International Econometric Review (IER)

A Pretest to Differentiate Between Weak and Nearly-Weak Instrument Asymptotics

Mehmet Caner®

North Carolina State University

ABSTRACT

We propose a pretest, bootstrap Kolmogorov-Smirnov test, to differentiate between weak and nearly-weak asymptotics. This is based on bootstrapping Wald Continuous Updating Estimator (CUE) based test. Since Wald CUE test has different limits under weak and nearly-weak cases this can be used in a pretest. We also conduct some simulations and show that some of the asset pricing models conform to nearly-weak asymptotics.

Key words: *Bootstrap, Kolmogorov-Smirnov Test*
JEL Classifications: C11, C20, C30

1. INTRODUCTION

In Generalized Method of Moments (GMM) and its Continuous Updated Estimator (CUE) version when the moment conditions at true value of the parameters are zero we deem that system as strongly identified. If the moment conditions decay to zero at square root T rate, then the system is called weakly identified by Stock and Wright (2000). Then, if the moment conditions decay to zero slower than root T rate, the system is called nearly weak by Hahn and Kuersteiner (2002). CUE is only different from GMM with its weight matrix. The parameters in the weight are optimized with sample moment functions in one step unlike GMM.

Stock and Wright (2000) in a seminal paper show that in the case of weak instruments, the GMM estimator is inconsistent. Furthermore, even in large samples, the Wald test is not pivotal. Hence we cannot benefit from standard asymptotic theory. An important assumption of Stock and Wright (2000) is the correlation between the instruments and the orthogonality restrictions decline at rate of square root T. So in large samples, identification of the parameters is not possible.

In an important recent paper; Hahn and Kuersteiner (2002) analyzed the “nearly-weak” instruments in a linear IV structure. In their setup, the correlation between instruments and the orthogonality restrictions decline at a slower rate than root T. This results in consistent estimates but slower rate of convergence than square root T with normal limits. They also carefully examine two-stage least squares estimators in higher order expansions. So there is a distinction between the weak and nearly-weak cases in terms of large sample theory. In the nearly-weak case the two-stage least squares estimators have the same limit as in the standard strong two stage-least squares estimators. In the weak case, the limit consists of several nuisance parameters and it is non normal distributed (Stock and Wright, 2000).

® Mehmet Caner, Department of Economics, Nelson Hall 4168, P.O. Box 8110, Raleigh, NC 27695, (email: mcaner@ncsu.edu)

Caner (2010) recently show that in the case of nearly-weak GMM, the GMM estimators are consistent and Wald tests based on CUE are asymptotically chi-square distributed. The only difference between the standard GMM asymptotics and the nearly-weak GMM is that the estimator converges to normal limits slower in the nearly-weak case. Then this brings us to a very important question. All the framework in the weak instruments case benefit from Anderson and Rubin (1949), and Kleibergen (2005) tests mainly. These tests have very limited power and most importantly only work with a grid search of the possible null values of the parameters. So testing multiple parameters is very problematic and also we cannot test anything else except the simple null. However, if the truth is not weak instruments, then we can still estimate the parameters consistently and conduct inference with Wald test. It is then essential to distinguish between nearly-weak and weak cases if we want to test restrictions.

Note that in the nearly-weak case, the limit is χ^2 for Wald test, unlike weak case. We propose a pretest for differentiating between nearly-weak and weak identification cases. This is a bootstrap version of Kolmogorov-Smirnov test. This test uses empirical distribution function of bootstrap Wald test in CUE form and compares it with χ^2 distribution. If the reality is nearly-weak asymptotics, we do not reject the null of nearly-weak asymptotics, otherwise (weak identification) there is a large discrepancy and we reject the null. Using bootstrap to understand whether the finite sample behavior accords with limit theory is proposed in GMM context by Hall and Horowitz (1996).

We conduct a simulation exercise and show that some of the asset pricing models may be explained by nearly-weak asymptotics. We benefit from Wald CUE test since this test has different limits under weak and nearly-weak moments cases.

Section 2 provides the model and assumptions. Section 3 conducts some simulations and introduces a simple pretest. Section 4 concludes.

2. THE MODEL AND ASSUMPTIONS

We benefit largely from the framework of Stock and Wright (2000). Let θ be a p-dimensional parameter vector, and θ_0 represents the population value which is in the interior of the compact parameter space $\boldsymbol{\Theta}\subset\boldsymbol{R}^p$. The population orthogonality conditions are of G dimension:

$$E[\psi(Y_t,\theta_0)] = 0 \qquad (2.1)$$

The data is $\{(Y_t) : t = 1,2,\ldots,T\}$.

If moment is exactly zero in finite samples at only the true value, strong identification holds. If moment declines to zero at rate $T^{1/2}$ then weak identification holds. If it declines to zero at a slower rate T^{κ}, $0 < \kappa < 1/2$, then nearly weak identification holds.

Definition 2.1. If
(i) (Weak Identification) the moment conditions follow

$$ET^{-1}\sum_{t=1}^{T}\psi_t(\theta) = \frac{m_{1T}(\theta_0)}{T^{1/2}}$$

then the parameters θ are weakly identified.
(ii) (Nearly-Weak Identification) the moment conditions follow, $0 \le \kappa < 1/2$,

$$ET^{-1}\sum_{t=1}^{T}\psi_t(\theta) = \frac{m_{1T}(\theta_0)}{T^{\kappa}}$$

then the parameters θ are near-weakly identified.

For the near-weakly identified parameters we see that moment function again decays to zero but at a slower rate of T^{κ} than $T^{1/2}$ of Stock and Wright (2000), $0 < \kappa < 1/2$. We consider basically a system with nearly-weak/strong instruments. Note that when $\kappa = 0$, the parameters are strongly identified.

We use the term "nearly-weak" instruments for $0 \leq \kappa < 1/2$. The term "nearly-weak" is introduced by Hahn and Kuersteiner (2002).

Now we can supply the assumptions.

Assumption 2.1. For $0 \leq \kappa < 1/2$;

$$ET^{-1}\sum_{t=1}^{T}\psi_t(\theta) = \frac{m_{1T}(\theta)}{T^{\kappa}}$$

$m_{1T}(\theta) \rightarrow m_1(\theta)$ uniformly in θ, $m_1(\theta)=0$ if and only if $\theta = \theta_0$. $m_1(\theta)$ is continuous. Furthermore, $\psi_t(.)$ is continuously differentiable in θ in N, a neighborhood of θ_0, and

$$\frac{T^{\kappa}}{T}\sum_{t=1}^{T}\frac{\partial\psi_t(\theta)}{\partial\theta'} \xrightarrow{p} R(\theta)$$

uniformly in N, $R(\theta_0)$ is of full column rank. $R(\theta)$ is $G\times p$ matrix. Note that $R(\theta)$ is continuous.

Assumption 2.2.
(i). $\psi_t(\theta)$ is m-dependent.
(ii).

$$|\psi_t(\theta_1) - \psi_t(\theta_2)| \leq B_t|\theta_1 - \theta_2|$$

where $lim_{T\rightarrow\infty}\ T^{-1}\Sigma_{t=1}^{T}EB_t^{2+\delta} < \infty$, for some $\delta > 0$.
(iii).

$$\sup_{\theta\in\Theta} E|\psi_t(\theta)|^{2+\delta} < \infty$$

for some $\delta > \infty$.

Define the empirical process as

$$\Psi_T(\theta) = T^{-1/2}\sum_{t=1}^{T}\psi_t(\theta) - E\psi(\theta) \tag{2.2}$$

Under Assumption 2.2 we obtain the following result via Andrews (1994)

$$\Psi_T(\theta) \Rightarrow \Psi(\theta) \tag{2.3}$$

where $\Psi(\theta)$ is a Gaussian stochastic process on Θ with mean zero and covariance function $E\Psi(\theta_1)\Psi(\theta_2)' = \Omega_{\theta_1,\theta_2}$.

Both Antoine and Renault (2007), and Caner (2010), independently show that in nearly-weak instruments case, the estimators are consistent. Furthermore, in a case of all nearly-weak instruments the limit is asymptotically normal. Nearly-weak case combines both the elements of strong and weak cases. Loosely speaking, like the weak instruments case the correlation between instruments and the orthogonality conditions goes to zero in large samples. However, since this is not decaying as fast as the weak instruments case we still have consistency as strong instruments case.

Assumption 2.3. Let

$$\hat{V}_{\theta,\theta}(\theta) = T^{-1}\sum_{t=1}^{T}[\psi_t(\theta) - \overline{\psi}_t(\theta)][\psi_t(\theta) - \overline{\psi}_t(\theta)]' \quad (2.4)$$

Then uniformly in θ

$$\hat{V}_{\theta,\theta}(\theta) \Rightarrow \Omega_{\theta,\theta}$$

For efficient CUE the weight matrix is: $W_T(\overline{\theta}(\theta)) = \hat{V}_{\theta,\theta}(\theta)^{-1}$. Variance covariance matrix estimators and result (3) at θ_0, can be extended to more general weakly dependent data.

Definition 2.2. The efficient CUE estimator $\hat{\theta}$ minimizes the following over Θ

$$S_T(\theta,\theta) = \left[T^{-1/2}\sum_{t=1}^{T}\psi_t(\theta)\right]'\left[\hat{V}_{\theta,\theta}(\theta)\right]^{-1}\left[T^{-1/2}\sum_{t=1}^{T}\psi_t(\theta)\right]$$

We define Wald test in CUE case. First we test the following null:

$$H_0: a(\theta_0) = 0,$$

where $a(.)$ is $r\times 1$ dimensional.

We need the following variant of Assumption 2.1 for partial derivative matrix estimation and identification in GMM estimates.

The following puts some structure on restrictions.

Assumption 2.4. $a(\theta)$ is continuously differentiable with $A(\theta) = \partial a(\theta)/\partial\theta'$, $A(\theta_0)$ is of full rank r, where $A(\theta_0) = A$ is $r\times p$. First we give theory for the Wald test.

Definition 2.3. The Wald test at efficient CUE estimator $\hat{\theta}$ is defined as follows

$$Wald_{CUE} = Ta(\hat{\theta})'\left[A(\hat{\theta})\left\{\left(\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\psi_t(\hat{\theta})}{\partial\theta'}\right)'\left[\hat{V}_{\theta,\theta}(\hat{\theta})\right]^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\psi_t(\hat{\theta})}{\partial\theta'}\right)\right\}^{-1}A(\hat{\theta})'\right]^{-1}a(\hat{\theta})$$

$V_T(\hat{\theta})$ represents the consistent estimate of the variance-covariance matrix in efficient CUE and can be seen in equation (2.4).

Corollary 1 of Caner (2010) shows under Assumptions 2.1-2.2

$$T^{1/2-\kappa}(\hat{\theta}-\theta_0) \xrightarrow{d} N(0,\left[R(\theta_0)'\Omega_{\theta_0,\theta_0}^{-1}R(\theta_0)\right]^{-1})$$

Then under Assumptions 2.1-2.4, Theorem 4 of Caner (2010) shows that

$$Wald \xrightarrow{d} \chi_r^2$$

Note that in the case of weak instruments, Theorem 1 and Corollary 4 of Stock and Wright (2000) shows that GMM estimators are inconsistent, and Wald test is not pivotal, and we cannot tabulate the critical values in any possible way. Note that the limits in Stock and

Wright (2000) are complicated and full of nuisance parameters and non-normal. For details, please see Theorem 1 of Stock and Wright (2000)

Both Antoine and Renault (2007), and Caner (2010) show that the Wald test converges in distribution to χ^2_r distribution under Assumptions 2.1-2.4.

3. PRETEST AND SIMULATIONS

This section proposes a pretest to differentiate between nearly-weak and weak cases, and conducts some simulation exercises. We show that we can better approximate finite sample properties in asset pricing models with nearly-weak instruments asymptotics. We explain the reasons behind this at the end of simulation exercise. In this respect, we consider the following model with constant relative risk aversion (CRRA) preferences

$$\psi_t(\theta) = \left[d\left(\frac{C_{t+1}}{C_t}\right)^{-f} R_{t+1} - \iota_{G_1} \right] \otimes Z_t$$

where $\theta = (d, f)'$, d is a time discount and f is the risk aversion parameter. Z_t represents the instruments a $G_2\times1$ vector. R_{t+1} is a $G_1\times1$ vector of asset returns. C_t is consumption at time t, ι_{G_1} is $G_1\times1$ vector of ones, $G_1 * G_2 = G$. The parameters are assumed to be bounded by $d_{\min} \le d \le d_{max}$, $f_{\min} \le d \le f_{max}$. This setup is both used in Stock and Wright (2000), and Kleibergen (2005). Unlike Stock and Wright (2000) who use two-step GMM or coefficient estimates in CUE and their empirical cumulative distribution functions we use Wald test in CUE format, W_{CUE}. We specifically test the null of $\theta = \theta_0$. χ^2 asymptotics is shown to be working well for the small sample behavior of W_{CUE} in spite of low correlation between the instruments and the moment equations (Hansen et.al, 1996, Figures 7-9). In the same figures we see that finite sample behavior of Wald test in two-step GMM format are not well approximated by χ^2 asymptotics.

We now want to see whether the assumptions for W_{CUE} in testing are satisfied in our case. These are Assumptions 2.1, 2.2, 2.3 and 2.4. First, (C_{t+1}/C_t), R_{t+1}, Z_t are m-dependent by design that is described below. Assumption 2.2(i) is then satisfied. Then $E(R_{t+1}\otimes Z_t)^5 < \infty$, and $E(exp\{5(f_{max} + 1)|c_{t+1}|\}) < \infty$, where $c_{t+1} = ln(C_{t+1}/C_t)$. These imply Assumption 2.2(ii), 2.2(iii). Assumption 2.3 is satisfied if (C_{t+1}/C_t), R_{t+1}, Z_t have enough moments. Assumption 2.4 is satisfied in our case since we test H_0: $\theta = \theta_0$. The remaining Assumption 2.1 is arbitrary, but the key point there is whether $(C_{t+1}/C_t)^{-f}$ has a low correlation with instruments. Again by the simulation design there is low correlation, this can be seen in (3.5). There is low correlation between the consumption ratio and their lagged values. This is a well known simulation design in this literature. We can deem "f" as nearly weakly or weakly identified and "d" as strongly identified. Unlike Stock and Wright (2000) we set "f" as possible nearly-weak or weak rather than weak. We can see in simulation exercises which scenario will be more plausible for "f" weak or nearly-weak. Now we explain the specifics in designs. The errors are martingale difference sequences at true values so correction for autocorrelation is not used. There are no overlapping data as well. The designs in the Monte Carlo are due to Hansen et al. (1996). Four designs are described in Table 3.1. Their method fits a VAR(1) to approximate consumption and dividend growth. Let c_t be the log growth rate of US per capita real annual consumption and d_t the log growth rate of real annual dividends on the S&P 500. This is given by

$$\begin{pmatrix} c_t \\ d_t \end{pmatrix} = \begin{pmatrix} 0.021 \\ 0.04 \end{pmatrix} + \begin{pmatrix} -0.161 & 0.017 \\ 0.414 & 0.117 \end{pmatrix} \begin{pmatrix} c_{t-1} \\ d_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ct} \\ \varepsilon_{dt} \end{pmatrix} \qquad (3.5)$$

where *var* ε_{dt} = 0.014, *var* ε_{ct} = 0.0012, *cov*($\varepsilon_{ct},\varepsilon_{dt}$) = 0.00177. This is the setup for Table 3.1.

Design	γ_0	δ_0	Assets	Instruments
1	1.3	0.97	r_t^s	1, r_{t-1}^s, c_{t-1}
2	13.7	1.139	r_t^s	1, r_{t-1}^s, c_{t-1}
3	1.3	0.97	r_t^f, r_t^s	1, c_{t-1}
4	1.3	0.97	r_t^f, r_t^s	1, r_{t-1}^s, r_{t-1}^f, c_{t-1}

Table 3.1 Monte Carlo Design.
Notes: $c_t = ln(Ct/C_{t-1})$, r_t^f, r_t^s represent consumption growth, the risk free rate, and the stock returns respectively. The program for Monte Carlo study can be obtained from the author on request.

We conduct the following simulation. First, we compute actual *p*-values of W_{CUE} under these four designs and whether they are close to their nominal level (for 10%, 5%). Various nominal levels are also analyzed but since 10% and 5% gave the worst results we display them. We also calculate Kolmogorov-Smirnov test and look at the maximal difference between the empirical distribution function of W_{CUE} and asymptotic distribution of χ_2^2. Then we also consider bootstrap *p*-values of W_{CUE} and see how they differ from nominal levels. Both these analysis give ideas to us whether these four designs come from a nearly-weak case or weakly identified one. If these actual, and bootstrap *p*-values are near theoretical distribution (χ_2^2) and Kolmogorov-Smirnov test statistic shows small discrepancy between the large sample distribution in nearly-weak case and empirical distribution function of W_{CUE}, then these designs have nearly-weak identification problem rather than weak identification. Since in the case of weak instruments asymptotics, the limit of W_{CUE} is not χ_2^2 and not pivotal we expect large differences between actual, bootstrap *p*-values and nearly weak asymptotic limit (χ_2^2) if the truth is indeed the weak instrument asymptotics. As suggested by Phillips and Park (1988), and Hall and Horowitz (1996) we use bootstrap approximation to critical values as a tool for analyzing the accuracy of asymptotic approximations. Note that if the asymptotic approximation is correct, the bootstrap should not deviate from asymptotics that much in this case. Bootstrap in regular GMM is consistent as shown in Hall and Horowitz (1996). CUE in nearly-weak case has the same limits as strong case, we should expect bootstrap here to estimate consistently χ_2^2 if the truth is nearly-weak identification.

We introduce a bootstrap version of Kolmogorov-Smirnov test. This is new and can be used as a pretest. In this way we can compare bootstrap empirical distribution function with asymptotics (χ_2^2). If these are very different from each other then we do have weak identification. Otherwise we have nearly-weak or standard asymptotics.

Now we give details of the simulation and the pretest idea that we propose. In the simulations we set T = 100. For obtaining the actual *p*-values for W_{CUE} we run 1000 iterations for each given design in Table 3.1. Then substitute χ_2^2 critical values (upper 10%, 5%) into the empirical distribution found from 1000 iterations, the record the number of rejections. Kolmogorov-Smirnov test is also found by comparing empirical distribution function of W_{CUE} with χ_2^2 distribution.

In the case of bootstrap exercise, we use bootstrap critical values rather than χ_2^2 ones. We calculate bootstrap critical values by using 1000 bootstrap iterations (bootstrap sample), then compare actual W_{CUE} from the original sample with 90% and 95% percentiles of the bootstrap empirical distribution function and then record whether we reject it or not. Then iterate this process 1000 times and sum the number of rejections at each level and calculate that as a percentage. To get the bootstrap empirical distribution function in the first step we withdraw a bootstrap sample from the original sample with replacement and then compute CUE estimator from this bootstrap sample $\hat{\theta}_{bc}$ and compute W_{CUE} at $\hat{\theta}_{bc}$. Repeat this 1000 times and sort it to get the bootstrap empirical distribution function. We follow the advice in Hall and Horowitz (1996) and demean each bootstrap sample moment. The Wald test that uses bootstrap sample is, where $\hat{\theta}$ is the efficient CUE estimator

$$W_{CUE}(\hat{\theta}_{bc}) = T(\hat{\theta}_{bc} - \hat{\theta})' \left[\left(\frac{1}{T} \hat{D}_T(\hat{\theta}_{bc}) \right)' \left[\hat{V}_{\theta,\theta}(\hat{\theta}) \right]^{-1} \left(\frac{1}{T} \hat{D}_T(\hat{\theta}_{bc}) \right) \right] (\hat{\theta}_{bc} - \hat{\theta})$$

where

$$\hat{D}_T(\hat{\theta}_{bc}) = \sum_{t=1}^{T} \frac{\partial \psi_t(\hat{\theta}_{bc})}{\partial \theta'}$$

The results of these exercises are in Table 3.2.

The pretest bootstrap Kolmogorov-Smirnov test is calculated in a way that we can use in an empirical study.

From any given data set, have 1000 bootstrap iterations and obtain the bootstrap empirical distribution function and compare with any asymptotic limit that we want to measure against.

$$BKS = T^{1/2} \sup_x |F_{nb}(x) - F(x)|$$

where F_{nb} represents the cumulative distribution function of bootstrap and F is the cumulative distribution function of chi-square distribution with 2 degrees of freedom.

Here we again benefit from designs 1-4, we simulate each design once and then have 1000 bootstrap iterations and measure against χ^2_2. The results are in Table 3.3.

This reflects the characteristic of an empirical study. We use again bootstrap empirical distribution function of W_{CUE}. This is done because it has better finite sample properties than say Wald two-stage least squares estimator and its limit is different in the case of nearly-weak and weak cases. So this approach can differentiate between two cases. We can not use Anderson and Rubin (1949), and Kleibergen (2005) type of tests since their limits are the same regardless of nearly-weak and weak cases. This is observed in Antoine and Renault (2007), and Caner (2010).

	Actual *p*-values		Bootstrap *p*-value		KS
	10%	5%	10%	5%	
D1	3.1	3.0	9.9	6.3	7.7
D2	70.8	69.5	54.0	52.2	6.5
D3	83.2	80.8	36.9	33.7	7.6
D4	97.6	96.9	35.9	35.3	9.2

Table 3.2 Analysis of Designs 1-4.
Notes: D1-D4 represent four designs that we use in Table 3.1. KS is the Kolmogorov-Smirnov test.

As can be seen from Table 3.2, Design 1 is more in line with χ^2_2 asymptotics, however Design 2 clearly shows that it is not coming from χ^2_2. We see large discrepancies in the case of Design 2 between bootstrap *p*-values and actual *p*-values from their nominal levels. Kolmogorov-Smirnov test also points that χ^2_2 asymptotics is not doing a good job in approximating finite sample distribution of simulations from Design 2. Designs 3 and 4 use more orthogonality restrictions than Designs 1 and 2, and it is clear from Table 3.2 that their behavior is different than χ^2_2 asymptotics.

Designs	KS test
D1	2.8
D2	9.5
D3	9.3
D4	9.5

Table 3.3 Pretest with Bootstrap KS test.
Notes: KS represents Kolmogorov-Smirnov test. In this case we compare bootstrap empirical distribution function with asymptotic limit.

Table 3.3, compares bootstrap empirical distribution function with χ_2^2 distribution. The same null is tested with W_{CUE}. This is not a regular simulation as in Table 3.2. We only select one draw from each design and then try to understand whether the data conforms to the limit. To compare the results in Table 3.3 with critical value of KS test at 5% level (1.36), (Mood et al., 1974:511). All designs in Table 3.3, seem to follow weak instrument asymptotics, but the bootstrap *p*-value for Design 1 is very close to true values for a χ_2^2 distribution in Table 3.2. So we think that except from Design 1, looking at Tables 3.2 and 3.3 jointly, Designs 2-4 conform to weak instrument asymptotics.

Note that Stock and Wright (2000) show that Designs 1-4 belong to weak instrument asymptotics. However, in the case of Design 1 we see that using two stage least squares and just analyzing estimators (not a specific test) might have caused Stock and Wright (2000) to think that they also could have come from weak instrument asymptotics. Even though, the instruments have low correlation with moment equations, in Design 1 we see that it can be explained by nearly-weak identification rather than weak identification. Simulating weak instrument asymptotics limits are not a good idea, since these contain nuisance parameters, and may give a different results in different simulations.

4. CONCLUSION

We propose a bootstrap pretest to differentiate between nearly-weak and weak identification. This seems to work in certain asset pricing examples in our simulations.

REFERENCES

Anderson, T.W. and H. Rubin (1949). Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Annals of Mathematical Statistics*, 20, 46-63.

Andrews, D.W.K. (1994). Empirical Process Methods in Econometrics. *Handbook of Econometrics*, 4, 2247-2294.

Antoine, B. and E. Renault (2007). Efficient GMM with nearly-weak identification. *Working paper*. Department of Economics, University of North Carolina-Chapel Hill.

Caner, M. (2010). Testing, Estimation in GMM and CUE with Nearly-Weak Identification. *Econometric Reviews*, 29, 330-363.

Hahn, J. and G. Kuersteiner (2002). Discontinuities of Weak Instrument Limiting Distributions. *Economics Letters*, 75, 325-331.

Hall, P. and J.L. Horowitz (1996). Bootstrap critical values for tests based on generalized method of moments estimators. *Econometrica*, 64, 891-916.

Hansen, L.P., J. Heaton and A. Yaron (1996). Finite Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics*, 14, 262-280.

Kleibergen, F. (2005). Testing Parameters in GMM Without Assuming That They Are Identified. *Econometrica*, 73, 1103-1124.

Mood, A.M., F.A. Graybill and D.C. Boes (1974). *Introduction to the Theory of Statistics*. New-York: Mc-Graw Hill.

Phillips, P.C.B. and J.Y. Park (1988). On the Formulation of Wald Tests of Nonlinear Restrictions. *Econometrica*, 56, 1065-1083.

Stock, J.H. and J.H. Wright (2000). GMM with Weak Identification. *Econometrica*, 68, 1055-1096.