Tests for Relevance and Redundancy of Moment Conditions

Alexander Larin

National Research University
Higher School of Economics

"Microeconomic assumptions of macroeconomic models"
Research Group Seminar
Nizhny Novgorod, November 3, 2016

The study was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2015-2016 (grant 15-05-0053) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

Motivation (White Spot in Theory)

- Identification in GMM is well examined:
  - Consequences of weak (lack of) identification — Stock, Wright, and Yogo (2002);
  - Test for weak (lack of) identification — Cragg and Donald (1993), Stock and Yogo (2005), Hahn and Hausman (2002), Wright (2003), Inoue and Rossi (2011);
  - Asymptotics for the case of weak identification — Staiger and Stock (1997) Han and Phillips (2006) Newey and Windmeijer (2009);
  - Inference robust to the weak (lack of) identification — Wang and Zivot (1998), Kleibergen (2005), Moreira (2009);
  - Procedures to select relevant moment conditions — Hall and Peixe (2003), Hall,Inoue, Jana, and Shin (2007), Cheng and Liao (2015).
- White spot — a test for relevance of a particular set of moment conditions.

Motivation (Empirical Application)

- Testing if a set of variables affects agents' expectations in models with rational expectations (i.e., C-CAPM, NKPC).
- Selecting relevant (non-redundant) moment conditions.
- Testing for heteroscedasticity, using results of Cragg (1983).
- Testing if using macro-variables as mean values for micro-variables can improve parameter estimates (Imbens and Lancaster, 1994).
- Testing if GMM estimation of a system of equations is more efficient.

GMM Estimator

- Moment conditions

$$\mathbb{E}\left[f(Z_t, \theta_0)\right] = 0. \tag{1}$$

- GMM estimator of θ_0

$$\hat{\theta} = \underset{\theta \in \Theta}{\arg\min}\ f_T(\theta)' W_T(\theta)^{-1} f_T(\theta), \tag{2}$$

where $f_T(\theta) = T^{-1} \sum_{t=1}^{T} f(Z_t, \theta)$, $W_T(\theta)^{-1}$ is a positive definite weighting matrix, $W_T(\theta) \to_p W(\theta)$.

Definition of Relevance

- Consider subset $f_{2T}(\theta)$ of moment functions $f_T(\theta)$.
- Relevance
  - $f_{2T}(\theta)$ can potentially help in local identification: $\mathbb{E}(f_{2T}(\theta)) \neq 0$ for all $\theta \neq \theta_0$ in a close neighborhood of θ_0.
  - Necessary and sufficient condition, which our statistic is based on:

$$\boxed{G_2 \neq 0,} \tag{3}$$

    where $G_2 = \partial\,\mathbb{E}(f_{2T}(\theta_0))/\partial\theta'$.
  - If $G_2 \neq 0$, but higher order derivatives are not zero, $f_{2T}(\theta)$ still may help in identification, but standard asymptotic results does not hold anymore.

Definition of Conditional Relevance

- Note that true values θ_0 can be defined by

$$\theta_0 = \arg\min_{\theta \in \Theta} \; E(f_T(\theta))' W(\theta)^{-1} \mathbb{E}(f_T(\theta)). \tag{4}$$

- Hessian of $S(\theta) = \mathbb{E}(f_T(\theta))' W(\theta)^{-1} \mathbb{E}(f_T(\theta))$ as a measure of the strength of identification:

$$\left.\frac{\partial^2 S(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\theta_0} = G_1' W_{11} G_1 + G_\Delta' W_\Delta G_\Delta, \tag{5}$$

where $G_\Delta = G_1 - W_{21} W_{11}^{-1} G_2$, $W_\Delta = W_{22} - W_{21} W_{11}^{-1} W_{12}$.

- Conditional relevance
  - $f_{2T}(\theta)$ increase strength of identification (improves identification).
  - Necessary and sufficient condition, which our statistic is based on:

$$\boxed{G_\Delta \neq 0.} \tag{6}$$

Identification Strength Illustration

- Red line is for stronger identification

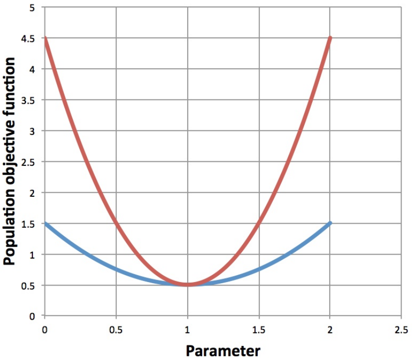

Figure 1. Strength of Identification

Conditional Relevance under Optimal Weighting

- W is the asymptotic covariance matrix Ω of $f_T(\theta_0)$.
- $G_\Delta \neq 0$ is also a condition for non-redundancy: $f_{2T}(\theta)$ improves efficiency (Breusch, Qian, Schmidt and Wyhowski, 1999).
- $f_{2T}(\theta)$ is conditionally relevant given $f_{1T}(\theta)$ if *additional* information it contains $f_{\Delta T}(\theta) = f_{2T}(\theta) - \Omega_{21}\Omega_{11}^{-1} f_{1T}(\theta)$ is relevant.

Special Cases

- Linear IV regression
  - Moment conditions

$$\mathbb{E}\left[\begin{array}{c}(y_t - x_t'\theta_0)z_{1t}\\(y_t - x_t'\theta_0)z_{2t}\end{array}\right] = 0. \tag{7}$$

  - Conditional relevance (non-redundancy)

$$G_\Delta = -\mathbb{E}\left(x_t z_{\Delta t}'\right) = 0, \tag{8}$$

  where $z_{\Delta t} = z_{2t} - \mathrm{cov}(z_2', z_1)\mathrm{var}^{-1}(z_1)z_{2t}$.
- Nonlinear IV regression
  - Moment conditions

$$\mathbb{E}\left[\begin{array}{c}(y_t - h(x_t,\theta_0)z_{1t}\\(y_t - h(x_t,\theta_0)z_{2t}\end{array}\right] = 0. \tag{9}$$

  - Conditional relevance (non-redundancy)

$$G_\Delta = -\mathbb{E}\left(\frac{\partial h(x_t)}{\partial \theta_0} z_{\Delta t}'\right) = 0. \tag{10}$$

Partial relevance, conditional relevance, and redundancy

- Consider subset θ_A of $\theta = (\theta'_A \quad \theta'_B)'$.
  - Relevance of $f_{2T}(\theta)$ for the estimation of θ_A:

$$\boxed{G_{2A} \neq 0,} \tag{11}$$

  where $G_{2A} = \partial\,\mathbb{E}(f_{2T}(\theta_0))/\partial\theta'_A$.
  - Conditional relevance of $f_{2T}(\theta)$ given $f_{1T}(\theta)$ for the estimation of θ_A:

$$\boxed{G_{\Delta A} \neq 0,} \tag{12}$$

  where $G_{1A} = G_{1A} - W_{21}W_{11}^{-1}G_{2A}$.
  - Non-redundancy of $f_{2T}(\theta)$ given $f_{1T}(\theta)$ for the estimation of θ_A:

$$\boxed{G_{\Delta A}\Sigma_{1AA} + G_{\Delta B}\Sigma_{1BA} \neq 0,} \tag{13}$$

  where Σ_{1AA} and Σ_{1AA} are corresponding blocks of asymptotic covariance matrix $\Sigma_{\hat{\theta}}$ under $f_{1T}(\theta)$ only.

Test Statistics

- Relevance
  - Estimator of $\mathrm{vec}(G_2)$:

$$g_{2T}(\hat{\theta}) = \mathrm{vec}\left(G_{2T}(\hat{\theta})'\right). \tag{14}$$

  - Under $H_0 : G_2 = 0$ and certain assumptions

$$\sqrt{T} g_{2T}(\hat{\theta}) \to_d N(0, \Sigma_{\hat{g}_2}). \tag{15}$$

  - Test statistics:

$$\boxed{T g'_{2T}(\hat{\theta}) \Sigma_{\hat{g}_2}^{-1} g_{2T}(\hat{\theta}) \to_d \chi^2(km_2).} \tag{16}$$

- Similar derivations for
  - partial relevance,
  - conditional relevance (and non-redundancy under optimal weighting),
  - partial conditional relevance.

Monte Carlo Setting

- Design of Tauchen (1986)
- Consumption based CAPM

$$\mathbb{E}\left(\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}R_{t,t+1}-1\right)z_t\right)=0, \tag{17}$$

- Markov Chain to approximate VAR:

$$\begin{pmatrix} c_t \\ d_t \end{pmatrix} = 2\begin{pmatrix} 0.021 \\ 0.004 \end{pmatrix} + 2\begin{pmatrix} -0.161 & 0.017 \\ 0.004 & 0.117 \end{pmatrix}\begin{pmatrix} c_{t-1} \\ d_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{c,t} \\ \varepsilon_{d,t} \end{pmatrix}, \tag{18}$$

  - c_t is the logarithm of consumption growth rate,
  - d_t is the logarithm of real dividends growth,
  - $\mathrm{var}(\varepsilon_{c,t}) = 0.014$, $\mathrm{var}(\varepsilon_{d,t}) = 0.0012$, $\mathrm{corr}(\varepsilon_{c,t}, \varepsilon_{d,t}) = 0.43$.
- Then recover series C_{t+1}/C_t and $R_{t,t+1}$

Instrumental Variables

- Supposed to be (conditionally) relevant
  - A constant
  - Lagged consumption growth C_t / C_{t-1}
  - Lagged asset return $R_{t-1,t}$
- Supposed to be conditionally irrelevant
  - $Z_{1t} = \exp(0.5c_t + 0.5\varepsilon_{1,t})$
- Supposed to be irrelevant
  - $Z_{2t} = \exp(\varepsilon_{2,t})$

Rejection Frequencies (1)

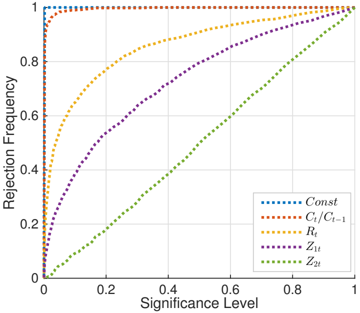

Figure 2. H_0:Irrelevance (conditional on a constant)

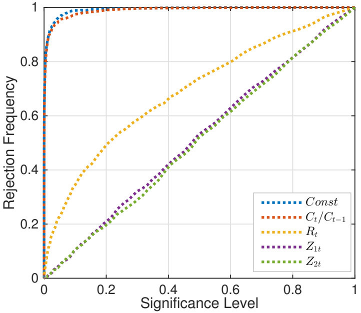

Figure 3. H_0 : Conditional irrelevance (redundancy)

Rejection Frequencies (2)

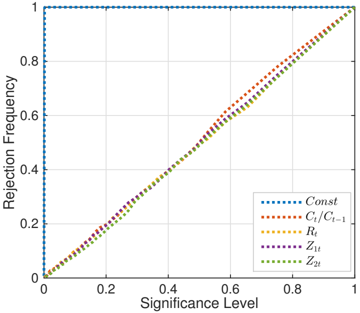

Figure 4. H_0 : Partial irrelevance for the identification of β (conditional on a constant)

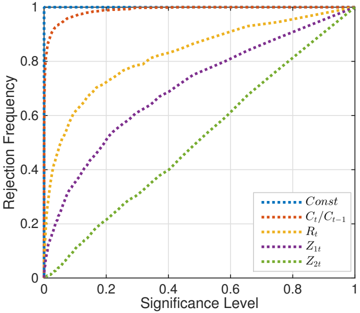

Figure 5. H_0 : Partial irrelevance for the identification of γ (conditional on a constant)

Rejection Frequencies (3)

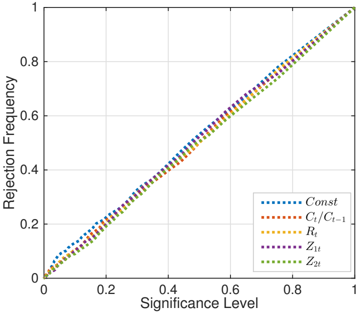

Figure 6. H_0 : Conditional partial irrelevance for the identification of β

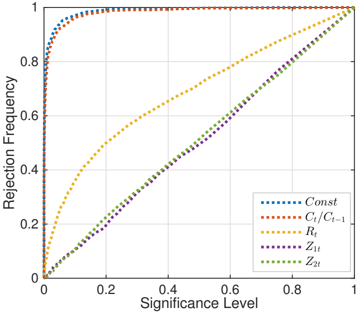

Figure 7. H_0 : Conditional partial irrelevance for the identification of γ

Case of HAC Covariance Estimator

- Convergence of HAC estimator $\Omega_T(\theta_0)$ of Ω is slower than $T^{-1/2}$.
- But $\Omega_T(\theta_0)$ is still asymptotically normal (Giraitis and Koul, 2012).
- Appropriate multiplier in the statistic for conditional relevance.
- Appropriate covariance matrix $\Sigma_{\hat{g_\Delta}}$ of $g'_{\Delta T}(\hat{\theta})$.
- Under $H_0 : G_2 = 0$ the statistic

$$\boxed{T^{2r} g'_{\Delta T}(\hat{\theta}) \Sigma^{-1}_{\hat{g}_\Delta} g_{\Delta T}(\hat{\theta}) \to_d \chi^2(km_2).} \tag{19}$$

where T^{-2r} is the rate of convergence of HAC estimator.

Case of the Lack of Identification

- Statistic is based on the idea of Wright (2003):

$$L^* = \inf_{\theta \in S^*_\theta(\alpha)} T g'_{2T}(\theta) \Sigma_{g_2}(\theta)^{-1} g_{2T}(\theta), \tag{20}$$

where $S^*_\theta(\alpha)$ is the confidence set of coverage $1 - \alpha$ robust to the lack of identification (Stock and Wright, 2000).

- Under $H_0 : G_2 = 0$ and certain assumptions

$$\boxed{\lim_{T \to \infty} \mathbb{P}\left(L^* \leq F_{\chi^2}(\alpha, km_2)\right) \leq 2\alpha.} \tag{21}$$

- This test is asymptotically conservative (true size can be smaller than significance level).
- Monte Carlo results are not ready yet.

Thank you for your attention!

Literature

BREUSCH, T., H. QIAN, P. SCHMIDT AND D. WYHOWSKI (1999). Redundancy of moment conditions. *Journal of Econometrics*, **91**, 89–111.

CHENG X., LIAO Z. (2015). Select the valid and relevant moments: An information-based lasso for gmm with many moments. *Journal of Econometrics*, **186(2)**, 443-464.

CRAGG, J. G. (1983). More Efficient Estimation in the Presence of Heteroscedasticity of Unknown Form. *Econometrica*, **51**, 751-764.

KLEIBERGEN, F. (2005). Testing parameters in GMM without assuming that they are identified. *Econometrica*, **73(4)**, 1103-1123.

STOCK, J. H., YOGO, M. (2005). Testing for weak instruments in linear IV regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg.*

STOCK, J., J. WRIGHT, AND M. YOGO (2002). GMM, weak instruments, and weak identification. *Journal of Business and Economic Statistics*, **20**, 518-530.

TAUCHEN, G. (1986). Statistical properties of generalized method-of-moments estimators of structural parameters obtained from financial market data. *Journal of Business and Economic Statistics*, **4(4)**, 397-416.

WRIGHT, J. H. (2003). Detecting lack of identification in GMM. *Econometric theory*, **19(2)**, 322-330.