exogeneity Assumption," Econometric Theory, 2012, 28(2), 387-421.

Guggenberger, P. and G. Kumar (2012), "On the Size Distortion of Tests after an Overidentifying Restrictions Pretest," Journal of Applied Econometrics, 2012, 27(7), 1138-1160.

Leeb, H. and P.M. Pötscher (2013), "Testing in the Presence of Nuisance Parameters: Some Comments on Tests Post-Model-Selection and Random Critical Values," Working Paper.

Loh, W.-Y. (1985), "A New Method for Testing Separate Families of Hypotheses," Journal of the American Statistical Association, 80, 362-368.

Newey, W. (1985), "Generalized Method of Moments Specification Testing," Journal of Econometrics, 29, 229-256.