On bootstrap validity for specification tests with weak instruments

Doko Tchatoka, Firmin

University of Tasmania, School of Economics and Finance

31 March 2013

On bootstrap validity for specification tests with weak instruments

Firmin Doko Tchatoka*

University of Tasmania

June 9, 2013

* School of Economics and Finance, University of Tasmania, Private Bag 85, Hobart TAS 7001, Phone: +613 6226 7226, Fax:+61 3 6226 7587; e-mail : Firmin.dokotchatoka@utas.edu.au

ABSTRACT

This paper investigates the asymptotic validity of the bootstrap for Durbin-Wu-Hausman (DWH) specification tests when instrumental variables (IVs) may be arbitrary weak. It is shown that under strong identification, the bootstrap offers a better approximation than the usual asymptotic χ^2 distributions. However, the bootstrap provides only a first-order approximation when instruments are weak. This indicates clearly that unlike the Wald-statistic based on a k-class type estimator (Moreira et al., 2009), the bootstrap is valid even for the Wald-type of DWH statistics in the presence of weak instruments.

Key words: Specification tests; weak instruments; bootstrap.

JEL classification: C3; C12; C15; C52.

1. Introduction

Specification tests of the type proposed by Durbin (1954), Wu (1973, 1974), and Hausman (1978), henceforth DWH tests, are widely used in applied work to decide whether the ordinary least squares (OLS) or instrumental variables (IV) method is appropriate. Although research on exogeneity testing in linear IV regressions is widespread[1], most studies in this topic usually consider the case of strong instruments. Recent studies focusing on the behavior of the DWH-type tests document that they never over-rejects the null hypothesis of exogeneity when IVs are weak. However, some of these tests can be overly conservative even in large-sample, and have low power when identification is weak.[2] Doko Tchatoka and Dufour (2011b) propose a size correction of these tests through the exact Monte carlo test procedure [Dufour (2006)], which remains valid even when identification is weak and the sample size is small. However, the Monte Carlo test procedure suggested requires the a priori knowledge of the distribution of model disturbance, at least up to an unknown scale factor. But in practice, researchers usually do not know the exact distribution of the errors and implementing the simulated method can be difficult, even infeasible.

This paper aims to relax this distributional assumption by resorting to bootstrap methods. We mainly focus on linear structural models and establish the asymptotic validity of the bootstrap for DWH exogeneity tests, when IVs may be arbitrary weak (weak instruments).

Moreira et al. (2009) show in the context of hypotheses specified on structural

[1]See, for example, Durbin (1954), Wu (1973, 1974, 1983a, 1983b), Revankar and Hartley (1973), Farebrother (1976), Hausman (1978), Revankar (1978), Dufour (1979, 1987), Hwang (1980), Kariya and Hodoshima (1980), Hausman and Taylor (1981), Spencer and Berk (1981), Nakamura and Nakamura (1981), Engle (1982), Holly (1982), Reynolds (1982), Smith (1983, 1984), Thurman (1986), Smith and Pesaran (1990), Ruud (1984, 2000), Newey (1985a, 1985b), Wong (1996), Ahn (1997), Baum, Schaffer and Stillman (2003).

[2]See, for examples, Staiger and Stock (1997), Guggenberger (2010), and Doko Tchatoka and Dufour (2011a, 2011b). Staiger and Stock (1997, Section D) show that with weak IVs, the size of Hausman (1978) tests that exploit the residuals from the 2SLS estimation, and that of the Wu (1973) $\mathcal{T}_3$ test depends on identification strength through the concentration matrix. Since the concentration matrix cannot be estimated consistently when IVs are weak, Staiger and Stock (1997) conclude that size adjustment of these statistics is infeasible. But Doko Tchatoka and Dufour (2011b) show the size of all DWH-type statistics can be adjusted using the simulated methods; see also Dufour (2006).

parameters, that the bootstrap is valid for the score test. This not however the case for Wald-type tests based on the 2SLS or LIML estimators when IVs are weak. We use the LM and Wald interpretation of the DWH staistics in Engle (1982) and Smith (1983) to propose a slight modification of Moreira et al.'s (2009) bootstrap. Our analysis of the bootstrap validity provides some new insights and extensions of Moreira et al.'s (2009). We show that when identification is strong, the bootstrap offers a better approximation than the usual asymptotic χ^2 distributions (similar to Moreira et al., 2009). However, the bootstrap provides only a first-order approximation when identification is weak, meaning that the bootstrap is valid even for the wald-type of the DWH test, despite the lack of identifiability. This contrasts with the bootstrap of the Wald-statistic based on the 2SLS or LIML estimators, which is invalid with weak IVs (Moreira et al., 2009).

The paper is organized as follows. Section 2 formulates the model and assumptions, and presents the statistics studied. Section 3 presents the statistics and provides their Lagrange multiplier or Wald interpretation, following Engle (1982) and Smith (1983). Section 4 details the proposed bootstrap implemented as well as its validity in both strong and weak instrument setups. Conclusions are drawn in Section 5 and the proofs and auxiliary lemmas are presented in the Appendix.

2. Framework

We consider the standard linear structural model described by the following equations:

$$
\begin{aligned}
y_1 &= y_2\beta + Z_1\gamma + u, &\quad (2.1)\\
y_2 &= Z\pi_2 + Z_1\pi_1 + v_2 &\quad (2.2)
\end{aligned}
$$

where y_1 and y_2 are $n \times 1$ vectors of observations on two endogenous variables, Z_1 is a $n \times k_1$ matrix of included exogenous variables, Z_1 is a $n \times k_2$ matrix instruments, $u = (u_1, \ldots, u_n)' \in \mathbb{R}^n$ is a vector of structural disturbances, $v_2 = [v_{21}, \ldots, v_{2n}]' \in \mathbb{R}^n$ is a vector of reduced form disturbances, β, $\gamma \in \mathbb{R}$ are unknown structural parameters, while $\pi_1 \in \mathbb{R}^{k_1}$ and $\pi_2 \in \mathbb{R}^{k_2}$ is the unknown reduced-form coefficient vector. The results in this paper can easily be extended to setups where y_2 contains more than one regressors. We assume that $Z = [Z_1 : Z_2] : n \times k$ has full column-rank $k = k_1 + k_2$.

The reduced-forms for y_1 and y_2 can be expressed from (2.1)-(2.2) as:

$$\begin{aligned} y_1 &= Z_1(\pi_1\beta + \gamma) + Z_2\pi_2\beta + v_1 \\ y_2 &= Z_1\pi_1 + Z_2\pi_2 + v_2, \end{aligned} \tag{2.3}$$

where $v_1 = u + v_2\beta$. For any random matrix X, let X_i denote the i-th row of X, written as column vector. Let $Y = [y_1 : y_2]$ and define

$$\mathcal{Q}_n = vech\left((Y_n', Z_n')'(Y_n', Z_n')\right) = (f_1(Y_n', Z_n'), f_1(Y_n', Z_n'), \ldots, f_l(Y_n', Z_n')), \tag{2.4}$$

where f_i, $i = 1, \ldots, l$, $l = (k+2)(k+3)/2$, $k = k_1 + k_2$, are elements of the matrix $(Y_n', Z_n')'(Y_n', Z_n')$. Let $\bar{\mathcal{Q}}_n = n^{-1}\sum_{i=1}^{n} \mathcal{Q}_i$ denote the empirical mean of the $\mathcal{Q}_i$. The following assumptions are made on the behavior of model variables.

Assumption 2.1 *(a) $\mathcal{Q}_n$ in (2.4) satisfies:* $\mathbb{E}[\|\mathcal{Q}_n\|^s] < \infty$ *for some* $s \geq 3$, $\limsup_{\|t\|\to\infty} | \mathbb{E}[exp(it'Q_n)] | < 1$; *and (b) when the sample size n converges to infinity, the following convergence results hold jointly*:

M1. $\quad n^{-1}[u : v_2]'[u : v_2] \xrightarrow{p} \Sigma = \begin{pmatrix} \sigma_u^2 & \delta \\ \delta & \sigma_{v_2}^2 \end{pmatrix}$, $n^{-1}Z'Z \xrightarrow{p} Q_Z$, $n^{-1}Z'[u : v_2] \xrightarrow{p} 0$

M2. $\quad n^{-1/2}Z'[u : v_2] \xrightarrow{d} \left[\psi_{Zu} : \psi_{Zv_2}\right]$, *where* $\psi_{Zu} = (\psi_{Z_1u}', \psi_{Z_2u}')' : k \times 1$,

$\psi_{Zv_2} = (\psi_{Z_1v_2}', \psi_{Z_2V-2}')' : k \times 1$, *and* $vech\left(\left[\psi_{Zu} : \psi_{Zv_2}\right]\right) \sim N(0, \Sigma \otimes Q_Z)$.

The first moment condition in Assumption **2.1**-(a) holds if $\mathbb{E}[\|(Y_n', Z_n')\|^{2s}] < \infty$, and the second is the commonly used Cramér's condition [see Bhattacharya and Ghosh

(1978)]. In Assumption **2.1**-(b), $M1$ is the weak law of large numbers (WLLN) property, where IVs and disturbances are asymptotically uncorrelated, while $M2$ is the central limit theorem (CLT) property.

From Assumption **2.1**, the exogeneity hypothesis of y_2 can be expressed as:

$$\mathrm{H}_0 : \delta = 0. \tag{2.5}$$

We are concerned with the asymptotic validity of the bootstrap for the DWH statistics often used to assess H_0, especially when identification is weak. Section 3 presents the DWH statistics and their LM or Wald interpretation.

3. Lagrange Multiplier and Wald Nature of the Standard DWH Tests

We consider the statistics $\mathcal{T}_l$, $l = 2, 3, 4$, by Wu (1973, 1974) and three alternative Hausman (1978) type statistics, namely, $\mathcal{H}_j$, $j = 1, 2, 3$. Let $A_1 = I_n - Z_1(Z_1'Z_1)^{-1}Z_1'$ and $A_2 = I_n - Z(Z'Z)^{-1}Z'$ denote the orthogonal matrices to the spaces spanned by the columns of Z_1 and Z, respectively. The statistics $\mathcal{T}_l$ and $\mathcal{H}_j$ can be expressed in the unified formulation as:

$$\mathcal{T}_l = \kappa_l(\tilde{\beta} - \hat{\beta})^2/\tilde{\omega}_l^2\,, \quad l = 2, 3, 4, \tag{3.1}$$

$$\mathcal{H}_j = n(\tilde{\beta} - \hat{\beta})^2/\hat{\omega}_j^2\,, \quad j = 1, 2, 3 \tag{3.2}$$

where $\hat{\beta} = (y_2'A_1y_2)^{-1}y_2'A_1y_1$ and $\tilde{\beta} = [y_2'(A_1 - A_2)y_2]^{-1}y_2'(A_1 - A_2)y_1$ are the OLS and IV estimators of β, respectively, and

$$\tilde{\omega}_2^2 = \tilde{\sigma}_2^2\hat{\Delta},\ \tilde{\omega}_3^2 = \tilde{\sigma}^2\hat{\Delta},\ \tilde{\omega}_4^2 = \hat{\sigma}^2\hat{\Delta},$$

$$
\begin{aligned}
\hat{\omega}_1^2 &= \tilde{\sigma}^2\hat{\omega}_{iv}^{-1} - \hat{\sigma}^2\hat{\omega}_{ls}^{-1},\ \hat{\omega}_2^2 = \tilde{\sigma}^2\hat{\Delta},\ \hat{\omega}_3^2 = \hat{\sigma}^2\hat{\Delta},\\
\hat{\Delta} &= \hat{\omega}_{iv}^{-1} - \hat{\omega}_{ls}^{-1},\ \hat{\omega}_{iv} = y_2'(A_1 - A_2)y_2/n,\ \hat{\omega}_{ls} = y_2'A_1y_2/n,\\
\tilde{\sigma}^2 &= (y_1 - y_2\tilde{\beta})'A_1(y_1 - y_2\tilde{\beta})/n,\ \hat{\sigma}^2 = (y_1 - y_2\hat{\beta})'A_1(y_1 - y_2\hat{\beta})/n,\\
\tilde{\sigma}_2^2 &= \hat{\sigma}^2 - (\tilde{\beta} - \hat{\beta})^2/\hat{\Delta},\ \kappa_2 = n - 2 - k_1,\ \kappa_3 = \kappa_4 = n - 1 - k_1.
\end{aligned}
$$

Engle (1982) and Smith (1983) show that each statistic in (3.1)-(3.2) has a score or Wald interpretation. The statistics $\mathcal{T}_2$, $\mathcal{T}_4$, and $\mathcal{H}_3$ are LM-type, while $\mathcal{T}_3$, $\mathcal{H}_1$, and $\mathcal{H}_2$ are quasi-Wald type.[3] Under H_0 and if further Assumption **2.1**-(b) holds, all DWH statistics have the usual chi-square asymptotic distributions if model identification is strong. However, $\mathcal{T}_3$, $\mathcal{H}_1$, and $\mathcal{H}_2$ are overly conservative, and all DWH tests have a low power if IVs are weak, even in large-sample. We question whether a bootstrap technique can improve[4] the properties of the DWH tests, with or without weak instruments.

4. Bootstrap Validity for DWH Tests

Let $\hat{\pi} = (Z'Z)^{-1}Z'y_2$ denotes the OLS estimator of $\pi = (\pi_1', \pi_2')'$ in the first stage regression (2.2). Let $\hat{\theta}$ be an estimator of β and $\hat{\gamma}$ those of γ. The bootstrap procedure consists of the following steps:

[3] See Smith (1983) for the score interpretation (Eqs. [6] and [9]) and for the quasi-Wald interpretation (Eqs. [7], [8] and [10]). The regression interpretation of these statistics is provided in Hausman (1978), Dufour (1979, 1987), Wooldridge (2009), and Doko Tchatoka and Dufour (2011b).

[4] Due to the LM nature of $\mathcal{T}_2$, $\mathcal{T}_4$, $\mathcal{H}_3$, and the result in Moreira et al. (2009), one can project the bootstrap validity for these statistics. But formal proof needs to be established, especially because the primary focus in Moreira et al. (2009) is not exogeneity testing, and there is no discussion in Moreira et al. (2009) related to exogeneity testing. On the other hand, because of the Wald nature of $\mathcal{T}_1$, $\mathcal{T}_3$, $\mathcal{H}_1$, $\mathcal{H}_2$, and the bootstrap invalidity result for the Wald-statistic in Moreira et al. (2009), it is not clear whether the bootstrap applies to these statistics. Hence, this note is useful in clarifying these issues. Wong (1996) illustrates through a Monte Carlo experiment that bootstrapping the Hausman (1978) exogeneity test improves both the size and power of the test. Li (2006) extends Wong's (1996) results by allowing for serial correlated errors. Both papers are referenced in the weak instrument literature. However, neither Wong (1996) nor Li (2006) provides a formal proof of the large-sample validity of their bootstrap, even when IVs are strong. Furthermore, the Monte Carlo designs in both papers exclude cases where IVs are poor, because the smallest correlation between each IV and the (possibly) endogenous regressors is set at 0.1. Although a correlation of 0.1 is not hight, it is not zero or close to either.

1. From observed data, compute $\hat{\pi}$ and $\hat{\theta}$ along with all other things necessary to get the realizations of the statistics $\mathcal{T}_l$, $\mathcal{H}_j$, and the residuals from the reduced-form equation (2.3): $\hat{v}_1 = y_1 - Z_1(\hat{\pi}_1\hat{\theta} + \hat{\gamma}) - Z_2\hat{\pi}_2\hat{\theta}$, $\hat{v}_2 = y_2 - Z\hat{\pi}$. These residuals are then re-centered by subtracting sample means to yield $(\tilde{v}_1, \tilde{v}_2)$.

2. For each bootstrap sample $r = 1, \ldots, B$, data are generated as:

$$y_1^* = Z_1^*(\hat{\pi}_1\hat{\theta} + \hat{\gamma}) + Z_2^*\hat{\pi}_2\hat{\beta} + v_1^*, \quad y_2^* = Z^*\hat{\pi} + v_2^* \tag{4.1}$$

where $Z^* = [Z_1^* : Z_2^*]$ and (v_1^*, v_2^*) are drawn independently from the empirical distribution of Z and $(\tilde{v}_1, \tilde{v}_2)$. The corresponding bootstrap statistics $\mathcal{T}_l^{*^r}$ and $\mathcal{H}_j^{*^r}$ are then computed for each bootstrap sample $r = 1, \ldots, B$.

3. The simulated bootstrap p-value is obtained as the proportion of bootstrap statistics that are more extreme than the statistics computed from observed data.

4. The bootstrap test rejects the null hypothesis of exogeneity at level α if its p-value is less than α.

The above bootstrap steps, though similar to those by Moreira et al. (2009), have a slight difference in the appropriate[5] estimator of $\hat{\theta}$ to be used; see fn.4 for further details. We now show the asymptotic validity of the bootstrap.

[5]Moreira et al. (2009) show that $\hat{\theta}$ must be strongly consistent, i.e.,

$$\hat{\pi} \xrightarrow{p} \pi \quad \text{and} \quad \hat{\theta}\hat{\pi} \xrightarrow{p} \theta\pi, \tag{4.2}$$

for the bootstrap to be valid. In a linear classical setting, the 2SLS and LIML estimators satisfy the sufficient conditions for strong consistency; see Moreira et al (2009, Proposition 4 and fn.3, p.55). The OLS estimator is not qualified for (4.2) if $\delta \neq 0$ (endogeneity). However, under the null hypothesis of exogeneity ($\delta = 0$), as it is the case here, the OLS estimator is consistent and further efficient, no matter how weak the IVs are. For this reason, we prefer OLS to an alternative 2SLS or LIML estimator.

4.1. High-order approximation with strong instruments

In this section, we focus on the case where $\pi \neq 0$ is fixed (strong IVs). We can express the bootstrap DHW statistics $\mathcal{T}_l^*$ and $\mathcal{H}_j^*$ based on the re-centering residuals as:

$$\mathcal{T}_l^* = \left(\sqrt{n}G(\bar{\tilde{Q}}_n^*)\right)^2, \mathcal{H}_j^* = \left(\sqrt{n}\tilde{G}(\bar{\tilde{Q}}_n^*)\right)^2 \quad \text{for all } l \text{ and } j \tag{4.3}$$

where $\bar{\tilde{Q}}_n$ and $\bar{\tilde{Q}}_n^*$ are analogous of $\bar{\mathcal{Q}}_n$ in (2.4). $\bar{\tilde{Q}}_n$ is based on the sample re-centering residuals and $\bar{\tilde{Q}}_n^*$ is based on the bootstrap sample residuals. The functions $G(.)$ and $\tilde{G}(.)$ are real-valued Borel measurable functions on $\mathbb{R}^l$, which satisfy $G(\bar{\tilde{Q}}_n) = 0$ and $\tilde{G}(\bar{\tilde{Q}}_n) = 0$, due to the re-centered mechanism [similar to Eqs. (A.5)-(A.6) in the Appendix]. Under strong identification, all derivatives of order s and less of the functions $G(.)$ and $\tilde{G}(.)$ are continuous. So, Edgeworth-type expansion[6] applies and we have the following theorem.

Theorem 4.1 Bootstrap validity with Strong IVs. *Suppose Assumption* **2.1** *is satisfied. Under H_0 and if further $\pi \neq 0$ is fixed, we have*:

$$\begin{aligned}\| \mathbb{P}^*(\mathcal{T}_l^* \leq x) - [\Phi(x) + \sum_{m=1}^{s-2} n^{-m/2} p_{\mathcal{T}_l}^m(x; F_n, \hat{\beta}, \hat{\pi})\Phi(x)]^2 \|_\infty &= o(n^{(s-2)}), \\ \| \mathbb{P}^*(\mathcal{H}_j^* \leq x) - [\Phi(x) + \sum_{m=1}^{s-2} n^{-m/2} p_{\mathcal{H}_j}^m(x; F_n, \hat{\beta}, \hat{\pi})\Phi(x)]^2 \|_\infty &= o(n^{(s-2)}) \text{ as } n \to \infty\end{aligned}$$

for all l and j, where $p_{\mathcal{T}_l}^m$ and $p_{\mathcal{H}_j}^m$ are polynomials in x with coefficients depending on $\hat{\beta}$, $\hat{\pi}$, and the moments of the distribution F_n of $\tilde{\mathcal{Q}}_n^* = vech\left((\tilde{Y}_n^{*\prime}, \tilde{Z}_n^{*\prime})'(\tilde{Y}_n^{*\prime}, \tilde{Z}_n^{*\prime})\right)$ conditional on $\hat{\mathscr{F}}_n = \{(Y_1', Z_1'), \ldots, (Y_n', Z_n')\}$; $\Phi(.)$ is the cdf of $N(0.1)$ and $\|.\|_\infty$ is the supremum norm.

First, Theorem **4.1** shows that the bootstrap approximates the empirical Edgeworth expansion in Lemma **A.1** up to the $o(n^{(s-2)})$ order. This is not surprising because the conditional moments of $\mathcal{Q}_n^*$, given the data $\hat{\mathscr{F}}_n$, converge almost surely

[6] Such as in Bhattacharya and Ghosh (1978, Theorem 2).

to those of $\mathcal{Q}_n$ when identification is strong. Second, the results shows that the error based on the bootstrap simulation is of order n^{-1}. Therefore, the bootstrap offers a better approximation than the usual asymptotic χ^2 distributions, even for the Wald-type versions of the DWH statistics.

4.2. First-order Validity with Weak Instruments

High-order approximation of the limiting distributions of the bootstrap as in Theorem **4.1** is not achievable now due to the lack of identification. Indeed, when $\pi_2 = \pi_0/\sqrt{n}$ where π_0 is a $k_2 \times 1$ constant vector, the functions $G(.)$ and $\tilde{G}(.)$ in (4.3) are non-differentiable.[7] So, the Edgeworth expansion is not applicable. However, we can prove the following theorem on the first-order approximation of the bootstrap when IVs are weak.

Theorem 4.2 Bootstrap validity with weak IVs. *Suppose Assumptions* **2.1** *and* H_0 *are satisfied. If for some* $\delta > 0$, $\mathbb{E}(\|Z_i\|^{4+\delta},\ \|v_i\|^{2+\delta}) < \infty$, *then we have*:

$$\mathcal{T}_l^{*} \mid \hat{\mathscr{F}}_n \ \xrightarrow{d}\ \chi^2(1), \quad \mathcal{H}_j^{*} \mid \hat{\mathscr{F}}_n \xrightarrow{d} \chi^2(1) \text{ a.s., for all } l = 2,\, 3,\, 4;\ j = 1,\, 2,\, 3$$

when $\pi = \pi_0/\sqrt{n}$, π_0 *is a* $k \times 1$ *constant vector, and* $\hat{\mathscr{F}}_n = \{(Y_1',\, Z_1'),\, \ldots,\, (Y_n',\, Z_n')\}$.

First, since the statistics $\mathcal{T}_2$, $\mathcal{T}_4$, and $\mathcal{H}_3$ are LM-type and following Moreira et al. (2009), the bootstrap validity for these statistics is predictable. However, the result of the Wald-type of the DWH statistics, $\mathcal{T}_3$, $\mathcal{H}_1$, and $\mathcal{H}_2$, is less obvious, because the bootstrap is not valid for the Wald-statistic of $\mathrm{H}_\beta : \beta = \beta_0$ (see Moreira et al., 2009). The key reason behind the bootstrap validity for the Wald-statistic here is that their asymptotic distributions, even when $\delta \neq 0$, do not depend on the unknown nuisance parameter[8] β, with or without weak IVs. Meanwhile, the asymptotic distribution of

[7]Note that all DWH statistics depends on $y_2'(A_1 - A_2)y_2/n$. However, it is straightforward to see that the derivative of the functions $G(.)$ and $\tilde{G}(.)$ with respect to $y_2'(A_1 - A_2)y_2/n$ is not well-defined when $\pi = 0$ or does not exist if $\pi = \pi_0 c_n$ for any sequence $c_n \downarrow 0$. So, $G(.)$ and $\tilde{G}(.)$ are not smooth when IVs are weak, and Edgeworth-type expansion does not apply.

[8]See Wu(1973, Section 3; 1974, Eqs. [3.11]-[3.16]) and Doko Tchatoka and Dufour (2011a, 2011b).

the Wald-statistic of $H_\beta : \beta = \beta_0$, based on 2SLS or LIML, depends heavily[9] on β under the weak instrument scenario.

4.3. Monte Carlo experiment

We use simulation to examine the size performance of the proposed bootstrap. The DGP is described[10] by Eqs. (2.1) and (2.2) where the n rows of $[u, v_2]$ are drawn i.i.d. with mean zero and unit variance, and the correlation between u_i and v_{2i} is set at $\rho = 0$ under H_0. Z_2 contains k_2 instruments, each generated i.i.d $\mathbf{N}(0, 1)$ independently of $[u, v_2]$. We vary k_2 in $\{2, 5, 20\}$ within the experiment, but the results are consistent with alternative values. The true value of β is set at 2 and the reduced-form coefficient π_2 is chosen as $\pi_2 = \left(\frac{\mu^2}{n\|Z_2\pi_0\|}\right)^{1/2}\pi_0$, where π_0 is a vector of ones, μ^2 is the concentration parameter characterizing the strength of the IVs. In this experiment, μ^2 varies in $\{0, 413, 1000\}$.[11] To account for non-normal errors, $[u, v_2]$ is generated following Kotz et al. (2000):

$$u_i = a + b\varepsilon_{1i} + c\varepsilon_{1i}^2 + d\varepsilon_{1i}^3, \ v_{2i} = a + b\varepsilon_{2i} + c\varepsilon_{2i}^2 + d\varepsilon_{2i}^3 \tag{4.4}$$

where $(\varepsilon_{1i}, \varepsilon_{2i})' \overset{i.i.d.}{\sim} \mathbf{N}(0, I_2)$ for all $i = 1, \ldots, n$. We consider two setups: (1) $a = c = d = 0$ and $b = 1$ (normal errors), and $a = c = 0$, $b = d = 1/\sqrt{22}$ (non-normal errors) such that[12] $Sknew = 0$ and $Kurt \approx 27.72$.

Table 1 presents the results for the standard DWH tests, and Table 2 reports those of the bootstrap tests. The first column of each table contains the test statistics, the second reports the number of IVs k_2, while the others present, for each sample size (n)

[9] See Nelson and Startz (1990); Staiger and Stock (1997); Dufour (1997, 2003); Wang and Zivot (1998), among others.

[10] There is no exogenous Z_1 in the simulations but the results do not alter when such exogenous IVs are included.

[11] Following Hansen et al. (2008), $\mu^2 = 0$ is a design of complete non-identification, $\mu^2 = 413$ designs weak identification, and $\mu^2 = 1000$ is for strong identification.

[12] We run the simulations with alternative values of $(Skew, Kurt)$ and the results are qualitatively similar.

and the IV strength (μ^2), the empirical rejections of the tests. The bootstrap rejection probability is estimated using $10,000$ pseudo-sample sets, each of size n varying in $\{50, 100, 200, 500, 1000\}$. The nominal level for both the standard and bootstrap tests is 5%. It is clear from Table 1 that the standard Wald-type of the DWH tests, namely, $\mathcal{T}_3$, $\mathcal{H}_1$, and $\mathcal{H}_2$, are highly conservative with weak IVs (see columns $\mu^2 = 0$ and $\mu^2 = 413$). The rejection frequencies of the LM-type tests —$\mathcal{T}_2$, $\mathcal{T}_4$, and $\mathcal{H}_3$— are close to the nominal level of 5% even when IVs are weak. These results are similar for normal and non-normal errors. Meanwhile, Table 2 shows clearly that the bootstrap method improves the size of the tests, especially for the Wald-type of the DWH tests. As seen, even the rejection frequencies of $\mathcal{T}_3$, $\mathcal{H}_1$, and $\mathcal{H}_2$ are very close to the nominal level, no matter how weak the IVs are, with or without normal errors, even with relatively small-sample sizes.

5. Conclusion

This paper considers the standard linear IV models and investigates the asymptotic validity of the bootstrap for the standard DWH exogeneity tests. We propose a slight modification of Moreira et al.'s (2009) bootstrap, which provides some new insights and extensions of earlier results. When identification is strong, we show that the bootstrap offers a better approximation of the distributions of the statistics than the usual asymptotic χ^2 distributions. However, it provides only a first-order approximation when instruments are weak. Unlike the Wald-statistic based on the 2SLS estimator (see Moreira et al., 2009), ours results show that the bootstrap is valid even for the Wald-type of the DWH statistics. This is mainly because even when identification is weak, the asymptotic distributions of all DWH statistics, including the Wald-type ones, do not depend on the unknown structural parameters, while those of the Wald-statistic based on 2SLS or LIML estimator does.

Acknowledgments

We are grateful to Professor Richard J. Smith, Manaing Editor of *The Econometrics Journal*, for his constructive comments and suggestions. We would also like to thank Professors Jean-Marie Dufour, Mardi Dungey and Lynda Khalaf for several useful comments.

Table 1. Rejection frequencies (in %) of the standard DWH tests

Statistics	$k_2 \downarrow \mu^2 \rightarrow$	$n=50$			$n=100$			$n=200$			$n=500$			$n=1000$		
Normal errors																
		0	413	1000	0	413	1000	0	413	1000	0	413	1000	0	413	1000
$\mathcal{T}_2$	2	6.3	6.4	7.1	5.7	5.2	5.7	6.0	5.6	5.7	6.2	5.6	5.6	6.1	5.5	5.6
$\mathcal{T}_2$	2	5.6	5.4	5.0	5.1	5.1	5.4	4.9	5.2	5.0	4.9	5.2	5.0	4.9	4.7	4.9
$\mathcal{T}_3$	2	0.1	4.0	4.2	0.0	3.7	4.6	0.0	3.7	4.2	0.1	2.9	3.8	0.0	2.0	3.6
$\mathcal{T}_4$	2	4.7	4.7	4.3	4.8	4.7	4.9	4.7	5.0	4.8	4.9	5.1	4.9	4.9	4.7	4.9
$\mathcal{H}_1$	2	0.1	3.5	3.6	0.0	3.4	4.2	0.0	3.6	4.0	0.1	2.8	3.7	0.0	1.9	3.6
$\mathcal{H}_2$	2	0.1	4.2	4.3	0.0	3.8	4.7	0.0	3.8	4.3	0.1	2.9	3.8	0.0	2.0	3.6
$\mathcal{H}_3$	2	5.0	4.9	4.5	4.9	4.8	5.1	4.8	5.1	4.9	4.9	5.1	4.9	4.9	4.7	4.9
$\mathcal{T}_2$	5	5.5	5.5	5.4	5.3	5.4	5.4	4.9	5.5	5.1	4.7	5.2	5.0	5.3	4.9	5.2
$\mathcal{T}_3$	5	0.3	4.5	4.7	0.3	4.9	5.0	0.3	4.6	4.8	0.4	4.2	4.6	0.3	3.9	4.5
$\mathcal{T}_4$	5	5.0	4.8	4.8	4.9	5.2	5.1	4.8	5.3	5.0	4.6	5.1	5.0	5.3	4.9	5.1
$\mathcal{H}_1$	5	0.2	3.9	4.0	0.3	4.5	4.6	0.2	4.5	4.6	0.3	4.2	4.6	0.2	3.8	4.5
$\mathcal{H}_2$	5	0.3	4.8	4.9	0.3	4.9	5.1	0.3	4.7	4.8	0.4	4.2	4.7	0.3	3.9	4.5
$\mathcal{H}_3$	5	5.2	5.1	5.0	5.1	5.2	5.2	4.8	5.4	5.0	4.6	5.2	5.0	5.3	4.9	5.1
$\mathcal{T}_2$	20	5.6	5.1	5.7	5.0	5.3	5.4	5.0	5.2	5.1	4.9	4.5	5.1	5.2	4.9	5.2
$\mathcal{T}_3$	20	3.3	4.5	5.1	3.0	4.9	5.0	2.8	5.0	4.8	2.9	4.3	4.9	2.8	4.6	5.1
$\mathcal{T}_4$	20	4.8	4.5	5.1	4.7	4.9	5.0	4.9	5.1	4.9	4.8	4.5	5.0	5.2	4.9	5.2
$\mathcal{H}_1$	20	2.7	3.9	4.4	2.7	4.6	4.6	2.7	4.8	4.6	2.8	4.2	4.9	2.8	4.5	5.1
$\mathcal{H}_2$	20	3.5	4.6	5.3	3.1	5.0	5.1	2.8	5.0	4.9	2.9	4.3	5.0	2.9	4.6	5.1
$\mathcal{H}_3$	20	5.0	4.7	5.3	4.8	5.1	5.2	4.9	5.2	5.0	4.9	4.5	5.0	5.2	4.9	5.2
Non-normal errors																
Statistics	$k_2 \downarrow \mu^2 \rightarrow$	$n=50$			$n=100$			$n=200$			$n=500$			$n=1000$		
		0	413	1000	0	413	1000	0	413	1000	0	413	1000	0	413	1000
$\mathcal{T}_2$	2	5.0	6.3	6.1	5.2	5.0	5.6	4.8	5.0	5.0	5.0	5.0	4.7	4.9	4.9	5.4
$\mathcal{T}_3$	2	0.0	5.0	5.4	0.0	3.8	5.0	0.0	3.4	4.3	0.1	2.8	3.8	0.0	2.2	4.0
$\mathcal{T}_4$	2	4.2	5.6	5.6	4.8	4.7	5.4	4.6	4.9	4.8	5.0	5.0	4.6	4.9	4.9	5.4
$\mathcal{H}_1$	2	0.0	4.4	4.8	0.0	3.5	4.7	0.0	3.3	4.1	0.0	2.8	3.7	0.0	2.1	3.9
$\mathcal{H}_2$	2	0.0	5.2	5.6	0.0	3.9	5.0	0.0	3.5	4.4	0.1	2.8	3.8	0.0	2.2	4.0
$\mathcal{H}_3$	2	4.5	5.9	5.7	5.0	4.9	5.5	4.7	5.0	4.9	5.0	5.0	4.6	4.9	4.9	5.4
$\mathcal{T}_2$	5	5.3	6.0	5.9	5.2	5.4	5.9	5.2	5.1	5.0	5.0	4.3	5.2	4.9	4.8	5.0
$\mathcal{T}_3$	5	0.3	5.3	5.2	0.3	4.8	5.6	0.3	4.3	4.5	0.2	3.6	4.7	0.2	3.7	4.5
$\mathcal{T}_4$	5	4.6	5.5	5.3	4.8	5.2	5.7	5.0	4.8	4.9	4.9	4.3	5.2	4.9	4.8	5.0
$\mathcal{H}_1$	5	0.2	4.7	4.6	0.2	4.5	5.2	0.2	4.2	4.3	0.2	3.5	4.6	0.2	3.7	4.5
$\mathcal{H}_2$	5	0.3	5.5	5.4	0.3	4.9	5.6	0.3	4.4	4.6	0.2	3.6	4.7	0.2	3.7	4.5
$\mathcal{H}_3$	5	4.8	5.6	5.4	4.9	5.2	5.8	5.1	4.9	4.9	4.9	4.3	5.2	4.9	4.8	5.0
$\mathcal{T}_2$	20	5.5	5.8	5.8	5.1	5.2	5.6	5.2	5.1	5.4	4.9	5.2	5.0	4.8	4.9	4.8
$\mathcal{T}_3$	20	3.5	5.2	5.1	2.8	4.8	5.1	3.1	4.8	5.1	2.7	4.8	4.8	2.6	4.6	4.7
$\mathcal{T}_4$	20	4.9	5.2	5.2	4.8	4.9	5.2	5.0	4.9	5.1	4.8	5.1	4.9	4.8	4.9	4.8
$\mathcal{H}_1$	20	2.8	4.6	4.5	2.6	4.5	5.0	2.9	4.6	5.0	2.7	4.8	4.8	2.6	4.6	4.7
$\mathcal{H}_2$	20	3.8	5.4	5.3	2.9	4.9	5.3	3.1	4.8	5.2	2.8	4.9	4.9	2.6	4.7	4.7
$\mathcal{H}_3$	20	5.1	5.5	5.3	4.9	5.0	5.3	5.1	5.0	5.2	4.9	5.1	5.0	4.8	4.9	4.8

Table 2. Rejection frequencies (in %) of the bootstrap DWH tests

Normal errors																
		$n = 50$			$n = 100$			$n = 200$			$n = 500$			$n = 1000$		
Statistics	$k_2 \downarrow \mu^2 \rightarrow$	0	413	1000	0	413	1000	0	413	1000	0	413	1000	0	413	1000
$\mathcal{T}_2^*$	2	6.3	6.4	7.1	5.7	5.2	5.7	6.0	5.6	5.7	6.2	5.6	5.6	6.1	5.5	5.6
$\mathcal{T}_3^*$	2	6.5	6.4	7.1	5.9	5.2	5.7	6.1	5.6	5.7	6.3	5.6	5.6	6.2	5.5	5.6
$\mathcal{T}_4^*$	2	6.3	6.4	7.1	5.7	5.2	5.7	6.0	5.6	5.7	6.2	5.6	5.6	6.1	5.5	5.6
$\mathcal{H}_1^*$	2	6.5	6.4	7.1	5.9	5.2	5.7	6.1	5.6	5.7	6.3	5.6	5.6	6.2	5.5	5.6
$\mathcal{H}_2^*$	2	6.5	6.4	7.1	5.9	5.2	5.7	6.1	5.6	5.7	6.3	5.6	5.6	6.2	5.5	5.6
$\mathcal{H}_3^*$	2	6.3	6.4	7.1	5.7	5.2	5.7	6.0	5.6	5.7	6.2	5.6	5.6	6.1	5.5	5.6
$\mathcal{T}_2^*$	5	6.8	5.9	5.9	6.6	6.0	6.7	7.2	5.6	5.4	6.2	5.8	5.6	7.6	5.8	5.1
$\mathcal{T}_3^*$	5	6.8	5.9	5.9	6.6	6.0	6.7	7.2	5.6	5.4	6.2	5.8	5.6	7.6	5.8	5.1
$\mathcal{T}_4^*$	5	6.8	5.9	5.9	6.6	6.0	6.7	7.2	5.6	5.4	6.2	5.8	5.6	7.6	5.8	5.1
$\mathcal{H}_1^*$	5	6.8	5.9	5.9	6.6	6.0	6.7	7.2	5.6	5.4	6.2	5.8	5.6	7.6	5.8	5.1
$\mathcal{H}_2^*$	5	6.8	5.9	5.9	6.6	6.0	6.7	7.2	5.6	5.4	6.2	5.8	5.6	7.6	5.8	5.1
$\mathcal{H}_3^*$	5	6.8	5.9	5.9	6.6	6.0	6.7	7.2	5.6	5.4	6.2	5.8	5.6	7.6	5.8	5.1
$\mathcal{T}_2^*$	20	6.0	5.9	6.1	7.1	6.3	6.7	6.6	6.3	6.3	6.7	5.2	5.1	7.1	5.8	5.5
$\mathcal{T}_3^*$	20	6.0	5.9	6.1	7.1	6.3	6.7	6.6	6.3	6.3	6.7	5.2	5.1	7.1	5.8	5.5
$\mathcal{T}_4^*$	20	6.0	5.9	6.1	7.1	6.3	6.7	6.6	6.3	6.3	6.7	5.2	5.1	7.1	5.8	5.5
$\mathcal{H}_1^*$	20	6.0	5.9	6.1	7.1	6.3	6.7	6.6	6.3	6.3	6.7	5.2	5.1	7.1	5.8	5.5
$\mathcal{H}_2^*$	20	6.0	5.9	6.1	7.1	6.3	6.7	6.6	6.3	6.3	6.7	5.2	5.1	7.1	5.8	5.5
$\mathcal{H}_3^*$	20	6.0	5.9	6.1	7.1	6.3	6.7	6.6	6.3	6.3	6.7	5.2	5.1	7.1	5.8	5.5
Non-normal errors																
		$n = 50$			$n = 100$			$n = 200$			$n = 500$			$n = 1000$		
Statistics	$k_2 \downarrow \mu^2 \rightarrow$	0	413	1000	0	413	1000	0	413	1000	0	413	1000	0	413	1000
$\mathcal{T}_2^*$	2	6.4	5.8	6.4	6.0	5.7	6.0	5.7	5.2	5.2	5.5	5.3	5.8	5.5	5.0	5.2
$\mathcal{T}_3^*$	2	6.4	5.8	6.4	5.9	5.7	6.0	5.9	5.2	5.2	5.7	5.3	5.8	5.7	5.0	5.2
$\mathcal{T}_4^*$	2	6.4	5.8	6.4	6.0	5.7	6.0	5.7	5.2	5.2	5.5	5.3	5.8	5.5	5.0	5.2
$\mathcal{H}_1^*$	2	6.4	5.8	6.4	5.9	5.7	6.0	5.9	5.2	5.2	5.7	5.3	5.8	5.7	5.0	5.2
$\mathcal{H}_2^*$	2	6.4	5.8	6.4	5.9	5.7	6.0	5.9	5.2	5.2	5.7	5.3	5.8	5.7	5.0	5.2
$\mathcal{H}_3^*$	2	6.4	5.8	6.4	6.0	5.7	6.0	5.7	5.2	5.2	5.5	5.3	5.8	5.5	5.0	5.2
$\mathcal{T}_2^*$	5	6.9	6.6	5.9	6.0	5.3	6.0	6.3	5.3	5.5	6.7	5.8	5.1	6.9	5.5	5.4
$\mathcal{T}_3^*$	5	6.9	6.6	5.9	6.0	5.3	6.0	6.3	5.3	5.5	6.7	5.8	5.1	6.9	5.5	5.4
$\mathcal{T}_4^*$	5	6.9	6.6	5.9	6.0	5.3	6.0	6.3	5.3	5.5	6.7	5.8	5.1	6.9	5.5	5.4
$\mathcal{H}_1^*$	5	6.9	6.6	5.9	6.0	5.3	6.0	6.3	5.3	5.5	6.7	5.8	5.1	6.9	5.5	5.4
$\mathcal{H}_2^*$	5	6.9	6.6	5.9	6.0	5.3	6.0	6.3	5.3	5.5	6.7	5.8	5.1	6.9	5.5	5.4
$\mathcal{H}_3^*$	5	6.9	6.6	5.9	6.0	5.3	6.0	6.3	5.3	5.5	6.7	5.8	5.1	6.9	5.5	5.4
$\mathcal{T}_2^*$	20	6.1	6.2	6.2	6.2	6.2	6.5	6.4	5.9	6.2	7.1	5.9	5.8	6.9	5.7	5.1
$\mathcal{T}_3^*$	20	6.1	6.2	6.2	6.2	6.2	6.5	6.4	5.9	6.2	7.1	5.9	5.8	6.9	5.7	5.1
$\mathcal{T}_4^*$	20	6.1	6.2	6.2	6.2	6.2	6.5	6.4	5.9	6.2	7.1	5.9	5.8	6.9	5.7	5.1
$\mathcal{H}_1^*$	20	6.1	6.2	6.2	6.2	6.2	6.5	6.4	5.9	6.2	7.1	5.9	5.8	6.9	5.7	5.1
$\mathcal{H}_2^*$	20	6.1	6.2	6.2	6.2	6.2	6.5	6.4	5.9	6.2	7.1	5.9	5.8	6.9	5.7	5.1
$\mathcal{H}_3^*$	20	6.1	6.2	6.2	6.2	6.2	6.5	6.4	5.9	6.2	7.1	5.9	5.8	6.9	5.7	5.1

APPENDIX

A. Auxiliary Lemmata and Proofs

This appendix presents some useful auxiliary lemmas and their proofs, as well as the proofs of the main theorems in the text.

A.1. Auxiliary Lemmata

Lemma A.1 *Suppose Assumption* **2.1** *is satisfied and that* $\pi \neq 0$ *is fixed. Under* H_0, *we have*:

$$(a)\qquad \| \mathbb{P}(\sqrt{n}\frac{(\tilde{\beta}-\hat{\beta})}{\tilde{\omega}_l} \leq x) - [\Phi(x) + \sum_{m=1}^{s-2} n^{-m/2} p^m_{\mathcal{T}_l}(x;\tilde{F},\pi)\Phi(x)] \|_\infty = o(n^{(s-2)/2}) \tag{A.1}$$

$$\| \mathbb{P}(\sqrt{n}\frac{(\tilde{\beta}-\hat{\beta})}{\hat{\omega}_j} \leq x) - [\Phi(x) + \sum_{m=1}^{s-2} n^{-m/2} p^m_{\mathcal{H}_j}(x;\tilde{F},\pi)\Phi(x)] \|_\infty = o(n^{(s-2)/2}) \tag{A.2}$$

$$(b)\qquad \| \mathbb{P}(\mathcal{T}_l \leq x) - [\Phi(x) + \sum_{m=1}^{s-2} n^{-m/2} p^m_{\mathcal{T}_l}(x;F,\boldsymbol{\pi})\Phi(x)]^2 \|_\infty = o(n^{(s-2)}) \quad , \tag{A.3}$$

$$\| \mathbb{P}(\mathcal{H}_j \leq x) - [\Phi(x) + \sum_{m=1}^{s-2} n^{-m/2} p^m_{\mathcal{H}_j}(x;\tilde{F},b_0,\boldsymbol{\pi})\Phi(x)]^2 \|_\infty = o(n^{(s-2)}) \tag{A.4}$$

for all l *and* j, *where* $p^m_{\mathcal{T}_l}$ *and* $p^m_{\mathcal{H}_j}$ *are polynomials in* x *with coefficients depending on moments of the distribution* F *of* $\mathcal{Q}_n$ *and* $\boldsymbol{\pi}$, *and* $\Phi(.)$ *is the cdf of a standard normal random variable.*

Lemma A.2 *Suppose Assumption* **2.1** *is satisfied. If for some* $\delta > 0$, *we have* $\mathbb{E}(\|Z_i\|^{2+\delta}, \|v_i\|^{2+\delta}) < \infty$, *then* $\mathbb{E}^*(|Z^*_{ji}v^*_{mi}|^{2+\delta})$ *is bounded a.s. under* H_0, *for all* $j = 1, \ldots, k$ *and* $m = 1, 2$; *where* Z^* *and* $v^* = [v^*_1 : v^*_2]$ *are the bootstrap draws from the empirical distribution of* Z *and the re-centered residuals* $\tilde{v} = [\tilde{v}_1 : \tilde{v}_2]$.

Corollary A.3 *Under the assumptions of Lemma* **A.2**, $\mathbb{E}^*(|Z^*_{ji}u^*_i|^{2+\delta})$ *is bounded a.s. under* H_0 *for all* $j = 1, \ldots, k$ *and* $m = 1, 2$; *where* $u^* = v^*_1 - v^*_2\beta$.

Lemma A.4 *Suppose Assumption* **2.1** *is satisfied. If for some* $\delta > 0$, $\mathbb{E}(\|Z_i\|^{4+\delta},\ \|v_i\|^{2+\delta}) < \infty$, *then under* H_0, *we have*:

$$\begin{pmatrix} Z^* u^*/\sqrt{n} \\ Z^* v_2^*/\sqrt{n} \\ \sqrt{n}(\frac{{W^*}'\mathbb{1}}{n} - \frac{W'\mathbb{1}}{n}) \end{pmatrix} \mid \hat{\mathscr{F}}_n \quad \xrightarrow{d} \quad \mathbf{N}\left[0, \begin{pmatrix} \mathbf{diag}(\sigma_u^2,\ \sigma_{v_2}) \otimes Q_Z & 0 \\ 0 & \Sigma_w \end{pmatrix}\right] \quad a.s.$$

where $W = (w_1,\ \ldots,\ w_n)$, $w_i = vech(Z_i Z_i')$, $W^* = (w_1^*,\ \ldots,\ w_n^*)$, $w_i^* = vech(Z_i^* {Z_i^*}') \in \mathbb{R}^{k(k+1)/2}$, $\Sigma_w = var(w_i)$, *and* $\mathbb{1}$ *is a* $(n$ *by* $1)$ *constant vector of ones,* $\hat{\mathscr{F}}_n = \{(Y_1',\ Z_1'),\ \ldots,\ (Y_n',\ Z_n')\}$.

Lemma A.5 *Suppose Assumption* **2.1** *is satisfied. If for some* $\delta > 0$, $\mathbb{E}(\|Z_i\|^{4+\delta},\ \|v_i\|^{2+\delta}) < \infty$, *then under* H_0, *we have*:

$$\frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta}^*)}{\tilde{\omega}_l^*} \mid \hat{\mathscr{F}}_n \ \xrightarrow{d}\ \mathbf{N}(0,\ 1), \quad \frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta}^*)}{\hat{\omega}_j^*} \mid \hat{\mathscr{F}}_n \xrightarrow{d} \mathbf{N}(0,\ 1) \quad a.s.$$

when $\pi = \pi_0/\sqrt{n}$, π_0 *is a* $(k$ *by* $1)$ *constant vector* (*and* $\pi_0 = 0$ *is allowed*), *where* $\tilde{\beta}^*$, $\hat{\beta}^*$, $\tilde{\omega}_l^*$, $\hat{\omega}_j^*$ *are the bootstrap counterparts of* $\tilde{\beta}$, $\hat{\beta}$, $\tilde{\omega}_l$, *and* $\hat{\omega}_j$ *defined in* (3.1)-(3.2).

A.2. Proofs

To shorten the exposition, note that the proofs of Theorem **4.1** and Lemma **A.2** are similar to those in Moreira et al. (2009) and are omitted.

Proof of Lemma **A.1** First, it is easy to see that $\mathcal{T}_l = c_{n_l}\left(\sqrt{n}\frac{(\tilde{\beta}-\hat{\beta})}{\tilde{\omega}_l}\right)^2$ and $\mathcal{H}_j = \left(\sqrt{n}\frac{(\tilde{\beta}-\hat{\beta})}{\hat{\omega}_j}\right)^2$ for all l and j, where $c_{n_l} = 1 + o(1)$. Now, we can observe $\sqrt{n}\frac{(\tilde{\beta}-\hat{\beta})}{\tilde{\omega}_l}$ and $\sqrt{n}\frac{(\tilde{\beta}-\hat{\beta})}{\hat{\omega}_j}$ as:

$$\begin{aligned} & \sqrt{n}\frac{(\tilde{\beta} - \hat{\beta})}{\tilde{\omega}_l} \\ = \quad & \sqrt{n}\frac{(y_2'y_2/n)^{-1}(y_2'y_1/n) - [(y_2'Z/n)(Z'Z/n)^{-1}(Z'y_2/n)]^{-1}[(y_2'Z/n)(Z'Z/n)^{-1}(Z'y_1/n)}{\sqrt{\frac{y_1'M_{y_2}y_1}{n}[(\frac{y_2'P_Zy_2}{n})^{-1} - (\frac{y_2'y_2}{n})^{-1}] - [(\frac{y_2'y_2}{n})^{-1}(\frac{y_2'y_1}{n}) - (\frac{y_2'P_Zy_2}{n})^{-1}(\frac{y_2'P_Zy_1}{n})]^2}} \\ = \quad & \sqrt{n}\frac{(y_2'y_2/n)^{-1}(y_2'u/n) - [(y_2'Z/n)(Z'Z/n)^{-1}(Z'y_2/n)]^{-1}[(y_2'Z/n)(Z'Z/n)^{-1}(Z'u/n)}{\sqrt{\frac{y_1'M_{y_2}y_1}{n}[(\frac{y_2'P_Zy_2}{n})^{-1} - (\frac{y_2'y_2}{n})^{-1}] - [(\frac{y_2'y_2}{n})^{-1}(\frac{y_2'y_1}{n}) - (\frac{y_2'P_Zy_2}{n})^{-1}(\frac{y_2'P_Zy_1}{n})]^2}} \\ = \quad & \sqrt{n}G(\bar{Q}_n) \overset{\text{under H}_0}{=} \sqrt{n}\left[G(\bar{Q}_n) - G(\mu)\right] \end{aligned} \tag{A.5}$$

$$
\begin{aligned}
&\sqrt{n}\frac{(\tilde{\beta}-\hat{\beta})}{\hat{\omega}_j} \\
= \quad & \sqrt{n}\frac{(y_2'y_2/n)^{-1}(y_2'y_1/n)-[(y_2'Z/n)(Z'Z/n)^{-1}(Z'y_2/n)]^{-1}[(y_2'Z/n)(Z'Z/n)^{-1}(Z'y_1/n)}{\sqrt{\frac{y_1'M_{y_2}y_1}{n}[(\frac{y_2'P_Zy_2}{n})^{-1}-(\frac{y_2'y_2}{n})^{-1}]}} \\
= \quad & \sqrt{n}\frac{(y_2'y_2/n)^{-1}(y_2'u/n)-[(y_2'Z/n)(Z'Z/n)^{-1}(Z'y_2/n)]^{-1}[(y_2'Z/n)(Z'Z/n)^{-1}(Z'u/n)}{\sqrt{\frac{y_1'M_{y_2}y_1}{n}[(\frac{y_2'P_Zy_2}{n})^{-1}-(\frac{y_2'y_2}{n})^{-1}]}} \\
= \quad & \sqrt{n}\tilde{G}(\bar{Q}_n) \overset{\text{under H}_0}{=} \sqrt{n}[\tilde{G}(\bar{Q}_n)-\tilde{G}(\mu)] \qquad \text{(A.6)}
\end{aligned}
$$

where $G(.)$ and $\tilde{G}(.)$ are real-valued Borel measurable functions in $\mathbb{R}^l$ such that $G(\mu) = G(\mathbb{E}(\mathcal{Q}_n)) = 0$ and $\tilde{G}(\mu) = \tilde{G}(\mathbb{E}[\mathcal{Q}_n]) = 0$ under H_0.[13] Since $\pi \neq 0$ is fixed (strong identification), all derivatives of $G(.)$ and $\tilde{G}(.)$ of order s and less are continuous in the neighborhood of $\mu = 0$. So, if further Assumption **2.1**-(b) holds, then (A.1)-(A.2) follow directly from Bhattacharya and Ghosh (1978, Theorem 2) and (A.3)-(A.4) hold by the definition of $\mathcal{T}_l$ and $\mathcal{H}_j$.

□

Proof of Lemma **A.4** Let $(c', d')'$ be a nonzero vector with $c = (c_1', c_2')' \in \mathbb{R}^{2k}$ and $d \in \mathbb{R}^{k(k+1)/2}$. Define

$$X_{ni} \quad = \quad c_1'Z_i^*u_i^*/\sqrt{n} + c_2'Z_i^*v_{2i}^*/\sqrt{n} + d'(w_i^* - \bar{w})/\sqrt{n}$$

where $[u_i^* : v_{2i}^*]$ is the i-th bootstrap draw of the (re-centered) residuals, and $\bar{w} = n^{-1}\sum_{i=1}^n w_i,\ w_i = vech(Z_iZ_i') \in \mathbb{R}^{k(k+1)/2}$, and $w_i^* = vech(Z_i^*Z_i^{*\prime}) \in \mathbb{R}^{k(k+1)/2}$.

We want to use the Cramér-Wold device. For this, it suffices to show X_{ni} satisfies all the conditions of the Liapunov Central Limit Theorem.

1. The first condition is obvious. Indeed, we have $\mathbb{E}^*(X_{ni}) = 0$ by the independence between Z^* and $[u_i^* : v_{2i}^*]$, and the fact that $\mathbb{E}^*\{[u_i^* : v_{2i}^*]\} = 0$.

2. The second condition is $\mathbb{E}^*(X_{ni}^2) < \infty$. Again, by the independence between Z^*

[13] This holds because $\mathbb{E}(y_2'u) = 0$ under H_0 and $\mathbb{E}(Z'u) = 0$ by Assumption **2.1**-(a).

and $[u_i^* : v_{2i}^*]$ and because u^* is uncorrelated with v_2^* under H_0, we have

$$\mathbb{E}^*(X_{ni}^2) = n^{-1}\left\{c_1'\left(\frac{Z'\tilde{u}\tilde{u}'Z}{n}\right)c_1 + c_2'\left(\frac{Z'\tilde{v}_2\tilde{v}_2'Z}{n}\right)c_2 + d'\tilde{\Sigma}_w d\right\} < \infty \quad a.s.,$$

where $\tilde{\Sigma}_w = n^{-1}\sum_{i=1}^n (w_i - \bar{w})(w_i - \bar{w})'$.

3. To check the final condition of the Liapunov Central Limit Theorem, it requires to show that $\lim_{n\to\infty}\sum_{i=2}^n \mathbb{E}^*(|X_{ni}|^{2+\delta}) = 0$ a.s. for some $\delta > 0$. Now, note that

$$\begin{aligned}
\sum_{i=2}^n \mathbb{E}^*[|X_{ni}|^{2+\delta}] &= n^{-\delta/2}n^{-1}\sum_{i=2}^n \mathbb{E}^*\left[|c_1'Z_i^*u_i^* + c_2'Z_i^*v_{2i}^* + d'(w_i^* - \bar{w})|^{2+\delta}\right] \\
&\leq C_1 n^{-\delta/2}\mathbb{E}^*\left[|c_1'Z_i^*u_i^*|^{2+\delta} + |c_2'Z_i^*v_{2i}^*|^{2+\delta} + |d'(w_i^* - \bar{w})|^{2+\delta}\right] \\
&\leq C_2 n^{-\delta/2}\left\{\sum_{j=1}^k |c_{1j}|^{2+\delta}\mathbb{E}^*[|Z_{ji}^*u_i^*|^{2+\delta}] + \sum_{j=1}^k |c_{2j}|^{2+\delta}\mathbb{E}^*[|Z_{ji}^*v_{2i}^*|^{2+\delta}] + \right\} \\
&+ C_2 n^{-\delta/2}\left\{\sum_{p=1}^{k(k+1)/2} |d_p|^{2+\delta}\mathbb{E}^*[|w_{pi}^* - \left(\frac{1}{n}\sum_{j=1}^n w_{ji}\right)|^{2+\delta}]\right\} \\
&= C_2 n^{-\delta/2}[A_1 + A_2 + A_3]
\end{aligned}$$

for large enough constants C_1 and C_2. From Lemma **A.2** and Corollary **A.3**, we have $A_1 = O(1)$ and $A_2 = O(1)$ a.s. If further $\mathbb{E}[\|Z_i\|^{4+\delta}] < \infty$, then e have $A_3 = O(1)$ a.s. Therefore, we get $\lim_{n\to\infty}\sum_{i=2}^n \mathbb{E}^*[|X_{ni}|^{2+\delta}] = 0$ a.s., and the last condition of the Liapunov Central Limit Theorem is satisfied. Lemma **A.4** is the Central Limit Theorem property once we realize that $p\lim_{n\to\infty}\left(\frac{Z'\tilde{u}\tilde{u}'Z}{n}\right) = \sigma_u^2 Q_Z$, $p\lim_{n\to\infty}\left(\frac{Z'\tilde{v}_2\tilde{v}_2'Z}{n}\right) = \sigma_{v_2}^2 Q_Z$, and $p\lim_{n\to\infty}(\tilde{\Sigma}_w) = \Sigma_w$.

□

Proof of Lemma **A.5** First, note that $\mathbb{E}^*(Z^{*\prime}Z^*/n) = Z'Z/n$, $\mathbb{E}^*(Z^{*\prime}u^*/n) = Z'\tilde{u}/n$, $\mathbb{E}^*(Z^{*\prime}v_2^*/n) = Z'\tilde{v}_2/n$, and $\mathbb{E}^*[(u^* : v_2^*)'(u^* : v_2^*)/n] = (\tilde{u} : \tilde{v}_2)'(\tilde{u} : \tilde{v}_2)/n$. So, the Markov law of large numbers entails that

$$\frac{Z^{*\prime}Z^*}{n} - \frac{Z'Z}{n} \mid \hat{\mathscr{F}}_n \overset{a.s.}{\to} 0,\ \frac{Z^{*\prime}u^*}{n} - \frac{Z'\tilde{u}}{n} \mid \hat{\mathscr{F}}_n \overset{a.s.}{\to} 0,\ \frac{Z^{*\prime}v_2^*}{n} - \frac{Z'\tilde{v}_2}{n} \mid \hat{\mathscr{F}}_n \overset{a.s.}{\to} 0$$
$$\frac{1}{n}(u^* : v_2^*)'(u^* : v_2^*) - \frac{1}{n}(\tilde{e} : \tilde{v}_2)'(\tilde{u} : \tilde{v}_2) \mid \hat{\mathscr{F}}_n \overset{a.s.}{\to} 0; \quad a.s.$$

Since $Z'Z/n \overset{p}{\to} Q_Z$, and $Z'\tilde{v}_2/n \overset{p}{\to} 0$, we have $Z'\tilde{u}/n \overset{p}{\to} 0$ and if H_0 holds, $(\tilde{u} : \tilde{v}_2)'(\tilde{u} : \tilde{v}_2)/n \overset{p}{\to} \mathbf{diag}(\sigma_u^2,\, \sigma_{v_2}^2)$. So, it is clear that: $Z^{*\prime}Z^*/n \overset{a.s.}{\to} Q_Z,\ Z^{*\prime}u^*/n \overset{a.s.}{\to} 0,$ $Z^{*\prime}v_2^*/n \overset{a.s.}{\to} 0,$ and $(u^* : v_2^*)'(u^* : v_2^*)/n \overset{a.s.}{\to} \mathbf{diag}(\sigma_u^2,\, \sigma_{v_2}^2)$ under H_0.

Now, from the above results along with Lemma **A.4** and the fact that $\pi = c/\sqrt{n}$, we have: $y_2^{*\prime}y_2^*/n = \boldsymbol{\pi}_0'(Z^{*\prime}Z^*/n^2)\boldsymbol{\pi}_0 + 2\boldsymbol{\pi}_0'Z^{*\prime}v_2^*/n^{3/2} + v_2^{*\prime}v_2^*/n \mid \hat{\mathscr{F}}_n \overset{a.s.}{\to} \sigma_{v_2}^2$ and $y_2^{*\prime}P_{Z^*}y_2^* = (y_2^{*\prime}Z^*/\sqrt{n})(Z^{*\prime}Z^*/n)^{-1}(Z^{*\prime}y_2^*/\sqrt{n}) \mid \hat{\mathscr{F}}_n \overset{d}{\to} (\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)$ a.s., where $\psi_{Zv_2} \sim \mathbf{N}(0,\, \sigma_{v_2}^2 Q_Z)$. Therefore, we have $\tilde{\beta}^* - \hat{\beta}^* = (y_2^{*\prime}P_{Z^*}y_2^*)^{-1}(y_2^{*\prime}P_{Z^*}u^*) - (y_2^{*\prime}y_2^*/n)^{-1}(y_2^{*\prime}u^*/n) = (y_2^{*\prime}P_{Z^*}y_2^*)^{-1}(y_2^{*\prime}P_{Z^*}u^*) + o_p(1) \mid \hat{\mathscr{F}}_n \overset{d}{\to} [(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)]^{-1}(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}\psi_{Zu}$ a.s. under H_0. Similarly, we can show that $\tilde{\omega}_l^{*2}/n,\, \hat{\omega}_j^{*2}/n \mid \hat{\mathscr{F}}_n \overset{a.s}{\to} \sigma_u^2[(\boldsymbol{\pi}_0\psi_{Zv_2} + Q_Z)'Q_Z^{-1}(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)]^{-1}$ a.s. for all l and j. Thus we get

$$\frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta})}{\tilde{\omega}_l^*} \mid \hat{\mathscr{F}}_n \;\overset{d}{\to}\; \frac{1}{\sigma_u}[(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)]^{-1/2}(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}\psi_{Zu}$$
$$\frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta}^*)}{\hat{\omega}_j^*} \mid \hat{\mathscr{F}}_n \;\overset{d}{\to}\; \frac{1}{\sigma_u}[(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)]^{-1/2}(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}\psi_{Zu} \quad a.s.$$

Moreover, ψ_{Zu} and ψ_{Zv_2} are independent and jointly normal under H_0 (see also Lemma **A.4**), thus we have $\frac{1}{\sigma_u}[(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)]^{-1/2}(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}\psi_{Zu} \mid \psi_{Zv_2} \sim \mathbf{N}(0,\, 1)$. Because the conditional distribution of $\frac{1}{\sigma_u}[(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}(\psi_{Zv_2}+Q_Z\boldsymbol{\pi}_0)]^{-1/2}(\psi_{Zv_2} + Q_Z\boldsymbol{\pi}_0)'Q_Z^{-1}\psi_{Zu}$, given ψ_{Zv_2}, does not depend on ψ_{Zv_2}, it is equal to the unconditional distribution. It follows that $\frac{\sqrt{n}(\tilde{\beta}^*-\hat{\beta})}{\tilde{\omega}_l^*} \mid \hat{\mathscr{F}}_n \overset{d}{\to} \mathbf{N}(0,\, 1)$ and $\frac{\sqrt{n}(\tilde{\beta}^*-\hat{\beta}^*)}{\hat{\omega}_j^*} \mid \hat{\mathscr{F}}_n \overset{d}{\to} \mathbf{N}(0,\, 1)$ for all $l = 2,\, 3,\, 4$ and $j = 1,\, 2,\, 3$.

□

Proof of Theorem **4.2** First, recall that $\mathcal{T}_l^* = \left(\frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta}^*)}{\tilde{\omega}_l^*}\right)^2$ and $\mathcal{H}_j^* = \left(\frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta}^*)}{\hat{\omega}_j^*}\right)^2$. By Lemma **A.5**, we have $\frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta}^*)}{\tilde{\omega}_l^*} \mid \hat{\mathscr{F}}_n \xrightarrow{d} \mathbf{N}(0,\, 1)$ and $\frac{\sqrt{n}(\tilde{\beta}^* - \hat{\beta}^*)}{\hat{\omega}_j^*} \mid \hat{\mathscr{F}}_n \xrightarrow{d} \mathbf{N}(0,\, 1)$ a.s. It is clear that $\mathcal{T}_l^* \mid \hat{\mathscr{F}}_n \xrightarrow{d} [\mathbf{N}(0,\, 1)]^2 \equiv \chi^2(1)$ and $\mathcal{H}_j^* \mid \hat{\mathscr{F}}_n \xrightarrow{d} [\mathbf{N}(0,\, 1)]^2 \equiv \chi^2(1)$ a.s. □

References

Ahn, S. , 1997. Orthogonality tests in linear models. Oxford Bulletin of Economics and Statistics 59, 83–186.

Baum, C., Schaffer, M. , Stillman, S. , 2003. Instrumental variables and GMM: Estimation and testing. Technical report, Department of Economics, Boston College, MA, Boston, USA MA.

Bhattacharya, R. N. , Ghosh, J. , 1978. On the validity of the formal edgeworth expansion. The Annals of Statistics 6, 434–451.

Doko Tchatoka, F. , Dufour, J.-M. , 2011a. Exogeneity tests and estimation in IV regressions. Technical report, Department of Economics, McGill University, Canada Montréal, Canada.

Doko Tchatoka, F., Dufour, J.-M., 2011b. On the finite-sample theory of exogeneity tests with possibly non-gaussian errors and weak identification. Technical report, Department of Economics, McGill University, Canada Montréal, Canada.

Dufour, J.-M., 1979. Methods for Specification Errors Analysis with Macroeconomic Applications PhD thesis University of Chicago. 257 + XIV pages. Thesis committee: Arnold Zellner (Chairman), Robert E. Lucas and Nicholas Kiefer.

Dufour, J.-M. , 1987. Linear Wald methods for inference on covariances and weak exogeneity tests in structural equations. In: I. B. MacNeill, G. J. Umphrey, eds, Advances in the Statistical Sciences: Festschrift in Honour of Professor V.M. Joshi's 70th Birthday. Volume III, Time Series and Econometric Modelling. D. Reidel, Dordrecht, The Netherlands, pp. 317–338.

Dufour, J.-M., 1997. Some impossibility theorems in econometrics, with applications to structural and dynamic models. Econometrica 65, 1365–1389.

Dufour, J.-M. , 2003. Identification, weak instruments and statistical inference in econometrics. Canadian Journal of Economics 36(4), 767–808.

Dufour, J.-M., 2006. Monte carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics in econometrics. Journal of Econometrics 138, 2649–2661.

Durbin, J., 1954. Errors in variables. Review of the International Statistical Institute 22, 23–32.

Engle, R. F., 1982. A general approach to Lagrange multiplier diagnostics. Journal of Econometrics 20, 83–104.

Farebrother, R. W., 1976. A remark on the Wu test. Econometrica 44, 475–477.

Guggenberger, P., 2010. The impact of a Hausman pretest on the size of the hypothesis tests. Econometric Theory 156, 337–343.

Hansen, C., Hausman, J. , Newey, W. , 2008. Estimation with many instrumental variables. Journal of Business and Economic Statistics 26(4), 398–422.

Hausman, J., 1978. Specification tests in econometrics. Econometrica 46, 1251–1272.

Hausman, J., Taylor, W. E., 1981. A generalized specification test. Economics Letters 8, 239–245.

Holly, A., 1982. A remark on Hausman's test. Econometrica 50, 749–759.

Hwang, H.-S., 1980. Test of independence between a subset of stochastic regressors and disturbances. International Economic Review 21, 749–760.

Kariya, T., Hodoshima, H., 1980. Finite-sample properties of the tests for independence in structural systems and LRT. The Quarterly Journal of Economics 31, 45–56.

Kotz, S., Balakrishnan, N., Johnson, N., 2000. Continuous Multivariate Distributions second edn. John Wiley & Sons, New York.

Li, J., 2006. The block bootstrap test of hausman's exogeneity in the presence of serial correlation. Economics Letters 91, 76–82.

Nakamura, A., Nakamura, M., 1981. On the relationships among several specification error tests presented by Durbin, Wu and Hausman. Econometrica 49, 1583–1588.

Nelson, C., Startz, R., 1990. The distribution of the instrumental variable estimator and its t-ratio when the instrument is a poor one. The Journal of Business 63, 125–140.

Newey, W. K., 1985a. Generalized method of moments speciÖcation testing. Journal of Econometrics 29, 229–256.

Newey, W. K., 1985b. Maximum likelihood specification testing and conditional moment tests. Econometrica 53(5), 1047–1070.

Revankar, N. S., 1978. Asymptotic relative efficiency analysis of certain tests in structural sysytems. International Economic Review 19, 165–179.

Revankar, N. S., Hartley, M. J., 1973. An independence test and conditional unbiased predictions in the context of simultaneous equation systems. International Economic Review 14, 625–631.

Reynolds, R. A. , 1982. Posterior odds for the hypothesis of independence between stochastic regressors and disturbances. International Economic Review 23(2), 479–490.

Ruud, P., 2000. An Introduction to Classical Econometric Theory. Oxford University Press, Inc., New York.

Ruud, P. A. , 1984. Tests of specification in econometrics. Econometric Reviews 3(2), 211–242.

Smith, R. J., 1983. On the classical nature of the wu-hausman statistics for independence of stochastic regressors and disturbance. Economics Letters 11, 357–364.

Smith, R. J., 1984. A note on likelihood ratio tests for the independence between a subset of stochastic regressors and disturbances. International Economic Review 25, 263–269.

Smith, R. J., Pesaran, M., 1990. A uniÖed approach to estimation and orthogonality tests in linear single-equation econometric models. Journal of Econometrics 44, 41–66.

Spencer, D. E., Berk, K. N., 1981. A limited-information specification test. Econometrica 49, 1079–1085.

Staiger, D., Stock, J. H., 1997. Instrumental variables regression with weak instruments. Econometrica 65(3), 557–586.

Thurman, W., 1986. Endogeneity testing in a supply and demand framework. Review of Economics and Statistics 68(4), 638–646.

Wang, J., Zivot, E., 1998. Inference on structural parameters in instrumental variables regression with weak instruments. Econometrica 66(6), 1389–1404.

Wong, K.-f. , 1996. Bootstrapping Hausman's exogeneity test. Economics Letters 53, 139–143.

Wooldridge, J. M. , 2009. Introductory Econometrics: A Modern Approach fourth edn. South-Western, USA.

Wu, D.-M., 1973. Alternative tests of independence between stochastic regressors and disturbances. Econometrica 41, 733–750.

Wu, D.-M., 1974. Alternative tests of independence between stochastic regressors and disturbances: Finite sample results. Econometrica 42, 529–546.

Wu, D.-M. , 1983a. A remark on a generalized specification test. Economics Letters 11, 365–370.

Wu, D.-M., 1983b. Tests of causality, predeterminedness and exogeneity. International Economic Review 24(3), 547–558.