Report on "Using Invalid Instruments on Purpose: Focused Moment Selection and Averaging for GMM" by F. J. Di Traglia

The paper considers a situation where a number of valid moment restrictions identifies a parameter of interest and additional (locally) invalid moment restrictions are available. Including (some of) them for GMM estimation typically increases bias but reduces variance. The author therefore suggests using asymptotic MSE (AMSE) to choose the set of moments when implementing a GMM estimator. AMSE depends on unknwon quantities and the author provides an asymptotically unbiased estimate of AMSE, see (2) on p.11. The performance of the resulting "focused moment selection estimator" (FMSE) is investigated in a Monte Carlo study for the linear IV model (Section 3.3).

In Section 4, the author investigates estimators given as weighted averages over GMM estimators using different sets of moment conditions (see (12) on p.25). When constructing confidence intervals based on this moment average estimator, the author suggests using the Bonferroni method (where in a first step a confidence region is constructed for the nuisance parameter τ that cannot be consistently estimated). Again, a Monte Carlo study in the linear IV model investigates the finite sample coverage probabilities of the confidence interval and an empirical application is given.

Comments:

1. No convincing justification is provided that shows that the suggested FMSE in (2) is to be preferred (in terms of MSE say) over the standard GMM estimator that only uses the valid moments. Likewise, no justification is provided that shows that the confidence interval (constructed in Algorithm 4.1) based on the moment average estimator is to be preferred (in terms of coverage probability and/or volume) over a standard confidence interval based on the GMM estimator that only uses the valid moments. More precisely:

i) It is theoretically unclear whether the *asymptotically unbiased* criterion for AMSE is picking the "best" subset of moments in an MSE perspective. The method would be more convincing if the criterion converged to AMSE in probability (but I understand that this cannot be achieved).

ii) As shown in your Monte Carlos and as one might expect, in instances where the invalid moments are only slightly invalid but very relevant the suggested FMSE might be beneficial, but when the additional moments are only slightly relevant and/or highly invalid, the standard approach is to be preferred. I suspect that your conclusion (p.17) that "taken on the whole, ... the potential advantage of using the valid estimator is small" is very specific to the design of the Monte Carlo study. E.g. if one increased the relevance of the valid moments (that is, in (7) take something larger than .1) the potential gains from using an additional invalid moment would be always small(er).

iii) The problem of invalid instruments (or moments) and weak instruments (or identification) are closely related, yet your theoretical setup assumes strong instruments. In fact, the AMSE formula in Corollary 3.1 breaks down under weak instrument asymptotics as in Stock and Wright (2000). Therefore, the theory developed here is not robust to the strength of identification.

iv) As you allude to on the bottom of p.14, asymptotically the MSE is finite, but in finite samples the MSE maybe unbounded, see Kinal (1980). In such instances, the finite sample MSE approximation is clearly not helpful. This occurs when the degree of overidentification is small or when the involved random variables have certain distributions. These shortcomings should be discussed.

v) Regarding the confidence intervals based on the moment average estimators: the suggested method is computationally complex (as always for Bonferroni type methods) and conservative. Why should they be used over the standard approach that uses only all the valid moments? The Monte Carlos should at least provide some length comparisions.

2. Regarding the literature review, the following papers could also be cited: Chen, Jacho-Chávez and Linton (2002) consider "moment averaging" in a moment condition model; Caner and Sandler (2011), Imbens et al (2011), Berkowitz, Caner, Fang (2012), and many references therein, consider estimation/inference under (locally) invalid moment conditions. Regarding Theorem 4.1, Newey (1985) studies the properties of the J test under local moment violation.

3. In the linear IV model: could you include the OLS estimator as another alternative in your choice set? The same intuition applies: if the endogeneity is not too large, OLS maybe better than 2SLS.

4. i) Wouldn't it more precise to use the title "Potentially using invalid instruments on purpose..."? ii) On p.8, in Theorem 2.2. shouldn't M_h be M_g? iii) Can you provide primitive conditions for Assumption 2.1 in the linear instrumental variables model? iv) Explain why you do include by default all valid instruments? Why don't you also allow choosing subset from the valid moments? v) In Table 1, rather than reporting absolute differences, percentages would be more informative I think.

References:

Berkowitz, D., M. Caner, and Y. Fang (2012): "The Validity of Instruments Revisited," Journal Of Econometrics, 166, 255-267.

Caner, M and M. Sandler (2011): "A New Paradigm: A Joint Test of Structural and Correlation Parameters in Instrumental Variables Regression when Perfect Exogeneity is Violated," unpublished working paper, NCSU.

Chen, X., D. Jacho-Chávez and O. Linton (2002): "An Alternative Way of Computing Efficient Instrumental Variable Estimators", unpublished working paper, Yale University.

Imbens G, Kolesar M, Chetty R, Friedman J, Glaeser E. (2011): "Identification and Inference with Many Invalid Instruments," unpublished working paper, Harvard University.

Kinal, Terrence W, 1980. "The Existence of Moments of k-Class Estimators," Econometrica, 48(1), 241–49.

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

Stock, J. and J. Wright (2000): "GMM with weak identification," Econometrica, 68, 1055-1096.