E C O N O M I C S B U L L E T I N

Which null hypothesis do overidentification restrictions actually test?

Rembert De Blander
EcRu, U.C.Louvain and CES, K.U.Leuven

Abstract

In this note I investigate which alternatives are detected by over-identifying restrictions tests, when we test the null hypothesis that all excluded instruments are exogenous. A reformulated null hypothesis will result in a consistent test.

I thank Frédéric Gaspart, Steven Groenez and Kristiana Hansen for helpful comments and suggestions. The usual disclaimer on remaining errors applies.
Citation: De Blander, Rembert, (2008) "Which null hypothesis do overidentification restrictions actually test?." *Economics Bulletin*, Vol. 3, No. 76 pp. 1-9
Submitted: October 24, 2008. **Accepted:** December 8, 2008.
URL: http://economicsbulletin.vanderbilt.edu/2008/volume3/EB-08C20068A.pdf

1. Introduction

It is well-known that over-identifying restrictions tests are not consistent against all alternatives when they are used to test the null hypothesis, say H_0, that all excluded instruments are exogenous (see, for instance, Newey, 1985). Here we use the standard notion of test consistency whereby a test is *consistent* (of significance level α) *against some alternative* if it is asymptotically of level α and the power of the test at this (fixed) alternative goes to one with increasing sample size, say N. This test is said to be *consistent* if it is consistent against all fixed alternatives. See for example van der Vaart (2000, p.193).

An inconsistent test, thus, has alternatives which it does not detect with power going to one for $N \to \infty$. We will show that, in the case of over-identifying restrictions tests, this inconsistency could be termed a *blind spot*, in the sense that certain alternatives, no matter how far from H_0, will not increase the power of the test statistic. This is unfortunate, since a "failure to reject" results are inherently ambiguous. In addition, results of the "failure to reject" kind can be considered stronger, the broader the class of alternatives is with respect to which one fails to reject. As a consequence, consistent tests are to be preferred over inconsistent ones.

One way to deal with an inconsistent test, is to add extra, a priori, assumptions which make the test consistent. In the instrumental variable context it is common to assume the presence of as much exogenous instruments as there are endogenous variables in the outcome equation. Instead of the latter strategy, one could try to identify the set of parameter combinations, say S_0, for which the considered test statistic has the same asymptotic distribution. It is possible that the null hypothesis of an inconsistent test does not coincide with S_0. In such a case, a reformulation of the null hypothesis (as G_0, say) is in order. If we subsequently show that the power of the test goes to one for $N \to \infty$ at every fixed element of G_A (the reformulated alternative hypothesis), this simple reformulation of the null hypothesis is enough to result in a consistent test. This course of action has the added advantage (over adding extra assumptions) that fundamentally untestable alternatives (with respect to the initial null H_0) are are made explicit.

In the context of the over-identifying restrictions tests considered here, the new null G_0 will be expressed in terms of those parameters which seem most natural: those of the structural outcome equation and those of the reduced form for the treatment variables. In this context, a test of G_0 versus G_A still can be interpreted in terms of over-identification (exclusion of the instruments from the structural outcome equation).

2. Framework

Consider the $N \times 1$ stacked observations of the outcome y_i to be given by

$$Y \quad = \quad X\alpha_x + Z\alpha_z + E, \tag{1}$$

where the $N \times K_1$ matrix X contains exogenous variables x_i, possibly including a constant term, and E stacks the N error terms ε_i. The $N \times L$ matrix Z contains the treatment, choice or otherwise endogenous variables, which can be represented by the reduced form

$$Z \quad = \quad XB_x + WB_w + V, \tag{2}$$

where the $N \times K_2$ matrix W fulfills $K_2 \geq L$ and $\text{rank}\left(\mathbf{E}\left[W'Z\right]\right) = L$. These familiar identification restrictions allow the use of W as a set of instruments for Z.

Remark. The first-stage parameters of the excluded instruments, i.e. B_w, will play a crucial role in the reformulation of the null hypothesis.

In the remainder of the text, the more compact notation $R = (X, Z)$ is used to denote the explanatory variables in the outcome equation (1) and $S = (X, W)$ for the exogenous variables in (2). Furthermore, following assumptions[1] are maintained.

Assumption 1. The observations (rows) of (S, E, V) are *IID*.

Assumption 2. The outcome disturbances E are mean-independent of S, i.e.

$$\mathbf{E}\left[E \mid S\right] \;=\; 0.$$

Assumption 2 is crucial for consistency of the IV estimator

$$\hat{\alpha}_{IV} \;=\; \left(R'S\left(S'S\right)^{-1}S'R\right)^{-1}\left(R'S\left(S'S\right)^{-1}S'Y\right).$$

While $\mathbf{E}\left[E \mid X\right] = 0$ is untestable within this framework due to the imposed orthogonality between regressors and residuals by all LS estimators, the assumption $\mathbf{E}\left[E \mid W\right] = 0$ is (partly) testable. The latter tests are commonly called *over-identifying restrictions tests* or *instrument validity tests* and will be treated in the following section.

3. Some over-identifying restrictions tests

Two well-known such tests are the Sargan (1958)-Hansen (1982) J-test and the LM or score test as described by Magdalinos (1988). The J-test uses the minimized optimal GMM criterion as test statistic, which is, under homoskedasticity, given by

$$J \;=\; \left(\hat{E}'S\right)\cdot\left(S'S\right)^{-1}\cdot\left(S'\hat{E}\right)/\widehat{\sigma_{\varepsilon}^{2}}, \tag{3}$$

with $\sigma_{\varepsilon}^{2} = \mathbf{Var}\left[\varepsilon\right]$, the error variance of the outcome equation (1). Under the null that all K_2 instruments are valid, the minimized optimal GMM criterion is asymptotically distributed as a central chi-square distribution with $K_2 - L$ degrees of freedom. The LM-statistic to test the null hypothesis that all instruments are exogenous is an omitted variable test for the variables W in the outcome equation (1)[2]. It is given by

$$LM \;=\; N^{-1}\left(\hat{E}'W\right)\hat{\Omega}^{-1}\left(W'\hat{E}\right) \tag{4}$$

where Ω is the covariance matrix of $N^{-1}E'\hat{W}$. Under the hypothesis that the instruments W are validly restricted from the outcome equation (1), it holds that $LM \xrightarrow{d} \chi^2_{K_2-L}(0)$, as well.

[1] Throughout this note it is implicitly assumed that the instruments are not weak. See, for instance, Shea (1997) for a way to detect weak instruments.

[2] Under homoskedasticity, above LM test is identical to Wooldridge's (1990) procedure to test for omitted variables in the outcome equation, where the variables under test here as being omitted are the excluded instruments.

All such validity tests have one weakness in common: they are inconsistent when used for testing the null hypothesis $H_0 : \mathbf{E}[E \mid W] = 0$. This means that, even for growing sample size N, not all departures from H_0 are detected[3]. In other words, these tests have a *blind spot*. Some alternatives, no matter how far from H_0, will leave the test statistics unaffected. Sometimes correct application of over-identifying restrictions tests is stated as follows:

> If at least L instruments are exogenous, then it is possible to test the null hypothesis that all instruments are exogenous against the alternative that at least one (but no more than $K_2 - L$) is endogenous (Stock 2001).

Such statements, by adding the extra assumptions that "at least L instruments are exogenous", produce a consistent test in the sense that now all alternatives (falling within the added assumptions) are detected with power going to one for growing sample size. However, they incorrectly characterize the class of alternatives for which over-identifying restrictions tests are inconsistent if we do not wish to make the assumption that "at least L instruments are exogenous". Furthermore, they discourage the use of IV estimators, since they suggest that their use is somehow founded on the *untestable belief* in the presence of at least L exogenous instruments. In the next section, I will characterize the set of alternatives to H_0, for which over-identifying restrictions tests, *as currently formulated*, are inconsistent. This will then lead to a reformulated null hypothesis, G_0, which, in combination with either of the test statistics (3) or (4) above, will constitute a consistent test.

4. Inconsistency

Consider again the outcome equation (1). Under the alternative hypothesis that not all instruments are exogenous, we can write the error term as

$$
\begin{aligned}
E &= W\kappa + U && (5)\\
\mathbf{E}[U \mid W] &= \mathbf{0}_{N\times 1}. && (6)
\end{aligned}
$$

The $L \times 1$ vector κ the indicates which linear combination of the instruments the error term E is maximally correlated with, i.e. $\kappa = \mathbf{E}[W'W]^{-1}\,\mathbf{E}[W'E]$. The null hypothesis $H_0 : \mathbf{E}[E \mid W] = \mathbf{0}_{N\times 1}$ can be reformulated as $H_0 : \kappa = \mathbf{0}_{K_2\times 1}$. The following theorem characterizes the set of alternatives to H_0, for which over-identifying restrictions tests have power equal to size and the set of alternatives for which said test is consistent.

Theorem. Over-identifying restrictions tests, consisting of test statistics (3) or (4) and null hypothesis $H_0 : \kappa = \mathbf{0}_{K_2\times 1}$, have asymptotic power equal to size against alternatives of the form $E = (WB_w)\,\gamma + U$, with $\gamma \in \mathbb{R}^L$. They are consistent against alternatives of the form $W\delta$, for which $B_w'\Sigma_w\delta = \mathbf{0}_{L\times 1}$, where $\Sigma_w = \mathbf{Var}[w]$.

Proof. Consider model (1)-(2) where the covariates X are partialed out for ease of exposition (and without loss of generality)

$$
\begin{aligned}
\tilde{Y} &= \tilde{Z}\alpha_z + \tilde{E} && (7)\\
\tilde{Z} &= \tilde{W}B_w + \tilde{V}, && (8)
\end{aligned}
$$

[3] This observation is well-known, see Newey (1985), for instance.

i.e. where $\tilde{Y}$ denotes[4] $\mathcal{Q}_X Y$. Under $H_A : \mathbf{E}\left[\tilde{E} \mid \tilde{W}\right] \neq 0$, we can write (7) as

$$\tilde{Y} \;=\; \tilde{Z}\alpha_z + \tilde{W}\kappa + \tilde{U}, \tag{9}$$

with $\mathbf{E}\left[\tilde{U} \mid \tilde{W}\right] = 0$ by (6). Now, *IV* of $\tilde{Y}$ on $\tilde{Z}$ using $\tilde{W}$ as instruments is identical to *OLS* of $\mathcal{P}_{\tilde{W}}\tilde{Y}$ on $\mathcal{P}_{\tilde{W}}\tilde{Z}$. To investigate its properties rewrite (9) as

$$\mathcal{P}_{\tilde{W}}\tilde{Y} \;=\; \mathcal{P}_{\tilde{W}}\tilde{Z}\alpha_z + \mathcal{P}_{\tilde{W}}\tilde{W}\kappa + \mathcal{P}_{\tilde{W}}\tilde{U}.$$

Defining

$$\begin{aligned}\hat{\tilde{Z}} &= \mathcal{P}_{\tilde{W}}\tilde{Z} \\ &= \tilde{W}\hat{B}_w,\end{aligned} \tag{10}$$

where $\hat{B}_w = \left(\tilde{W}'\tilde{W}\right)^{-1}\tilde{W}'\tilde{Z}$ is the *OLS* estimator of B_w, we have that

$$\begin{aligned}\mathcal{P}_{\tilde{W}}\tilde{Y} &= \mathcal{P}_{\tilde{W}}\tilde{Z}\alpha_z + \mathcal{P}_{\tilde{W}}\tilde{W}\kappa + \mathcal{P}_{\tilde{W}}\tilde{U} \\ &= \hat{\tilde{Z}}\alpha_z + \tilde{W}\kappa + \mathcal{P}_{\tilde{W}}\tilde{U} \\ &= \hat{\tilde{Z}}\alpha_z + \mathcal{P}_{\hat{\tilde{Z}}}\tilde{W}\kappa + \mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W}\kappa + \mathcal{P}_{\tilde{W}}\tilde{U},\end{aligned}$$

where the first equality follows from (10) and the identity $\tilde{W} = \mathcal{P}_{\tilde{W}}\tilde{W}$ and the last equality follows from the identity $A = \mathcal{P}_B A + \mathcal{Q}_B A$, which holds for all conformable matrices A and B. We have in general that $\mathcal{P}_{\hat{\tilde{Z}}}\tilde{W} = \hat{\tilde{Z}}\Lambda$ for some matrix $\Lambda \in \mathbb{R}^{L\times K_2}$ and, thus,

$$\mathcal{P}_{\tilde{W}}\tilde{Y} \;=\; \hat{\tilde{Z}}\left(\alpha_z + \Lambda\kappa\right) + \mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W}\kappa + \mathcal{P}_{\tilde{W}}\tilde{U}. \tag{11}$$

Now, *IV* of $\tilde{Y}$ on $\tilde{Z}$ using $\tilde{W}$ as instruments, results in $\hat{\alpha}_{z;IV} = \left(\hat{\tilde{Z}}'\hat{\tilde{Z}}\right)^{-1}\left(\hat{\tilde{Z}}'\mathcal{P}_{\tilde{W}}\tilde{Y}\right)$. Expression (11) informs us that

1. α_z is not identified unless $\Lambda = \mathbf{0}_{L\times K_2}$ or $\kappa = \mathbf{0}_{K_2\times 1}$.

2. $\mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W}\kappa$ will be part of the residual, since $\hat{\tilde{Z}}'\mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W} = \mathbf{0}_{L\times K_2}$. Thus, unless $\mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W}\kappa = \mathbf{0}_{N\times 1}$, tests based on $\tilde{W}'\hat{\tilde{E}}$, with $\hat{\tilde{E}} = \tilde{Y} - \tilde{Z}\hat{\alpha}_{z;IV}$ the stacked *IV* residuals, will pick up departures from H_0.

We will now determine under which conditions it holds that $\mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W}\kappa = \mathbf{0}_{N\times 1}$, i.e. for which alternatives to H_0 the power of the over-identifying restrictions statistics (3) and (4) remains equal to their size.

- Given the definition (10), we have that $\mathcal{P}_{\hat{\tilde{Z}}}\tilde{W}\hat{B}_w = \mathcal{P}_{\hat{\tilde{Z}}}\hat{\tilde{Z}} = \hat{\tilde{Z}} = \tilde{W}\hat{B}_w$, which shows that the row space of $\mathcal{P}_{\hat{\tilde{Z}}}$ is given by $\tilde{W}\hat{B}_w\gamma$, with $\gamma \in \mathbb{R}^L$. Since the row space of $\mathcal{P}_{\hat{\tilde{Z}}}$ is identical to the null space of $\mathcal{Q}_{\hat{\tilde{Z}}}$, we have that

$$\mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W}\hat{B}_w\gamma \;=\; \mathbf{0}_{N\times 1}.$$

[4] The projection operator $\mathcal{P}_D$ on any $N \times M$ matrix D of rank M is defined as $\mathcal{P}_D = D\left(D'D\right)^{-1}D'$. Its complement $\mathcal{Q}_D$ is given by $I_N - \mathcal{P}_D$. See Seber (2008) for the definition and properties of projection matrices.

- Conversely, the null space of $\mathcal{P}_{\hat{\hat{Z}}}$ and, thus, the row space of $\mathcal{Q}_{\hat{\hat{Z}}}$ consists of all $N \times 1$-vectors D for which $\hat{\hat{Z}}'D = \hat{B}_w'\tilde{W}'D = \mathbf{0}_{L\times 1}$. Within this row space we are looking for linear combinations $\tilde{W}\kappa$ of the instruments $\tilde{W}$. Consequently, $K_2 \times 1$ parameter vectors κ satisfying $\hat{B}_w'\tilde{W}'\tilde{W}\kappa = \mathbf{0}_{L\times 1}$ generate linear combinations $\tilde{W}\kappa$ lying within the row space of $\mathcal{Q}_{\hat{\hat{Z}}}$.

We now turn to the asymptotic distribution of the considered test statistics.

- In case $\kappa = \hat{B}_w\gamma$, equation (11) simplifies to

$$\mathcal{P}_{\tilde{W}}\tilde{Y} \;=\; \hat{\hat{Z}}\left(\alpha_z + \gamma\right) + \mathcal{P}_{\tilde{W}}\tilde{U},$$

  and the IV residuals $\hat{\hat{E}} = \tilde{Y} - \tilde{Z}\hat{\alpha}_{z;IV}$ are asymptotically given by

$$\begin{aligned}\hat{\hat{E}} &= \tilde{Z}\alpha_z + \hat{\hat{Z}}\gamma + \tilde{U} - \tilde{Z}\hat{\alpha}_{z;IV} \\ &\xrightarrow{p} \tilde{Z}\alpha_z + \left(\tilde{Z} - \tilde{V}\right)\gamma + \tilde{U} - \tilde{Z}\left(\alpha_z + \gamma\right) \\ &= \tilde{U} - \tilde{V}\gamma\end{aligned}$$

  since $\hat{B}_w \xrightarrow{p} B_w$ and $\hat{\alpha}_{z;IV} \xrightarrow{p} \alpha_z + \gamma$. As a consequence $N^{-\frac{1}{2}}\tilde{W}'\hat{\hat{E}} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}_{K_2\times 1}, \Sigma_I\right)$, where $\Sigma_I = \mathbf{Var}\left[\tilde{w}\left(\tilde{u} - \tilde{v}'\gamma\right)\right]$. Any test based on $\tilde{W}'\hat{\hat{E}}$, such as the Sargan-Hansen J-test (3) and the LM-test (4), will asymptotically have power equal to size against alternatives of the form $\tilde{E} = \left(\tilde{W}B_w\right)\gamma + \tilde{U}$. For the LM-test (4) the formal proof is given by Rao and Mitra (1971, Theorem 9.2.3, p.173), for the J-test the proof is given by Newey (1985, Theorem 1)[5].

- Alternatively, if $B_w'\tilde{W}'\tilde{W}\kappa = \mathbf{0}_{L\times 1}$[6], the IV residuals $\hat{\hat{E}} = \tilde{Y} - \tilde{Z}\hat{\alpha}_{z;IV}$ are asymptotically given by $\hat{\hat{E}} = \tilde{W}\kappa + \tilde{U} - \tilde{V}\Lambda\kappa$. Consequently, $N^{-\frac{1}{2}}\tilde{W}'\hat{\hat{E}} \xrightarrow{d} \mathcal{N}\left(\delta, \Sigma_{II}\right)$, where $\Sigma_{\tilde{w}} = \mathbf{Var}\left[\tilde{w}\right]$, $\delta = \sqrt{N}\Sigma_{\tilde{w}}\kappa$ and $\Sigma_{II} = \mathbf{Var}\left[\tilde{w}\left(\tilde{w}'\kappa + \tilde{u} - \tilde{v}'\Lambda\kappa\right)\right]$. Following Rao and Mitra (1971, Theorem 9.2.3, p.173), we have that $LM \xrightarrow{d} \chi^2_{K_2-L}\left(\delta'\Sigma_{II}^{-1}\delta\right)$, which is in accordance with Newey (1985). This shows that alternatives $\tilde{W}\kappa$ for which $B_w'\tilde{W}'\tilde{W}\kappa = \mathbf{0}_{L\times 1}$, are detected by both tests with probability one, for $N \to \infty$, since non-centrality parameter $\delta'\Sigma_{II}^{-1}\delta \sim N$ (in Bachmann-Landau notation[7]).$\square$

The consequences of above Theorem are clear. For a model (1)-(2) with (5), the parameter subspace consisting of $\kappa = \hat{B}_w\gamma$, $\gamma \in \mathbb{R}^L$ has identical asymptotic distribution for each of the test statistics considered. In addition, in the complement of this subspace the J and LM test statistics have power going to one for $N \to \infty$. Reformulation of the null hypothesis as $G_0 : \kappa = \hat{B}_w\gamma$, $\gamma \in \mathbb{R}^L$ thus results in a consistent test.

5. Discussion

[5]This proof is valid for any J-test, not only for (3), which is only applicable under homoskedasticity of the errors.

[6]It holds that $\hat{B}_w'\tilde{W}'\tilde{W}\kappa \xrightarrow{p} B_w'\Sigma_{\tilde{w}}\kappa$, where $\Sigma_{\tilde{w}} = \mathbf{Var}\left[\tilde{w}\right]$.

[7]$x \sim y$ denotes that $\exists C \in \mathbb{R} : \lim_{N\to\infty} \frac{x}{y} = C$ (Knuth, 1976).

Contrary to common formulations of over-identification tests, in the presence of $K_2 > L$ instruments, it *is* possible to test against alternatives where *all* instruments are endogenous. However these tests are inconsistent when used for testing $H_0 : \mathbf{E}[E \mid W] = \mathbf{0}_{N\times 1}$, in that they have no power against alternatives whereby the conditional mean of the disturbance E, given the instruments W is identical to a linear combination of the conditional mean(s) of the endogenous regressor(s) Z, given the instruments.

To illustrate, consider a model with one single variable of interest Z. Alternatives of the form $\tilde{E} = \left(\tilde{W}B_w\right)c + \tilde{U} = \hat{\tilde{Z}}c + \tilde{U}$, with c any scalar, go undetected. When the (incorrectly) excluded instruments appear in the (structural) outcome equation in the same proportion as they appear in the reduced form for Z, this violation of H_0 goes undetected. In other words, we can not distinguish the effect of $\tilde{Z}$ on $\tilde{Y}$ from the effect of $\hat{\tilde{Z}}$ on $\tilde{Y}$. This is an example of the equivalence between (non-)testability and (non-)identification (Dufour, 2003). All *other* departures from H_0, however, are testable. Consider some $K_2 \times 1$ vector λ for which $B_w'\lambda = 0$. It clearly holds that $\mathcal{P}_{\hat{\tilde{Z}}}\tilde{W}\lambda = 0$ and thus $\mathcal{Q}_{\hat{\tilde{Z}}}\tilde{W}\lambda \neq 0$, since the null-space of $\mathcal{P}_A$ is equal to the row space of $\mathcal{Q}_A$, for any matrix A. This type of alternative is thus clearly testable, although all instruments are endogenous. The requirement that at least L exogenous instruments are needed, for an over-identifying restriction test to work, thus seems to restrictive.

On the other hand, a priori assuming that L instruments are valid, is one way to ensure that κ, having L elements equal to zero, will never be equal to $B_w\gamma$. Interpreted this way, the inconsistency of the test is assumed away.

In contrast, I propose to formulate the different null $G_0 : B_w\gamma = \mathbf{0}_{L\times 1}$. Remark that $H_0 \subset G_0$. This extended null, in combination with the test statistics (3) and (4) results in a consistent test, as guaranteed by the above theorem.

The difference between both approaches can also be evaluated in terms of the dimensionality of the space over which is tested. The space spanned by the K_2 instruments can be represented by $\mathbb{R}^{K_2}$. Making over-identifying restrictions tests consistent by extra assumptions, reduces the space over which is tested to $\mathbb{R}^{K_2 - L}$. However, changing the null from H_0 to G_0 leaves $\mathbb{R}^{K_2} \setminus \mathbb{R}^L$ alternatives testable, a clearly larger set. The validity of our claims will be checked empirically in the next section.

6. Monte Carlo

In this section the results of a small Monte Carlo study nicely illustrate the findings from previous sessions. The *DGP* consists of model (1)-(2)

$$\begin{aligned} Y &= X\alpha_x + Z\alpha_z + W\alpha_w + U \\ Z &= X\beta_x + W\beta_w + V, \end{aligned}$$

with $(K_1, K_2, L) = (2, 2, 1)$. The variables $(x, w, u, v)' \sim NID(0; I_6)$ $\beta_x = (1, 1)$, $\beta_w = (1, 1)$, $\alpha_x = (1, 1)$, $\alpha_z = 1$, $\alpha_w = \theta\{\phi\gamma_w + (1-\phi)\beta_w\}$, where $\gamma_w = (1, -1)$, $\theta = 0.1$, 0.2 and $\phi = 0$, 0.25, 0.5, 0.75, 1. The data were generated for the sample sizes $N = 100$, 200, 500, 1000 and the number of Monte Carlo replications was 5000.

It is easy to verify that $\theta = 0$ corresponds to H_0, the null that all instruments are exogenous. The null G_0 is given by $(\phi = 0) \cup (\theta = 0)$. The parameter θ can be interpreted

N	θ	0	0.1					0.2				
	ϕ	/	0	0.25	0.5	0.75	1	0	0.25	0.5	0.75	1
J-statistic (3)												
100		0.0656	0.0660	0.0698	0.0992	0.1398	0.1936	0.0628	0.1044	0.2074	0.3494	0.5204
200		0.0574	0.0592	0.0734	0.1208	0.1966	0.2920	0.0586	0.1236	0.3168	0.5722	0.7978
500		0.0530	0.0508	0.0864	0.2096	0.3982	0.6024	0.0552	0.2236	0.6544	0.9302	0.9920
1000		0.0518	0.0532	0.1418	0.3658	0.6630	0.8896	0.0532	0.4050	0.9110	0.9986	1.0000
LM-statistic (4)												
100		0.0562	0.0578	0.0622	0.0870	0.1270	0.1760	0.0554	0.0926	0.1890	0.3246	0.4956
200		0.0530	0.0544	0.0692	0.1138	0.1880	0.2794	0.0548	0.1168	0.3060	0.5624	0.7866
500		0.0514	0.0500	0.0844	0.2074	0.3938	0.5966	0.0542	0.2188	0.6494	0.9276	0.9918
1000		0.0512	0.0530	0.1406	0.3634	0.6608	0.8882	0.0522	0.4016	0.9104	0.9986	1.0000

Table 1: Rejection percentages of $H_0 : \kappa = 0$ for the different test statistics

as the degree to which H_0 is violated. Similarly, $\theta\phi$ constitutes a measure of the distance between the DGP and G_0. According to the theorem above DGPs for which $\theta = 0$ will have power equal to size. DGPs for which $\theta = 1$, have $B_w'\Sigma_w\kappa = \beta_w'\kappa = 0$, and have thus power going to one for $N \to \infty$. Intermediate values can be decomposed into an undetectable component and a detectable component, the latter having power going to one for $N \to \infty$.

In Table 1 the rejection percentages are given for both the J-statistic (3) and the LM-statistic (4). The table clearly shows that rejection percentages are very close to their nominal values for all parameter combinations that fall under G_0, with the LM-test staying slightly closer to its nominal size. In addition, for every column that falls under G_A, the power increases to one for increasing N. The difference between J and LM tests is again negligible.

7. Conclusion

In a way, above result is comforting. When the inconsistency of validity tests is interpreted as an a priori requirement of L valid instruments (Stock, 2001), the seemingly insurmountable burden of proof rests on the scientist applying IV, who has to argue why these L instruments are exogenous. However, the assumption that the excluded instruments W do not appear in the outcome equation (1) in some linear combination of B_w seems much easier to defend and the burden of proof seems to rest on the critic, who has to make the case why the instruments would appear in the outcome equation in exactly this linear combination.

References

Dufour, J.-M. (2003) "Identification, Weak Instruments, and Statistical Inference in Econometrics" *Canadian Journal of Economics-Revue canadienne d'économique* **36**(4), 767–808.

Hansen, L. (1982) "Large Sample Properties of Generalized Method of Moments Estimators" *Econometrica* **50**(4), 1029–1054.

Knuth, D.E. (1976) "Big Omicron and big Omega and big Theta" *ACM SIGACT News* **8**(2), 18–24.

Magdalinos, M.A. (1988) "The Local Power of the Tests of Overidentifying Restrictions" *International Economic Review* **29**(3), 509–524.

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

Rao, C.R., Mitra, S.K. (1971) *Generalized Inverse of Matrices and Its Applications*, Wiley Series in Probability and Mathematical Statistics, Wiley, New York.

Seber, G.A.F. (2008) *A Matrix Handbook for Statisticians*, Wiley Series in Probability and Statistics, New Jersey.

Shea, J. (1997) "Instrument Relevance in Multivariate Linear Models: A Simple Measure" *Review of Economics and Statistics* **79**(2), 348–352.

Stock, J.H. (2001) "Instrumental Variables in Statistics and Econometrics" *International Encyclopaedia of the Behavioural Sciences*, Elsevier Publishing: Amsterdam, 7577–7582.

van der Vaart, A.W. (2000) *Asymptotic Statistics*, Cambridge University Press: Cambridge, U.K.

Wooldridge, J.M. (1990) "A Note on the Lagrange Multiplier and F-statistics for Two Stage Least Squares Regressions" *Economics Letters* **34**(2), 151–155.