THE UNIVERSITY of ADELAIDE

School of Economics

Working Papers

ISSN 2203-6024

Specification Tests with Weak and Invalid Instruments

Firmin Doko Tchatoka

Working Paper No. 2014-05
June, 2014

Specification tests with weak and invalid instruments

Firmin Doko Tchatoka*
The University of Adelaide

June 13, 2014

* School of Economics, The University of Adelaide, 10 Pulteney Street, Adelaide SA 5005, Tel:+618 8313 5540, Fax:+618 8223 1460, e-mail: firmin.dokotchatoka@adelaide.edu.au

ABSTRACT

We focus on the classical linear simultaneous equations models and study the sensitivity to instrument endogeneity of six alternative versions of Durbin-Wu-Hausman (DWH) tests of exogeneity. To address this issue, we consider two setups for instrument endogeneity: (i) *fixed instrument endogeneity*, i.e., the parameter that controls instrument invalidity is fixed (does not depend on the sample size); (ii) *local-to-zero instrument endogeneity*, i.e., this parameter goes to zero at rate $[n^{-1/2}]$ as the sample size n increases [similar to Staiger and Stock (1997)]. In the first setup, we show that all tests have size converging to 1, as the sample size increases, no matter how weak the instruments are. In the second setup, we provide a characterization of the null limiting distribution of the statistics, which clearly shows that all statistics converge asymptotically to non-degenerated noncentral χ^2 under the null hypothesis. This means that all tests have size greater than their nominal level asymptotically if the usual χ^2 critical values are used, despite the fact that *instrument endogeneity* vanishes as the sample size increases. We propose size correction based on bootstrap techniques. Our analysis of the proposed bootstrap tests provides some new insights. More precisely, we show that even for moderate instrument endogeneity, the bootstrap provides a first-order approximation of the asymptotic size of the DWH tests, no matter how weak the instruments are. We present a Monte Carlo experiment that confirms our theoretical findings.

Key words: DWH tests; weak instruments; exclusion restrictions; instrument endogeneity; bootstrap.
JEL classification: C3; C12; C15; C52.

1. Introduction

A basic problem in econometrics is estimating an equation of the form

$$y_1 = y_2\beta + u, \tag{1.1}$$

where y_2 and u might be correlated. It is well known that least squares estimators are biased and inconsistent when y_2 is not independent of u. If so, instrumental variable (IV) estimation, such as two-stage least squares (2SLS), usually provides a way to obtain a consistent estimator of β. Even though coefficients estimated by IV method may have interesting interpretations from the viewpoint of economic theory, IV procedure usually requires the availability of instruments, at least as great as the number of coefficients to be estimated, which: (i) satisfy the *exclusion restrictions* (that is, not contained in (1.1), thus uncorrelated with the errors u); and (ii) are *strong*, that is, highly correlated with the explanatory variable y_2. The IV literature often referred to condition (i) as the "strict exogeneity" of the instruments, while (ii) characterizes strong identification. The so-called "weak instruments" problem occurs when condition (ii) is violated, meaning that instrumental variables are weakly (poorly) correlated with y_2.

It is now well known that IV estimators can be imprecise and that inference procedures (such as tests and confidence sets) can be highly unreliable in the presence of weak instruments. This has led to a large literature on reliable inference in the presence of weak instruments, such as Anderson and Rubin (1949, AR-test), Kleibergen (2002, K-test), and Moreira (2003, CLR-test); see the reviews of Stock, Wright and Yogo (2002), Dufour (2003), Andrews and Stock (2007), and Mikusheva (2013), among others. However, most of these studies often impose the exclusion restrictions, thus ruling out the possibility that instrumental variables may not be strictly exogenous.

In recent years, concerns have been raised about the strict exogeneity assumption of the IVs; see Bound, Jaeger and Baker (1995), Hausman and Hahn (2005), Murray (2006), Doko Tchatoka and Dufour (2008), Bazzi and Clemens (2009, 2013), and Guggenberger (2011). Murray (2006) states: "in most IV applications, the instruments often arrive with a dark cloud of invalidity hanging overhead and researchers usually do not know whether their correlations with the error are exactly zero ". He suggests avoiding invalid instruments in IV procedures. However, since it is difficult to test the validity of all candidate instruments, it might seem that if we want to avoid invalid instruments, there is no hope in trying to use IV methods. Bound et al. (1995, Section 3) provide evidence on how a slight violation of instrument exogeneity can cause severe bias in IV estimates, especially when identification is weak. Hausman and Hahn (2005) show that even in large-sample, IV estimator can have a substantial bias even when the instruments are only slightly correlated with the error. Doko Tchatoka and Dufour (2008) and Guggenberger (2011) show that inference [on β in (1.1)] based on Anderson and Rubin (1949) AR-statistic, Kleibergen (2002) K-statistic, and Moreira (2003) CLR-statistic is unreliable from the viewpoint of size control when instruments violate “strict exogeneity”.

In this paper, we focus on linear IV regressions and study the sensitivity to instrument endogeneity of the standard Durbin (1954), Wu (1973, 1974), and Hausman (1978) tests of exogeneity (henceforth, DWH tests). We show that even a slight correlation between the IVs and the error substantially distorts the size of DWH tests, thus invalidating pretesting based on these statistics. We suggest size correction based on a bootstrap method, which works quite well even for relatively moderate values of instrument endogeneity. This suggests that reliable inference can still be conducted even when some instruments violate the exclusion restrictions. Several authors have recently adopted the same position and this paper attempts to make progress in this direction. Indeed, Imbens (2003) shows that bounds on average treatment effect in program evaluation can be recovered via a sensitivity analysis of the correlations between treatment and unobserved components of the outcomes. Small (2007) and Ashley (2009) show how the discrepancy between OLS and IV estimates can be used to estimate the degree of bias under any given assumption about the degree to which IVs violate the exclusion restrictions. Kiviet and Niemczyk (2007, 2012) show that the realizations of IV estimators based on strong but invalid instruments seem much closer to the true parameter values than those obtained from valid but weak instruments. Recently, Imbens et al. (2011) show that consistent point estimators can still be obtained in linear IV models when instruments violate the exclusion restrictions.

Before analyzing the impact of instruments endogeneity on DWH tests, it will be illuminating to first define instrumental variable (IV) estimation, and briefly recall the role of DWH procedures in IV estimation.

1.1. Instrumental variable estimation and tests of exogeneity

Stock and Trebbi (2003) define instrumental variable estimation as “the use of additional instrumental variables, not contained in the equation of interest, to estimate the unknown parameters of that equation”. Using the Stylometric Analysis, they show that the idea that instrumental variables estimation can be used to estimate the coefficient on an endogenous variable first appeared in Appendix B of Wright (1928). Since then, IV method is often referred to as a solution to the identification problem– the problem of identifying and estimating one or more coefficients of a system of simultaneous equations– and has been substantially studied in statistics; for example, see Anderson and Rubin (1949, 1950), Anderson (1951, 1976), Durbin (1954), Sargan (1958, 1983), Nagar (1959), Fisher(1966), Richardson (1968), Sawa (1969), Anderson and Kadane (1977), Dufour (1979), and Engle, Hendry and Richard (1982).

In economics, IV estimation is widely used in empirical research, maybe the most common statistical procedure besides linear regression. Usually, applied researchers want to know whether ordinary least squares (OLS) or an IV estimation method is appropriate to their study. From our knowledge, the idea to use a formal statistical test to decide whether one should use OLS or an IV estimation method first appeared in Durbin (1954). This was a major step forward in exogeneity testing, and has later been more formalized in Wu (1973, 1974), Revankar and Hartley (1973), Farebrother (1976), Revankar (1978), and Hausman (1978). Research on tests of exogeneity is now considerable; see Dufour (1979, 1987), Hwang (1980, 1985), Kariya and Hodoshima (1980), Hausman and Taylor (1981), Spencer and Berk (1981), Nakamura and Nakamura (1981, 1985), Engle (1982), Holly (1982, 1983b, 1983a), Holly and Monfort (1983), Reynolds (1982), Smith (1983, 1984, 1985), Thurman (1986), Smith and Pesaran (1990), Ruud (1984, 2000), Newey (1985a, 1985b), and Meepagala (1992), among others.

However, most of these studies usually assume that IVs are strong and strictly exogenous, thus leaving out issues related to *weak* and *invalid* instruments. There is now a growing literature about the reliability of the DWH procedures in the presence of weak instruments. However, there is little focus on the behavior of these tests when instrument violate the exclusion restrictions. Section 1.2 summarizes some recent studies on the behavior of DWH tests in the presence of weak instruments.

1.2. Impact of weak instruments on tests of exogeneity

In recent years, the literature on weak instruments has raised concerns about the reliability of tests of exogeneity, such as DWH procedures, which rely mainly on IV estimators; for examples, see Staiger and Stock (1997), Ahn (1997), Guggenberger (2010), Hahn, Ham and Moon (2010), Doko Tchatoka and Dufour (2011a, 2011b, 2014), Chmelarova and Hill (2010), Kiviet and Niemczyk (2007, 2012), Kiviet and Pleus (2012), Kiviet (2013), and Doko Tchatoka (2013).

Hahn et al. (2010) provide a framework where a Hausman-type test is used to assess the exogeneity of a subset of instruments excluded from the structural equation of interest. The proposed test is essentially a test of exclusion restrictions [similar to Sargan (1958), Basmann (1960), and Hansen (1982)] and it is valid even when IVs are weak. By contrast, in this paper, we focus on testing the exogeneity of the regressor y_2 that is included in the structural equation of interest, by allowing for the possibility that some instruments violate the exclusion restrictions.

Staiger and Stock (1997) show under the exclusion restrictions that the null limiting distributions of Hausman (1978) type statistics depend on the concentration parameter, which usually determines the strength of the IVs. Meanwhile, Wu (1973, 1974) T_2 and T_4 statistics are asymptotically pivotal under exogeneity, even when instruments are weak. Doko Tchatoka and Dufour (2011b) provide a characterization of the finite-sample distributions of DWH statistics, allowing for the possibility of identification failure and non-Gaussian errors. Under the maintained assumption of exclusion restrictions, they show that the statistics T_1, T_2, and T_4 by Wu (1973, 1974) are pivotal under exogeneity without identifying assumptions, even when the errors are non-Gaussian. However, Wu (1973, 1974) T_3 and alternative Hausman (1978) type statistics do not share this property, but they are boundedly pivotal with or without non-Gaussian errors, no matter how weak the IVs are. This means that if the exclusion restrictions are satisfied, all DWH tests are valid (in the sense that their level is controlled) in the presence of weak instruments if the usual F or asymptotic χ^2 critical values are applied in the inference. However, applying asymptotic χ^2 critical values to Wu (1973, 1974) T_3 and Hausman (1978) statistic can lead to overly conservative procedures when instruments are very weak. Size correction is nonetheless possible by using the exact Monte Carlo tests method of Dufour (2006) [see Doko Tchatoka and Dufour (2011b)]. Doko Tchatoka (2013) shows that bootstrapping substantially improves the size of DWH tests, as well as inference on the structural parameter β in (1.1) [also, see Wong (1996, 1997) and Li (2006)]. However, the bootstrap validity is established under the exclusion restrictions. Not much is known about the properties of bootstrap when IVs violate "strict exogeneity", which is the main focus of this study.

To address these issues, we consider the case where IVs may be arbitrary both weak and invalid. We show that all DWH tests considered have size converging to 1 if instrument endogeneity is fixed (does not depend on the sample size), no matter how weak the IVs are. Even when the parameter that controls instrument endogeneity goes to zero as the sample size increases, our results indicate that all tests have size greater than the nominal level asymptotically. We propose size correction based on bootstrap techniques. Our analysis of the proposed bootstrap tests provides some new insights. In particular, we show that even for relatively mild correlations between the IVs and the error, the bootstrap provides a first-order approximation of the asymptotic distributions of DWH statistics, whether IVs are strong or weak. As a result, all DWH tests have asymptotically a correct level if the bootstrap critical values are used in the inference, despite the fact that instruments violate the exclusion restrictions. We present a Monte Carlo experiment that confirms our theoretical findings.

The remainder of this paper is organized as follows. Section 2 formulates the model and presents the statistics. Section 3 characterize the asymptotic behavior of the DWH test with locally invalid instruments. Section 4 presents the bootstrap technique, while Section 5 presents the Monte Carlo experiment. Conclusions are drawn in Section 6 and proofs are presented in the Appendix.

Throughout this paper, I_q stands for the identity matrix of order q. For any full rank $n \times m$ matrix A, $P_A = A(A'A)^{-1}A'$ is the projection matrix on the space spanned by A, $M_A = I_n - P_A$. The notation $vec(A)$ is the $nm \times 1$ dimensional column vectorization of A and $B > 0$ for a squared matrix B means that B is positive definite (p.d.). The convergence in probability is symbolized by "$\xrightarrow{p}$" while "$\xrightarrow{d}$" stands for convergence in distribution. The usual (stochastic) orders of magnitude are denoted by $O_p(.)$ and $o_p(.)$. Finally, $\|U\| = [tr(U'U)]^{\frac{1}{2}}$ denotes the usual Euclidian or Frobenius norm for a matrix U.

2. Framework

We consider a linear structural equation of the form

$$y_1 = y_2\beta + u, \tag{2.1}$$

where $y_1 \in \mathbb{R}^n$ is a vector of observations on a dependent variable, $y_2 \in \mathbb{R}^n$ is a vector of observations on an explanatory variable that may be correlated with the structural disturbance $u = (u_1, \ldots, u_n)' \in \mathbb{R}^n$, β is an unknown fixed scalar coefficient (structural parameter). Further, we assume that y_2 obeys the equation

$$y_2 = Z\pi + v_2, \tag{2.2}$$

where Z is a $n \times k$ matrix of instruments, $v_2 = (v_{21}, \ldots, v_{2n})' \in \mathbb{R}^n$ is a vector of reduced-form disturbance, and $\pi \in \mathbb{R}^k$ is an unknown reduced-form coefficient vector. If Z is exogenous (not correlated with u), the usual necessary and sufficient condition for the identification of β in model (2.1)-(2.2) is $\pi \neq 0$. If π is close to zero, β is ill-determined by the data, a situation often called "weak identification" in the literature; see Staiger and Stock (1997), Stock et al. (2002), Dufour (2003), Andrews and Stock (2007), and Dufour and Hsiao (2008), among others.

The exogeneity hypothesis of y_2 in model (2.1)-(2.2) can be expressed as

$$H_0 : cov(y_{2t}, u_t) = 0,\ t = 1, \ldots, n. \tag{2.3}$$

In this paper, we are concerned with the properties of Durbin-Wu-Hausman tests[1] that are usually used to assess H_0 : (i) without any identifying assumptions of the model (i.e., Z may be arbitrary weak), and (ii) when the "exclusion restrictions" may be violated (i.e., Z may be correlated with the disturbance u). To investigate this, we consider six alternative

[1]See Durbin (1954), Wu (1973, 1974), and Hausman (1978).

versions of the DWH statistics, namely, the three statistics T_l $(l = 2,3,4)$ by Wu (1973, 1974) and three alternative Durbin-Hausman statistics $(DW_j,\ j = 1,2,3)$. All statistics can be expressed in a unified way [see Staiger and Stock (1997) and Doko Tchatoka and Dufour (2011b)] as:

$$T_l \;=\; \kappa_l(\tilde{\beta}-\hat{\beta})^2/\tilde{\omega}_l^2, \quad DW_j = n(\tilde{\beta}-\hat{\beta})^2/\hat{\omega}_j^2,\, l=2,3,4,\ j=1,2,3, \tag{2.4}$$

where $\hat{\beta} = (y_2'y_2)^{-1}y_2'y_1$ and $\tilde{\beta} = (y_2'P_Z y_2)^{-1}y_2'P_Z y_1$ are the OLS and IV estimators of β, respectively, $\kappa_2 = n-2$, $\kappa_3 = \kappa_4 = n-1$, and

$$\begin{aligned}
\tilde{\omega}_2^2 &= \tilde{\sigma}_2^2\hat{\Delta},\, \tilde{\omega}_3^2 = \tilde{\sigma}^2\hat{\Delta},\, \tilde{\omega}_4^2 = \hat{\sigma}^2\hat{\Delta},\\
\hat{\omega}_1^2 &= \tilde{\sigma}^2\hat{\omega}_{IV}^{-1} - \hat{\sigma}^2\hat{\omega}_{LS}^{-1},\, \hat{\omega}_2^2 = \tilde{\sigma}^2\hat{\Delta},\, \hat{\omega}_3^2 = \hat{\sigma}^2\hat{\Delta},\\
\hat{\Delta} &= \hat{\omega}_{IV}^{-1} - \hat{\omega}_{LS}^{-1},\, \hat{\omega}_{IV} = \frac{1}{n}(y_2'P_Z y_2),\, \hat{\omega}_{LS} = \frac{1}{n}(y_2'y_2),\\
\tilde{\sigma}_2^2 &= \hat{\sigma}^2 - (\tilde{\beta}-\hat{\beta})^2/\hat{\Delta},\, \hat{\sigma}^2 = \frac{1}{n}(y_1 - y_2\hat{\beta})'(y_1 - y_2\hat{\beta}),\\
\tilde{\sigma}^2 &= \frac{1}{n}(y_1 - y_2\tilde{\beta})'(y_1 - y_2\tilde{\beta}).
\end{aligned} \tag{2.5}$$

Engle (1982) and Smith (1983) show that T_2, T_4, and DW_3 are score (LM) statistics, while T_3, DW_1, and DW_2 are quasi-Wald statistics.[2] The regression interpretation is also provided in Hausman (1978), Dufour (1979, 1987), and Davidson and Mackinnon (1993) for the statistic T_2, and in Doko Tchatoka and Dufour (2011b) for all statistics. Finite-sample distributions are available in Wu (1973) for T_2 when the errors are Gaussian and IVs are strong and valid. Doko Tchatoka and Dufour (2011b) provide a characterization of the finite-sample distributions of all statistics [including the Hausman (1978) statistic], allowing for the possibility of identification failure and non-Gaussian errors. However, all these studies impose the exclusion restrictions. Here, we are concerned with what happens to DWH procedures when instruments violate "strict exogeneity", at least locally.

In the remainder of the paper, let $Y = [y_1 : y_2] = [Y_1, \ldots, Y_n]' \in \mathbb{R}^{n\times 2}$ denotes the matrix of endogenous variables and define $\mathscr{X}_n = \{(Y_1', Z_1'), \ldots, (Y_n', Z_n')\}$ is the observed sample. We define $Y_t \in \mathbb{R}^2$ and $Z_t \in \mathbb{R}^k$ as the tth rows of Y and Z respectively, written as column vectors, and similarly for other random matrices. We will make the following assumptions on the model variables.

Assumption A *For some fixed vector b in $\mathbb{R}^k$, we have*:

$$u_t = Z_t'b + e_t,\ t = 1, \ldots, n, \tag{2.6}$$

[2]See Smith (1983) for the score interpretation (eqs. 6 and 9) and for the quasi-Wald interpretation (eqs. 7, 8 and 10).

where $\mathsf{E}[e_t \,|\, Z_t] = 0$ *for all* $t = 1, \ldots, n$.

Assumption B *(i)* $\{(V_t = (e_t, v_{2t})', Z_t) : t \leq n\}$ *are i.i.d. across* $t \leq n$ *and* n; *(ii)* $\mathbb{E}(V_t) = 0$ *for all* $t = 1, \ldots, n$.

Assumption C *When the sample size n converges to infinity, the following convergence hold jointly*:
(i) $n^{-1}[e : v_2]'[e : v_2] \xrightarrow{p} \Sigma_V = \begin{pmatrix} \sigma_e^2 & \sigma_{v_2 e} \\ \sigma_{v_2 e} & \sigma_{v_2}^2 \end{pmatrix}$ *and* $n^{-1}Z'Z \xrightarrow{p} Q_Z$, $n^{-1}Z'[e : v_2] \xrightarrow{p} 0$; *where* Σ_V *and* Q_Z *are p.d.*;
(ii) $n^{-1/2}\sum_{t=1}^{n}(Z_t'e_t, Z_t'v_{2t}, v_{2t}e_t - \sigma_{v_2 e})' \xrightarrow{d} \Psi = (\psi_{Ze}', \psi_{Zv_2}', \psi_{v_2 e})'$, *where* $vec(\Psi) \sim N(0, \Omega)$.

The decomposition (2.6) in Assumption A defines e_t through its orthogonality with Z_t for any $t = 1, \ldots, n$. It allows substantial conditional heterogeneity in the distribution of the structural disturbances u_t, $t = 1, \ldots, n$. For example, the conditional mean of u_t, given Z_t, may depends on Z_t as long as $b \neq 0$. If Z_t has finite second moments and $cov(Z_t)$ is nonsingular[3], we have $cov(Z_t, u_t) = cov(Z_t)b = 0$ if, and only if, $b = 0$. Which means that Z_t satisfies the exclusion restrictions only when $b = 0$ in (2.6). When $b \neq 0$, $cov(Z_t, u_t) \neq 0$ and Z_t, $t = 1, \ldots, n$, do not constitute valid instruments. The usual tests of exclusion restrictions, such as Sargan (1958) and Basmann (1960) tests, typically test the null hypothesis that $b = 0$ in (2.6); see Staiger and Stock (1997). Recently, Hahn et al. (2010) consider the case where $Z = [Z_1 : Z_2]$, $Z_1 : k_1 \times 1$ is exogenous and possibly weak, while $Z_2 : k_2 \times 1$ is strong but may violate the exclusion restrictions. They then develop a Hausman-type test for testing whether Z_2 satisfies the exclusion restrictions, conditionally on Z_1 being exogenous. In other words, the problem considered can be formulated as testing the subset null hypothesis

$$H_{b_2} : b_2 = 0 \; s.t. \; Rb = 0, \tag{2.7}$$

where $R = \begin{bmatrix} 0 & : & I_{k_2} \end{bmatrix} : k_2 \times k$ and $b = (b_1' : b_2')'$. The proposed test is robust to identifying assumptions in the sense that it has correct level asymptotically even when Z_1 is weak. However, testing whether the exclusion restrictions are satisfied is fundamentally different from testing the exogeneity of an included regressor in the structural equation (2.1), where the instrumental variables may be arbitrary invalid ($b \neq 0$). To check this, assume that Z_t, $t = 1, \ldots, n$ are i.i.d. observations, each with finite second moments. Then, testing H_0 in (2.3) is equivalent to test $b'cov(Z_t)\pi + \sigma_{v_2 e} = 0$, which is different from testing $b = 0$

[3]Following Anderson (1971) and Muirhead (2005), we define $cov(X) = \mathbb{E}[(X - \mu_X)(X - \mu_X)']$ for any $q \times 1$ vector X, where $\mu_X = \mathbb{E}(X)$.

[as in Sargan (1958) and Basmann (1960)] or $b_2 = 0$ [as in Hahn et al. (2010) and Ruud (2000)]. Assumption A is not new in the IV literature. Staiger and Stock (1997) analyze the properties of the Sargan (1958) and Basmann (1960) tests of exclusion restrictions under local-to-zero instrument endogeneity, that is, $b = b_0/\sqrt{n}$, $b_0 \in \mathbb{R}^k$ is fixed. Their show that these are size distorted even when $\pi \neq 0$ is fixed (strong instruments). Doko Tchatoka and Dufour (2008) make a similar assumption[4] and show that the Anderson and Rubin (1949) (AR) and Kleibergen (2002) (K) tests are size distorted even when instruments only violate the exclusion restrictions slightly.

Assumption B-(i) states that the errors and IVs are random and i.i.d across $i \leq n$ and n, while Assumption B-(ii) is the usual zero mean assumption of the errors e_t and v_{2t}. Under Assumption B-(ii) and if Z_t has finite first moment, it is easy to see that $\mathbb{E}(u_t) = \mathbb{E}(Z_t')b$, which is not necessary zero when $b \neq 0$.

The convergence results in Assumption C are easy to interpret. Assumption C-(i) is the weak law of large numbers (WLLN) property of (e, v_2) and Z, where the IVs are assumed asymptotically uncorrelated with both e and v_2. Assumption C-(ii) is the central limit theorem (CLT) property; see Staiger and Stock (1997), Kleibergen (2002, 2004), and Kleibergen and Mavroeidis (2011), and Doko Tchatoka (2013). Even though Assumption C-(ii) allows for heteroskedastic or serially correlated errors, without any loss of generality, our analysis focus on the case where $\Omega = \begin{pmatrix} \Sigma & 0 \\ 0' & \sigma_e^2 \sigma_{v_2}^2 \end{pmatrix}$ and $\Sigma = \Sigma_V \otimes Q_Z$ under H_0 (exogeneity). Note that the estimators of the covariance matrix of $\tilde{\beta} - \hat{\beta}$ in (2.5) are obtained under homoskedastic errors; for example, see Wu (1973, eq.(1.3)) and Hausman (1978, eq.(1.1b)). So, the restriction to homoskedastic errors is a simplification in the same spirit to that of the original derivation of DWH statistics [also, see Guggenberger, Kleibergen, Mavroeidis and Chen (2012) and Doko Tchatoka (2014) for a similar assumption]. However, there is no impediment to allowing for heteroskedastic or serially correlated errors; for example, see Baum, Schaffer and Stillman (2003). If so, the results of this paper do not change qualitatively.

We will now study the sensitivity to instrument endogeneity of DWH tests. Section 3 presents the results.

3. Asymptotic results

In this section, we wish to study the sensitivity to instrument endogeneity of the DWH tests of exogeneity. For this, we consider the following two setups for instrument endogeneity:

[4] Also, see Guggenberger (2011) and Berkowitz, Caner and Fang (2008, 2012) for local-to-zero endogeneity framework, i.e., $b = b_0/\sqrt{n}$, b_0 is a $k \times 1$ constant vector.

(i) $b \neq 0$ is fixed (fixed instrument endogeneity), and (ii) $b = b_0/\sqrt{n}$, where b_0 is a $k \times 1$ constant vector (local-to-zero invalid instruments); similar to Staiger and Stock (1997). In the latter setup, $b_0 = 0$ means exogenous instruments, while $b_0 \neq 0$ entails that Z violates the exclusion restrictions.

We can now prove the following two lemmas on the asymptotic behavior of the DWH statistics.

Lemma 3.1 *Suppose that Assumptions A-C and H_0 are satisfied. If further $\pi \neq 0$ is fixed, then for all $l = 2, 3, 4$ and $j = 1, 2, 3$, we have*:

$$(a) \quad T_l \xrightarrow{p} +\infty, DW_j \xrightarrow{p} +\infty \text{ if } b \neq 0 \text{ is fixed;}$$
$$(b) \quad T_l \xrightarrow{d} \chi^2(1; \|\bar{\tau}_{b_0}\|^2), DW_j \xrightarrow{d} \chi^2(1; \|\bar{\tau}_{b_0}\|^2) \text{ if } b = b_0/\sqrt{n}, b_0 \text{ is fixed (possibly zero),}$$

where b is the $k \times 1$ vector defined in (2.6) *and* $\bar{\tau}_{b_0} = \frac{1}{\sigma_e}[(\pi' Q_Z \pi)^{-1} - (\pi' Q_Z \pi + \sigma^2_{v_2})^{-1}]^{1/2} \pi' Q_Z b_0$.

Lemma 3.2 *Suppose that Assumptions A-C and H_0 are satisfied. If further $\pi = \pi_0/\sqrt{n}$, where $\pi_0 \in \mathbb{R}^k$ is fixed, then for all $l = 2, 3, 4$ and $j = 1, 2, 3$, we have*:

$$(a) \quad T_l \xrightarrow{p} +\infty, DW_j \xrightarrow{p} +\infty; \text{ if } b \neq 0 \text{ is fixed;}$$
$$(b) \quad T_2, T_4, DW_3 | \psi_{Zv_2} \xrightarrow{d} \chi^2(1; \|\bar{\nu}_{b_0}\|^2), T_3, DW_1, DW_2 | \psi_{Zv_2} \xrightarrow{d} \frac{\chi^2(1; \|\bar{\nu}_{b_0}\|^2)}{1+\chi^2(1; \|\bar{\nu}_{b_0}\|^2)} \text{ if } b = \frac{b_0}{\sqrt{n}},$$
$$\bar{\nu}_{b_0} = \frac{1}{\sigma_e}[(\pi_0 + Q_Z^{-1}\psi_{Zv_2})' Q_Z (\pi_0 + Q_Z^{-1}\psi_{Zv_2})]^{-\frac{1}{2}} (\pi_0 + Q_Z^{-1}\psi_{Zv_2})' Q_Z b_0.$$

Lemma **3.1** characterizes the null limiting distributions of DWH statistics when IVs are strong ($\pi \neq 0$ is fixed) possibly invalid. If instrument endogeneity is fixed and different from zero, all statistics diverge. However, the statistics converge asymptotically to non-degenerated noncentral χ^2 distributions when instrument endogeneity is local-to-zero.

Lemma **3.2** extends the results of Lemma **3.1** to weak values of π, i.e., when $\pi = \pi_0/\sqrt{n}$, where $\pi_0 \in \mathbb{R}^k$ is fixed. We see that the conclusions of Lemma **3.1** remain valid even when $\pi_0 = 0$ (irrelevant instruments).

We can now prove the following two theorems on the size of the DWH tests.

Theorem 3.3 *Suppose that Assumptions A-C and H_0 are satisfied. If further $b \neq 0$ is fixed, then, for all $l = 2, 3, 4$ and $j = 1, 2, 3$, we have*:

$$\lim_{n\to\infty} \mathbb{P}(T_l > \chi^2_{1,\alpha}) \;=\; 1, \lim_{n\to\infty} \mathbb{P}(DW_j > \chi^2_{1,\alpha}) = 1 \text{ for all values of } \pi \in \mathbb{R}^k.$$

Theorem 3.4 *Suppose that Assumptions A-C and H_0 are satisfied. If further $b = b_0/\sqrt{n}$,* *where* $b_0 \neq 0$ *is fixed*, *then we have*:

$$\lim_{n\to\infty} \mathbb{P}(T_l > \chi^2_{1,\alpha}) \;=\; p_\alpha > \alpha,\ \lim_{n\to\infty} \mathbb{P}(DW_j > \chi^2_{1,\alpha}) = p_\alpha > \alpha$$

when $\pi \neq 0$ *is fixed, where* $p_\alpha = 1 - G_1(\chi^2_{1,\alpha};\ \|\bar{\tau}_{b_0}\|^2)$; *and*

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(T_l > \chi^2_{1,\alpha}) &= \tilde{p}_\alpha > \alpha,\ \lim_{n\to\infty} \mathbb{P}(DW_3 > \chi^2_{1,\alpha}) = \tilde{p}_\alpha > \alpha,\ l = 2,4 \\ \lim_{n\to\infty} \mathbb{P}(\hat{\omega}^2 T_3 > \chi^2_{1,\alpha}) &= \tilde{p}_\alpha > \alpha,\ \lim_{n\to\infty} \mathbb{P}(\hat{\omega}^2 DW_j > \chi^2_{1,\alpha}) = \tilde{p}_\alpha > \alpha,\ j = 1,2\end{aligned}$$

when $\pi = \pi_0/\sqrt{n}$, $\pi_0 \in \mathbb{R}^k$ *is fixed* ($\pi_0 = 0$ *is allowed*), *where* $\hat{\omega}^2 = \tilde{\sigma}^2/\hat{\sigma}^2$ *and* $\tilde{p}_\alpha = 1 - \mathbb{E}[G_1(\chi^2_{1,\alpha};\ \|\bar{v}_{b_0}\|^2)]$.

Theorem **3.3** follows directly from Lemma **3.1**-(a) and Lemma **3.2**-(a), therefore, the proof is omitted. It shows that if the usual asymptotic χ^2 critical values are used in the inference, all DWH tests have size converging to 1 when instrument endogeneity is fixed, no matter how weak these (invalid) instruments are.

Theorem **3.4** is concerned with the case where instrument endogeneity is local-to-zero. The results show that all DWH tests, asymptotically, have size greater than their nominal level for all values of π, if the asymptotic χ^2 critical values are used in the inference. Although Theorem **3.4** excludes the case of strict exogeneity ($b_0 = 0$), there is no impediment to expanding to it. If so, it is straightforward to see from Theorem **3.4** that the size of all tests converges to their nominal level α under exogeneity.[5]

More interestingly, when $\pi \neq 0$ is fixed (strong IVs) and $b = b_0/\sqrt{n}$, we observe that the non-centrality parameter in the asymptotic distributions of Lemma **3.1**-(b), $\|\bar{\tau}_{b_0}\|^2$, increases with instrument endogeneity. This means that the size distortions of the tests increase with instrument endogeneity. Because the asymptotic distributions of the statistics are continuous functions of b_0 for all values of π, hence, the size distortions of the tests also increase with instrument endogeneity for weak values of π (for example, $\pi = \pi_0/\sqrt{n}$ where π_0 is fixed).

Finally, note that the results of Theorem **3.3** and Theorem **3.4** are intuitive. Indeed, under Assumptions A-C, we have $cov(y_{2t}, u_t) = b'cov(Z_t)\pi + \sigma_{v_2 e}$. So, as long as the exclusion restrictions are satisfied ($b = 0$), the exogeneity hypothesis of y_2 can be expressed as $H_0 : \sigma_{v_2 e} = 0$. The DWH statistics in (2.4) typically test the null hypothesis $H_0 : \sigma_{v_2 e} = 0\ s.t.\ b = 0$. Now, if $b \neq 0$, it is clear that testing whether $cov(y_{2t}, u_t) = 0$ is not equivalent to test whether $\sigma_{v_2 e} = 0$i since $cov(Z_t)$ is p.d. Even when $b = O(n^{-1/2})$ or

[5]Without correction of the critical values, the upper bounds on the size of T_3, DW_1 and DW_2 are not the same as those of Theorem **3.4** when $\pi = \pi_0/\sqrt{n}$. This is because the null limiting distributions of the latter statistics are bounded above by the usual χ^2 distributions when $b_0 = 0$ (exogenous instrument); see Doko Tchatoka and Dufour (2011a, 2011b).

$\pi = O(n^{-1/2})$ as in Theorem **3.4**, testing $H_0 : \sigma_{v_2 e} = 0$ may substantially differ from testing $b' cov(Z_t)\pi + \sigma_{v_2 e} = 0$, especially if the sample size is moderate or small. So, the DWH statistics in (2.4) do not have correct size even when b is in the neighborhood of zero (see the Monte Carlo experiment in Section 5).

We will now investigate whether bootstrapping can improve the size property of DWH tests when instruments violate the exclusion restrictions, at least locally.

4. Bootstrapping with invalid instruments

In this section, we propose a bootstrap procedure for DWH tests when instruments violate "strict exogeneity". To proceed, let $\hat{\beta} = (y_2' y_2)^{-1} y_2' y_1$ and $\hat{\pi} = (Z'Z)^{-1} Z' y_2$ denote the OLS estimators of β and π in (2.1)-(2.2). The bootstrap method that we suggest can be implemented in the following steps:

1. from the observed data, compute $\hat{\pi}$ and $\hat{\beta}$, along with all other things necessary to obtain the realizations of the statistics T_l, DW_j, and the residuals $\hat{v}_1 = y_1 - Z\hat{\pi}\hat{\beta}$, $\hat{v}_2 = y_2 - Z\hat{\pi}$. These residuals are then re-centered by subtracting sample means to yield $(\tilde{v}_1, \tilde{v}_2)$;

2. for each bootstrap sample $r = 1, \ldots, B$, the data are generated following

$$y_1^* = Z^* \hat{\pi}\hat{\beta} + v_1^*, \quad y_2^* = Z^* \hat{\pi} + v_2^*, \tag{4.1}$$

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

3. The bootstrap test rejects H_0 if $\frac{1}{B}\sum_{r=1}^{B} \mathbb{1}(T_l^{*^{(r)}} > c^*_{T_l,\alpha})$ or $\frac{1}{B}\sum_{r=1}^{B} \mathbb{1}(DW_j^{*^{(r)}} > c^*_{DW_j,\alpha})$ is less then α, where $c^*_{T_l,\alpha}$ and $c^*_{DW_j,\alpha}$ are the $1-\alpha$ quantiles of T_l^* $(l = 2,3,4)$ and DW_j^* $(j = 1,2,3)$, respectively.

The bootstrap procedure described above is easy to implement because only the OLS and IV regressions are required. As seen, the re-sampling mechanism exploits the joint empirical distribution of $(Z, \tilde{v}_1, \tilde{v}_2)$. So, the unknown correlation structure between Z and u is perserved, and no distributional assumption on the errors is required. Moreover, the i.i.d. assumption in the re-sampling method can be replaced by a block bootstrap method if the errors or IVs are serially correlated; see Li (2006).

In the reminder of the paper, $\mathbb{P}^*$ is the probability under the empirical distribution function (conditional on $\mathscr{X}_n$), and $\mathbb{E}^*$ its corresponding expectation operator. Following Andrews (2002), we define the $1-\alpha$ quantile of T_l^* $(l = 2,3,4)$ and DW_j^* $(j = 1,2,3)$, namely

$c^*_{T_l,\alpha}$ and $c^*_{DW_j,\alpha}$ respectively, to be the value that minimizes $\mid \mathbb{P}^*(T_l^* \leq x) - (1-\alpha) \mid$ and $\mid \mathbb{P}^*(DW_j^* \leq x) - (1-\alpha) \mid$ over $x \in \mathbb{R}$. We will investigate in turn the size of DWH tests when the bootstrap critical values are used in the inference. But before proceeding, it will be useful to prove the following lemma on the existence of second moments of $|Z^*_{pt}v^*_{mt}|$ $(m = 1,2)$, where $[Z^* : v_1^* : v_2^*]$ is the bootstrap draw from the joint empirical distribution of $[Z : \tilde{v}_1 : \tilde{v}_2]$. For this, we shall define $v = [v_1 : v_2]$ where $v_1 = u + v_2\beta = e + Zb + v_2\beta$.

Lemma 4.1 *Suppose that Assumptions A-C and H_0 are satisfied. If further for some $\delta > 0$, we have $\mathbb{E}(\|Z_{tp}Z_t\|^{2+\delta}, \|Z_t v_t'\|^{2+\delta}, \|v_t\|^{4+\delta}) < +\infty$, then $\mathbb{E}^*(|Z^*_{pt}v^*_{mt}|^{2+\delta})$ and $\mathbb{E}^*(|v^*_{1t}v^*_{2t}|^{2+\delta})$ are bounded a.s. for all $p = 1, \ldots, k$ and $m = 1, 2$.*

Lemma **4.1** states $\mathbb{E}^*(|Z^*_{pt}v^*_{mt}|^{2+\delta})$, $t = 1, \ldots, n$, and $\mathbb{E}^*(|v^*_{1t}v^*_{2t}|^{2+\delta})$ are finite $a.s.$ for all $p = 1, \ldots, k$ and $m = 1, 2$, despite instrument endogeneity. As a result, $\mathbb{E}^*(|Z^*_{pt}u^*_t|^{2+\delta})$ and $\mathbb{E}^*(|u^*_t v^*_{2t}|^{2+\delta})$ are also bounded $a.s.$ for all $p = 1, \ldots, k$ and $m = 1, 2$, where $u^* = v_1^* - v_2^*\beta$. To check this, let $\beta \in \mathbb{R}$ be any arbitrary scalar. By using Minkowski's inequality, we have:

$$\begin{aligned} \mathbb{E}^*[|Z^*_{pt}u^*_t|^{2+\delta}] &= \mathbb{E}^*[|Z^*_{pt}v^*_{1t} - Z^*_{pt}v^*_{2t}\beta|^{2+\delta}] \\ &\leq K\left\{\mathbb{E}^*[|Z^*_{pt}v^*_{1t}|^{2+\delta}] + |-\beta|^{2+\delta}\mathbb{E}^*[|Z^*_{pt}v^*_{2t}|^{2+\delta}]\right\} \\ &\quad \text{for a large enough constant } K \\ &< +\infty\, a.s. \quad \text{by Lemma } \mathbf{4.1}. \end{aligned}$$

By similar way, we also have $\mathbb{E}^*(|u^*_t v^*_{2t}|^{2+\delta}) < +\infty\, a.s.$

In the remainder of this section, we focus on *local-to-zero* instrument endogeneity, that is, the setup where $b = b_0/\sqrt{n}$, $b_0 \in \mathbb{R}^k$ is a constant vector. The main reason of our focus on local-to-zero invalid instruments is that, when $b \neq 0$ is fixed, the asymptotic distributions of DWH statistics diverge [see Lemma **3.1**-(a) and Lemma **3.2**-(b)]. Although the bootstrap DWH statistics also diverge in this case, we are not able to establish the validity of the bootstrap, and we wish to complete our investigation with a Monte Carlo experiment.

Lemma **4.2** establishes the joint asymptotic distribution of $\frac{1}{\sqrt{n}}[Z^{*'}u^* : Z^{*'}v_2^* : v_2^{*'}u^*]$ that is extensively used to derive the distributions of the bootstrap statistics in Lemma **4.3**.

Lemma 4.2 *Suppose that Assumptions A-C and H_0 are satisfied, and let $b = b_0/\sqrt{n}$, where $b_0 \neq 0$ is fixed. If for some $\delta > 0$, $\mathbb{E}(\|Z_{tp}Z_t\|^{2+\delta}, \|Z_t v_t'\|^{2+\delta}\, \|v_t\|^{4+\delta}) < +\infty$, then:*

$$n^{-1/2}vec[Z^{*'}u^* : Z^{*'}v_2^* : v_2^{*'}u^*] \mid \mathscr{X}_n \overset{d}{\to} N\left[\begin{pmatrix} Q_Z b_0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_e^2 Q_Z & 0' & 0 \\ 0 & \sigma_{v_2}^2 Q_Z & 0 \\ 0 & 0' & \sigma_e^2\sigma_{v_2}^2 \end{pmatrix}\right] a.s.,$$

where $\mathscr{X}_n = \{(Y_1', Z_1'), \ldots, (Y_n', Z_n')\}$.

Lemma **4.2** shows that if instrument endogeneity is local-to-zero, the bootstrap distribution of $\frac{1}{\sqrt{n}}[Z^{*'}u^* : Z^{*'}v_2^* : v_2^{*'}u^*]$, given the observed sample $\mathscr{X}_n$, provides a first order approximation of the unknown distribution of $\frac{1}{\sqrt{n}}[Z'u : Z'v_2 : v_2'u]$ under H_0. So, based on this, we can prove the following lemma on the asymptotic distributions of the bootstrap DWH statistics.

Lemma 4.3 *Suppose that Assumptions A-C and H_0 are satisfied, and let $b = b_0/\sqrt{n}$, where $b_0 \neq 0$ is fixed. If for some $\delta > 0$, $\mathbb{E}(\|Z_{tp}Z_t\|^{2+\delta}, \|Z_t v_t'\|^{2+\delta}, \|v_t\|^{4+\delta}) < +\infty$, then:*

(a) $T_l^* \mid \mathscr{X}_n \xrightarrow{d} \chi^2(1;\|\bar{\tau}_{b_0}\|^2),\, DW_j^* \mid \mathscr{X}_n \xrightarrow{d} \chi^2(1;\|\bar{\tau}_{b_0}\|^2)$ *a.s. if $\pi \neq 0$ is fixed ;*

(b) $T_2^*, T_4^*, DW_3^* \mid \mathscr{X}_n, \psi_{Zv_2} \xrightarrow{d} \chi^2(1;\|\bar{v}_{b_0}\|^2),\, T_3^*, DW_1^*, DW_2^* \mid \mathscr{X}_n, \psi_{Zv_2} \xrightarrow{d} \frac{\chi^2(1;\|\bar{v}_{b_0}\|^2)}{1+\chi^2(1;\|\bar{v}_{b_0}\|^2)}$ *a.s. if $\pi = \pi_0/\sqrt{n}$, π_0 is fixed (possibly zero),*

*where $\bar{\tau}_{b_0}$ and $\bar{v}_{b_0}$ are given in Lemma **3.1** and Lemma **3.2**, respectively.*

Lemma **4.3** shows that if instrument endogeneity is local-to-zero, the bootstrap distributions of T_l^* and DW_j^*, given the observed sample $\mathscr{X}_n$, converge to the asymptotic distributions of the T_l and DW_j statistics in Lemma **3.1**-(b) and Lemma **3.2**-(b). This establishes the following corollary on the consistency of the bootstrap for exogeneity tests under the local-to-zero instrument endogeneity assumption.

Corollary 4.4 *uppose that Assumptions A-C and H_0 are satisfied, and let $b = b_0/\sqrt{n}$, where $b_0 \neq 0$ is fixed. If for some $\delta > 0$, $\mathbb{E}(\|Z_{tp}Z_t\|^{2+\delta}, \|Z_t v_t'\|^{2+\delta}, \|v_t\|^{4+\delta}) < +\infty$, then:*
$\sup_{x\in\mathbb{R}} \mid \mathbb{P}^*(T_l^* \leq x) - \mathbb{P}(T_l \leq x) \mid \xrightarrow{p} 0$ *and* $\sup_{x\in\mathbb{R}} \mid \mathbb{P}^*(DW_j^* \leq x) - \mathbb{P}(DW_j \leq x) \mid \xrightarrow{p} 0$
in probability $\mathbb{P}$ for all values of π.

Corollary **4.4** states that conditional on the observed data, the distributions of the bootstrap statistics T_l^* and DW_j^* provide a first order approximation of the unknown distributions of the T_l and DW_j statistics under exogeneity and local-to-zero invalid instruments. Hence, the bootstrap is consistent under exogeneity and small instruments endogeneity. We can now state the following theorem on the size of the DWH tests when the bootstrap critical values are used in the inference.

Theorem 4.5 *Under the conditions of Corollary **4.4**, we have*
$\mid \mathbb{P}(T_l > c^*_{T_l,\alpha}) - \alpha \mid \to 0$ *and* $\mid \mathbb{P}(DW_j > c^*_{DW_j,\alpha}) - \alpha \mid \to 0$ *as $n \to +\infty$,*
for all values of π.

Theorem **4.5** shows that if instrument endogeneity is local-to-zero, the bootstrap critical values for T_l^* and DW_j^* yield levels for the T_l and DW_j tests that are correct asymptotically under H_0, no matter how weak the IVs are. So, the bootstrap makes a substantial improvement compared with the standard DWH procedures. We will complete our analysis with a Monte Carlo experiment.

5. Monte Carlo experiment

We use simulation to examine the size performance of the standard and proposed bootstrap DWH tests. The DGP is described by

$$y_1 = y_2\beta + u,\ y_2 = Z\pi + v_2, \tag{5.1}$$
$$u = Zb + e, \tag{5.2}$$

where y_1 and y_2 are $n \times 1$ vectors, Z contains $k \in \{5, 20\}$ instruments that violate the exclusion restrictions when $b \neq 0$. The errors $(e_t, v_{2t})'$ and IVs Z_t, $t = 1, \ldots, n$, are generated as

$$(u_t, v_{2t})' = J\varepsilon_t \quad \text{with} \quad J = \begin{pmatrix} 1 & \rho_{v_2 e} \\ 0 & 1 \end{pmatrix}$$
$$\text{and } (Z_t', \varepsilon_t')' \sim N(0, I_{k+2}), \tag{5.3}$$

where $\rho_{v_2 e}$ measures the correlation between e_t and v_{2t}. We set $b = \lambda.\mathbb{1}_k$ where λ is the measure of Z endogeneity, and $\mathbb{1}_k$ is a $k \times 1$ vector of ones. In this experiment, we vary λ in $\{0, 0.05, 0.1, 0.2, 0.3\}$. It is then a simple exercise to see that λ measures the covariance between u_t and Z_{pt} for all $p = 1, \ldots, k$. So, $\lambda = 0$ (or $b = 0$) means exogenous instruments, while $\lambda \neq 0$ characterizes invalid instruments. In particular, $\lambda = 0.05$ can be seen as local-to-zero instrument endogeneity, while $\lambda = 0.1$ corresponds to moderate instrument endogeneity. Meanwhile, each value of λ in $\{0.2, 0.3\}$ is interpreted as strong instrument endogeneity in this exercise. We set $\beta = 2$ and $\pi = (\frac{\mu^2}{n\|Z.\mathbb{1}_k\|})^{1/2}.\mathbb{1}_k$, where μ^2 is the concentration parameter that usually determines the quality of the instruments; see Hansen, Hausman and Newey (2008), Guggenberger (2010), and Doko Tchatoka and Dufour (2011a). The values $\mu^2 \leq 613$ correspond to weak instruments and those $\mu^2 > 613$ characterize strong instruments.[6] In this experiment, we choose μ^2 in $\{0, 13, 1000\}$ but the results do not change qualitatively for alternative choices of μ^2.

Table 1 presents the empirical rejection frequencies of the DWH tests with the usual asymptotic χ^2 and bootstrap critical values. In the first part of the table, the rejection

[6] see Hansen et al. (2008) and Doko Tchatoka and Dufour (2011a).

frequencies with the asymptotic χ^2 critical values are reported, while the second part of the table contains those with the bootstrap critical values. The first column of the table shows the test statistics, the second reports the number of possibly invalid instruments, and the others present, for each value of instrument endogeneity λ and IV strength μ^2, the empirical rejections of the tests. We use $5,000$ replications and the nominal level and sample size are $\alpha = 5\%$ and $n = 300$ respectively, in all cases. The bootstrap critical values are computed using $B = 999$ bootstrap pseudo-samples of size $n = 300$. As it can be seen from the table, two main findings emerge.

First, when the usual asymptotic χ^2 critical values are used in the inference, the rejection frequencies of all DWH tests exceed the nominal 5% level with a wide margin, irrespective of μ^2 (instrument quality), thus supporting our theoretical findings of Theorems **3.3** - **3.4**. The maximal rejection frequency of the null hypothesis of exogeneity can even be as great as 100% for all DWH tests, including when instrument endogeneity is not very high (for example, see column $\mu^2 = 1000$ and $\lambda = 0.1$). More importantly, all tests show serious size distortions even for small values of instrument endogeneity; for example, see column $\lambda = 0.05$ in the table. Moreover, we observe that the rejection frequencies increase with instrument endogeneity, thus confirming our analysis in Theorem **3.4**. In addition, instrument endogeneity is more severe when the quality of the IVs is high than when it is low. This is quite intuitive because weak instruments has an opposite effect on the size of DWH tests than invalid instruments.

Second, the second part of Table 1 clearly shows that bootstrapping substantially improves the DWH tests size, including when instrument endogeneity is relatively large. As seen, the rejection frequencies with the bootstrap critical values are remarkably low compared to those with asymptotic χ^2 critical values. For example, when $\mu^2 = 1000$ and $\lambda = 0.1$, the maximal rejection frequencies of the tests are around 99% (with the asymptotic χ^2 critical values) against only 5.8% for T_2, T_4 and DW_3, and only about 9.8% for T_3, DW_1 and DW_2, when the bootstrap critical values are used. This represents significant differences and underscores the importance of bootstrapping. Even when $\lambda = 0.3$, the maximal rejection frequencies of all bootstrap tests is only around 15%. More importantly, these results hold irrespective of the value of μ^2. Further, when λ varies in $\{0, 0.05, 0.1\}$ (small and moderate instrument endogeneity), the bootstrap rejection frequencies are close to the 5% nominal level, thus confirming the theoretical results of Theorem **4.5**.

Table 1. Empirical size (in %) of the DWH tests at nominal level 5%, $n = 100$

Asymptotic χ^2 critical values

Statistics	$k \downarrow$ $\mu^2 \rightarrow$	$\lambda = 0$			$\lambda = .05$			$\lambda = .1$			$\lambda = .2$			$\lambda = .3$		
		0	13	1000	0	13	1000	0	13	1000	0	13	1000	0	13	1000
T_2	5	4.8	5.2	4.2	13.5	20.8	62.0	34.6	53.8	99.6	59.9	79.6	100.0	69.3	85.0	100.0
T_3	-	0.3	0.7	4.0	1.2	5.0	60.8	6.3	20.5	99.6	18.6	47.8	100.0	25.4	53.8	100.0
T_4	-	4.7	5.2	4.2	13.5	20.8	61.9	34.6	53.7	99.6	59.9	79.6	100.0	69.3	85.0	100.0
DW_1	-	0.3	0.7	3.8	1.2	4.7	60.5	6.2	20.1	99.6	18.3	47.4	100.0	25.2	53.4	100.0
DW_2	-	0.3	0.7	4.0	1.3	5.0	60.9	6.4	20.7	99.6	18.7	47.9	100.0	25.5	54.0	100.0
DW_3	-	4.8	5.2	4.3	13.5	20.8	62.1	34.7	53.8	99.6	60.0	79.7	100.0	69.4	85.0	100.0
T_2	20	5.1	5.1	5.2	12.3	32.1	99.7	28.4	67.4	100.0	46.5	84.0	100.0	54.4	88.6	100.0
T_3	-	2.9	3.2	5.0	8.3	26.2	99.7	22.7	62.9	100.0	40.9	81.9	100.0	49.5	87.2	100.0
T_4	-	5.1	5.1	5.1	12.2	32.1	99.7	28.3	67.3	100.0	46.4	84.0	100.0	54.4	88.6	100.0
DW_1	-	2.8	3.0	4.9	8.2	25.8	99.7	22.5	62.5	100.0	40.6	81.7	100.0	49.2	87.0	100.0
DW_2	-	2.9	3.2	5.1	8.3	26.3	99.7	22.7	62.9	100.0	41.0	82.0	100.0	49.6	87.2	100.0
DW_3	-	5.1	5.2	5.2	12.3	32.1	99.7	28.4	67.4	100.0	46.5	84.0	100.0	54.5	88.6	100.0

Bootstrap critical values

Statistics	$k \downarrow$ $\mu^2 \rightarrow$	$\lambda = 0$			$\lambda = .05$			$\lambda = .1$			$\lambda = .2$			$\lambda = .3$		
		0	13	1000	0	13	1000	0	13	1000	0	13	1000	0	13	1000
T_2	5	4.7	5.9	4.1	5.5	5.3	5.7	4.9	4.6	5.7	4.7	6.3	6.1	6.5	6.0	6.7
T_3	-	7.7	8.7	8.4	10.3	12.0	8.0	10.7	10.0	9.7	7.6	10.4	12.2	11.8	12.5	9.5
T_4	-	4.7	5.9	4.1	5.5	5.3	5.7	4.9	4.6	5.7	4.7	6.3	6.1	6.5	6.0	6.7
DW_1	-	7.8	8.8	8.5	10.4	11.1	8.1	10.8	10.1	9.8	7.8	10.5	12.4	11.9	12.7	9.7
DW_2	-	7.7	8.7	8.4	10.3	11.0	8.0	10.7	10.0	9.7	7.6	10.4	12.2	11.8	12.5	9.5
DW_3	-	4.7	5.9	4.1	5.5	5.3	5.7	4.9	4.6	5.7	4.7	6.3	6.1	6.5	6.0	6.7
T_2	20	5.1	5.0	6.2	4.8	5.4	4.2	4.0	4.9	5.8	3.8	5.4	8.6	6.0	5.2	14.1
T_3	-	6.1	6.0	7.0	5.8	6.2	5.2	5.1	5.9	6.7	5.1	6.4	9.3	7.8	6.0	15.7
T_4	-	5.1	5.0	6.2	4.8	5.4	4.2	4.0	4.9	5.8	3.8	5.4	8.6	6.0	5.2	14.1
DW_1	-	6.3	6.1	7.1	6.1	6.3	5.3	5.2	6.1	6.9	5.3	6.7	9.6	7.9	6.1	15.9
DW_2	-	6.1	6.0	7.0	5.8	6.2	5.2	5.1	5.9	6.7	5.1	6.4	9.3	7.8	6.0	15.7
DW_3	-	5.1	5.0	6.2	4.8	5.4	4.2	4.0	4.9	5.8	3.8	5.4	8.6	6.0	5.2	14.1

6. Conclusion

In this paper, we focused on linear IV regressions and investigated the sensitivity to instrument endogeneity of Durbin-Wu-Hausman (DWH) tests of exogeneity. The DWH tests are widely used as pretests in applied work to decide whether OLS or IV method is applicable, and the research on this topic is widespread.[7] However, most of these studies usually assume that IVs are strong and strictly exogenous. There is now a growing literature[8] about the reliability of the DWH procedures in the presence of weak instruments. But little is known about their behavior when instruments violate when instrument violate the exclusion restrictions.

In this paper, we show that: (1) all DWH tests considered have size converging to 1 when instrument endogeneity is fixed, no matter how weak the instruments are; and (2) all tests have size greater than their nominal level asymptotically, when instrument endogeneity is local-to-zero. We propose size correction based on bootstrap techniques. Our analysis of the proposed bootstrap tests provides some new insights. More precisely, we show that under *local-to-zero instrument endogeneity*, the bootstrap provides a first-order approximation of the asymptotic size of DWH statistics, no matter how weak the instruments are. We present a Monte Carlo experiment that confirms our theoretical findings.

[7] See Dufour (1979, 1987), Hwang (1980, 1985), Kariya and Hodoshima (1980), Hausman and Taylor (1981), Spencer and Berk (1981), Nakamura and Nakamura (1981, 1985), Engle (1982), Holly (1982, 1983b, 1983a), Holly and Monfort (1983), Reynolds (1982), Smith (1983, 1984, 1985), Thurman (1986), Smith and Pesaran (1990), Ruud (1984, 2000), Newey (1985a, 1985b), Meepagala (1992), and Baum et al. (2003), among others.

[8] For examples, see Staiger and Stock (1997), Wong (1996, 1997), Guggenberger (2010), Hahn et al. (2010), Doko Tchatoka and Dufour (2011a, 2011b, 2014), Kiviet and Niemczyk (2007, 2012), Kiviet and Pleus (2012), and Doko Tchatoka (2013).

APPENDIX

A. Proofs

PROOF OF LEMMA **3.1** To shorten the exposition, we only show the proof for the statistic DW_3. Those of the other statistics can be deduced in similar way.

To start, recall from (2.4) that DW_3 is given by

$$DW_3 = n(\tilde{\beta} - \hat{\beta})^2 / \hat{\omega}_3^2, \tag{A.1}$$

where $\tilde{\beta} - \hat{\beta} = \hat{\omega}_{IV}^{-1} y_2' P_Z u/n - \hat{\omega}_{LS}^{-1} y_2' u/n$ and $\hat{\omega}_3^2 = \hat{\sigma}^2 \Delta$. We will consider the following two cases: (a) $b \neq 0$ is fixed, and (b) $b = b_0/\sqrt{n}$, where b_0 is fixed (possibly zero).

(a) Suppose first that $b \neq 0$ is fixed (that is, does not dependent on the sample size). Under Assumptions A - C, and if further $\pi \neq 0$ and H_0 is satisfied, we have $\hat{\omega}_{LS} \xrightarrow{p} \pi' Q_Z \pi + \sigma_{v_2}^2 > 0$, $\hat{\omega}_{IV} \xrightarrow{p} \pi' Q_Z \pi > 0$, $y_2' P_Z u/n \xrightarrow{p} \pi' Q_Z b \neq 0$, $y_2' u/n \xrightarrow{p} 0$, $\hat{\sigma}^2 = u'u/n - (u' y_2/n) \hat{\omega}_{ls}^{-1} (y_2' u/n) \xrightarrow{p} \sigma_u^2 = \sigma_e^2 + b' Q_Z b > 0$, so that we get: $\tilde{\beta} - \hat{\beta} \xrightarrow{p} (\pi' Q_Z \pi)^{-1} \pi' Q_Z b \neq 0$ and $\hat{\omega}_3^2 \xrightarrow{p} \sigma_u^2 [(\pi' Q_Z \pi)^{-1} - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1}] > 0$. As a result, we have $(\tilde{\beta} - \hat{\beta})^2 / \hat{\omega}_3^2 \xrightarrow{p} \frac{1}{\sigma_u^2} (\pi' Q_Z \pi)^{-2} [(\pi' Q_Z \pi)^{-1} - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1}]^{-1} (\pi' Q_Z b)^2 > 0$. From (A.1), it is clear that $DW_3 = n(\tilde{\beta} - \hat{\beta})^2 / \hat{\omega}_3^2 \xrightarrow{p} +\infty$, as $n \to +\infty$.

(b) Suppose now that $b = b_0/\sqrt{n}$, where b_0 is fixed (possibly zero). It is easy to see that $\hat{\sigma}^2 \xrightarrow{p} \sigma_e^2 > 0$, so that we have: $\hat{\omega}_3^2 \xrightarrow{p} \omega_0^2 = \sigma_e^2 [(\pi' Q_Z \pi)^{-1} - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1}] > 0$. Now, we can write $\sqrt{n}(\tilde{\beta} - \hat{\beta})$ as:

$$\sqrt{n}(\tilde{\beta} - \hat{\beta}) = \hat{\omega}_{IV}^{-1} \frac{y_2' P_Z u}{\sqrt{n}} - \hat{\omega}_{LS}^{-1} \frac{y_2' u}{\sqrt{n}}. \tag{A.2}$$

Under the assumptions of the lemma, we have $y_2' P_Z u/\sqrt{n} = \pi' \frac{Z'Z}{n} b_0 + \pi' Z' e/\sqrt{n} + v_2' P_Z u/\sqrt{n} \xrightarrow{d} \pi' Q_Z b_0 + \pi' \psi_{Ze}$ and $y_2' u/\sqrt{n} = \pi' \frac{Z'Z}{n} b_0 + \pi' Z' e/\sqrt{n} + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (v_{2t} e_t - \sigma_{v_2 e}) + o_p(1) \xrightarrow{d} \pi' Q_Z b_0 + \pi' \psi_{Ze} + \psi_{v_2 e}$. So, we get $\sqrt{n}(\tilde{\beta} - \hat{\beta}) \xrightarrow{d} \Psi_{b_0} = (\pi' Q_Z \pi)^{-1} \pi' (Q_Z b_0 + \psi_{Ze}) - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1} (\pi' Q_Z b_0 + \pi' \psi_{Ze} + \psi_{v_2 e}) = [(\pi' Q_Z \pi)^{-1} - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1}] \pi' Q_Z b_0 + (\pi' Q_Z \pi)^{-1} \pi' \psi_{Ze} - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1} (\pi' \psi_{Ze} + \psi_{v_2 e})$ from (A.2). It follows from (A.1) that

$$DW_3 \xrightarrow{d} (\omega_0^{-1} \Psi_{b_0})^2, \text{ where } \omega_0 = \sigma_e [(\pi' Q_Z \pi)^{-1} - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1}]^{1/2}. \tag{A.3}$$

Again, under the assumptions of the lemma, we have $(\psi_{Ze}', \psi_{v_2 e})' \sim N\left[o,\ \sigma_e^2 \begin{pmatrix} Q_Z & 0 \\ 0 & \sigma_{v_2}^2 \end{pmatrix}\right]$ jointly, so that we get (with a little calculation): $\omega_0^{-1} \Psi_{b_0} \sim N(\overline{\tau}_{b_0}, 1)$, where $\overline{\tau}_{b_0} = \frac{1}{\sigma_e} [(\pi' Q_Z \pi)^{-1} - (\pi' Q_Z \pi + \sigma_{v_2}^2)^{-1}]^{1/2} \pi' Q_Z b_0$. As a result, we have $DW_3 \xrightarrow{d} \chi^2(1; \|\overline{\tau}_{b_0}\|^2)$. This completes the proof.

□

PROOF OF LEMMA **3.2** The proof is similar to those of Lemma **3.1**. Again, we will only focus on the proof for DW_3. Now, let us express DW_3 as

$$DW_3 = \frac{(\tilde{\beta}-\hat{\beta})^2}{n^{-1}\hat{\omega}_3^2}, \text{ where } n^{-1}\hat{\omega}_3^2 = \hat{\sigma}^2[(n\hat{\omega}_{IV})^{-1} - \frac{1}{n}\hat{\omega}_{LS}^{-1}]. \tag{A.4}$$

We will distinguish the following two cases (similar to the proof of Lemma **3.1**): (a) $b \neq 0$ is fixed, and (b) $b = b_0/\sqrt{n}$, where b_0 is fixed (possibly zero).

(a) Suppose first that $b \neq 0$ is fixed. Under Assumptions A - C, and H_0, and if further $\pi = \pi_0/\sqrt{n}$, where π_0 is fixed, then: $\hat{\omega}_{LS} \xrightarrow{p} \sigma_{v_2}^2 > 0$, $n\hat{\omega}_{IV} \xrightarrow{d} \tilde{\omega}_0 = (\pi_0 + Q_Z^{-1}\psi_{Zv_2})'Q_Z(\pi_0 + Q_Z^{-1}\psi_{Zv_2}) > 0$ with probability one, and $\hat{\sigma}^2 = u'u/n - (u'y_2/n)\hat{\omega}_{LS}^{-1}(y_2'u/n) \xrightarrow{p} \sigma_u^2 = \sigma_e^2 + b'Q_Z b > 0$. So, we have $n^{-1}\hat{\omega}_3^2 = \hat{\sigma}^2(n\hat{\omega}_{IV})^{-1} + o_p(1) \xrightarrow{d} \sigma_u^2\tilde{\omega}_0^{-1}$ and $\tilde{\beta} - \hat{\beta} = (n\hat{\omega}_{IV})^{-1}y_2'P_Z u - \hat{\omega}_{LS}^{-1}y_2'u/n = (n\hat{\omega}_{IV})^{-1}y_2'P_Z u + o_p(1)$. So, it is clear that

$$DW_3 = \frac{1}{\sigma_u^2}(n\hat{\omega}_{IV})^{-1}(y_2'P_Z u)^2 + o_p(1), \text{ where } \sigma_u^2 = \sigma_e^2 + b'Q_Z b > 0. \tag{A.5}$$

Since under the assumptions of the lemma, we have $y_2'P_Z u = (y_2'Z/\sqrt{n})(Z'Z/n)^{-1}(Z'u/\sqrt{n}) = n(v_2'Z/\sqrt{n})(Z'Z/n)^{-1}(Z'u/n)$, hence we have $DW_3 = \frac{n^2}{\sigma_u^2}(n\hat{\omega}_{IV})^{-1}[(y_2'Z/\sqrt{n})(Z'Z/n)^{-1}(Z'u/n)]^2$. Again, we have $Z'u/n \xrightarrow{p} Q_Z b \neq 0$ when $b \neq 0$ and $(y_2'Z/\sqrt{n})(Z'Z/n)^{-1} \xrightarrow{d} (\pi_0 + Q_Z^{-1}\psi_{Zv_2})' \neq 0$ with probability one. Since we also have $\frac{1}{\sigma_u^2}(n\hat{\omega}_{IV})^{-1} \xrightarrow{d} (\sigma_u^2\tilde{\omega}_0)^{-1} > 0$ with probability one, it follows immediately that $DW_3 \xrightarrow{p} +\infty$ as $n \to +\infty$.

(b) Suppose now that $b = b_0/\sqrt{n}$, where b_0 is fixed (possibly zero). Under Assumptions A - C, H_0 and if further $\pi = \pi_0/\sqrt{n}$, we have $y_2'P_Z u = (y_2'Z/\sqrt{n})(Z'Z/n)^{-1}(Z'u/\sqrt{n}) \xrightarrow{d} (\pi_0 + Q_Z^{-1}\psi_{Zv_2})'(\psi_{Ze} + Q_Z b_0)$, $\hat{\sigma}^2 \xrightarrow{p} \sigma_e^2$, so that we get: $DW_3 \xrightarrow{d} (\sigma_e^2\tilde{\omega}_0)^{-1}[(\pi_0 + Q_Z^{-1}\psi_{Zv_2})'(\psi_{Ze} + Q_Z b_0)]^2$ from (A.5). As in the proof of Lemma **3.1**-(b), it is easy to see that

$$(\sigma_e^2\tilde{\omega}_0)^{-1/2}(\pi_0 + Q_Z^{-1}\psi_{Zv_2})'(\psi_{Ze} + Q_Z b_0) \mid \psi_{Zv_2} \sim N(\bar{v}_{b_0}, 1), \tag{A.6}$$

where $\bar{v}_{b_0} = (\sigma_e^2\tilde{\omega}_0)^{-1/2}(\pi_0 + Q_Z^{-1}\psi_{Zv_2})'Q_Z b_0$. Thus the result follows.

□

PROOF OF THEOREM **3.3** Theorem **3.3** follows directly from Lemma **3.1**-(a) and Lemma **3.2**-(a), there the proof is omitted.

□

PROOF OF THEOREM **3.4** Again, we prove the results for DW_3. First, suppose $b = b_0/\sqrt{n}$, where $b_0 \neq 0$ is fixed. It will be illuminating to deal separately with the following two subcases : (a) $\pi \neq 0$ is fixed, and (b) $\pi = \pi_0/\sqrt{n}$, where $\pi_0 \in \mathbb{R}^k$ is fixed ($\pi_0 = 0$ is allowed).

(a) Suppose first that $\pi \neq 0$ is fixed, and that Assumptions A - C and H_0 are satisfied. From

Lemma **3.1**-(b), we have $DW_3 \xrightarrow{d} \chi^2(1; \|\overline{\tau}_{b_0}\|^2)$ so that as $n \to +\infty$, we have

$$\mathbb{P}(DW_3 > \chi^2_{1,\alpha}) \to \mathbb{P}(\chi^2(1; \|\overline{\tau}_{b_0}\|^2) > \chi^2_{1,\alpha}) = 1 - G_1(\chi^2_{1,\alpha};\ \|\overline{\tau}_{b_0}\|^2) \equiv p_\alpha > \alpha \tag{A.7}$$

because $\|\overline{\tau}_{b_0}\|^2 > 0$ when $b_0 \neq 0$. Note that $p_\alpha = \alpha$ in (A.7) only when $b_0 = 0$, which is the case of exogenous instruments and is excluded by assumption.

(b) Suppose now that $\pi = \pi_0/\sqrt{n}$, where $\pi_0 \in \mathbb{R}^k$ is fixed ($\pi_0 = 0$ is allowed). From Lemma **3.2**-(b), we have $DW_3 \mid \psi_{Zv_2} \xrightarrow{d} \chi^2(1; \|\overline{v}_{b_0}\|^2)$. Therefore, as $n \to +\infty$, we have

$$\begin{aligned} \mathbb{P}[DW_3 > \chi^2_{1,\alpha} \mid \psi_{Zv_2}] &\to \mathbb{P}[\chi^2(1; \|\overline{\tau}_{b_0}\|^2) > \chi^2_{1,\alpha} \mid \psi_{Zv_2}] \\ &= 1 - G_1(\chi^2_{1,\alpha};\ \|\overline{v}_{b_0}\|^2) > \alpha \text{ with probability one.} \end{aligned} \tag{A.8}$$

Note that in (A.8), $G_1(\chi^2_{1,\alpha};\ \|\overline{v}_{b_0}\|^2)$ is a random variable because $\|\overline{v}_{b_0}\|^2$ depends on ψ_{Zv_2} which is random [see Lemma **3.2**-(b)]. By the mean value argument, it follows that

$$\mathbb{P}[DW_3 > \chi^2_{1,\alpha}] \to 1 - \mathbb{E}[G_1(\chi^2_{1,\alpha};\ \|\overline{v}_{b_0}\|^2)] \equiv \tilde{p}_\alpha > \alpha \tag{A.9}$$

as $n \to +\infty$, where the expectations in (A.9) are taken over the distribution of ψ_{Zv_2}. This completes the proof. □

PROOF OF LEMMA **4.1** First, observe that $\mathbb{E}^*(|Z^*_{pt} v^*_{mt}|^{2+\delta}) \leq |\ \mathbb{E}^*(Z^*_{pt} v^*_{mt})\ |^{2+\delta}$ by Jensen's inequality. Hence, it suffices to establish that $|\ \mathbb{E}^*(Z^*_{pt} v^*_{mi})\ |^{2+\delta}$ is bounded *a.s*. Because each row of (Z^*, v^*_1, v^*_2) are i.i.d under $\mathbb{P}^*$, we have $\mathbb{E}^*(Z^*_{pt} v^*_{mt}) = \frac{1}{n}\sum_{t=1}^n Z_{pt}\hat{v}_{mt},\ p = 1, \ldots, k$ and $m = 1, 2$. So, by applying Minkowski and Cauchy-Schwartz inequalities, we have (under H_0):

$$\begin{aligned} |\ \mathbb{E}^*(Z_{pt} v^*_{1t})\ |^{2+\delta} &= |\ \frac{1}{n}\sum_{t=1}^n Z_{pt}\hat{v}_{1t}\ |^{2+\delta} \leq \frac{1}{n}\sum_{t=1}^n |\ Z_{pt}\hat{v}_{1t}\ |^{2+\delta} \\ &= \frac{1}{n}\sum_{t=1}^n |Z_{pt}v_{1t} - Z_{pt}Z'_t(\hat{\pi}\hat{\beta} - \pi\beta)|^{2+\delta} \\ &\leq C_1[\frac{1}{n}\sum_{t=1}^n |Z_{pt}v_{1t}|^{2+\delta} + \frac{1}{n}\sum_{t=1}^n |Z_{pt}Z'_t(\hat{\pi}\hat{\beta} - \pi\beta)|^{2+\delta}] \\ &\leq C_2[\frac{1}{n}\sum_{t=1}^n |Z_{pt}v_{1t}|^{2+\delta} + |\ \hat{\pi}\hat{\beta} - \pi\beta\ |^{2+\delta}\ \frac{1}{n}\sum_{t=1}^n \|Z_{pt}Z_t\|^{2+\delta}] \end{aligned}$$

for large enough positive constants C_1 and C_2. Under Assumptions A-C and H_0, we have $\hat{\pi} - \pi = o_p(1)$ and $\hat{\beta} - \beta = O_p(1)$ for all values of π and b. So, we have $\hat{\pi}\hat{\beta} - \pi\beta = O_p(1)$. Moreover, $\frac{1}{n}\sum_{t=1}^n \|Z_{pt}Z_t\|^{2+\delta} \xrightarrow{p} \mathbb{E}(\|Z_{pt}Z_t\|^{2+\delta}) < +\infty$, and $\frac{1}{n}\sum_{t=1}^n |Z_{pt}v_{1t}|^{2+\delta} \xrightarrow{p} \mathbb{E}(|Z_{pt}v_{1t}|^{2+\delta}) < +\infty$ by assumption. It follows immediately that $|\ \mathbb{E}^*(Z_{pt}v^*_{1t})\ |^{2+\delta} < +\infty$ so that we get $\mathbb{E}^*(|Z^*_{pt}v^*_{1t}|^{2+\delta}) < +\infty$, as stated. By the same way, we can prove that $\mathbb{E}^*(|Z^*_{pt}v^*_{2t}|^{2+\delta}) < +\infty$ and $\mathbb{E}^*(|v^*_{1t}v^*_{2t}|^{2+\delta}) < +\infty$.

□

PROOF OF LEMMA **4.2** Let c_m $(m = 1,2)$ be $k \times 1$ nonzero vectors, and let also $c_3 \in \mathbb{R} \setminus \{0\}$. Define

$$X_{nt} = c_1' Z_t^* u_t^* / \sqrt{n} + c_2' Z_t^* v_{2t}^* / \sqrt{n} + c_3 u_t^* v_{2t}^* / \sqrt{n}, \text{ where } u_t^* = v_{1t}^* - v_{2t}^* \beta = Z_t^{*'} b + e_t^*.$$

We will establish that X_{nt} satisfies all the conditions of the Liapunov Central Limit Theorem so that the Cramér-Wold device applies.

1. We can write $\mathbb{E}^*(X_{nt})$ as:

$$\mathbb{E}^*(X_{nt}) = \underbrace{n^{-1/2} c_1' \mathbb{E}^*(Z_t^* v_{1t}^*)}_{(1)} + \underbrace{n^{-1/2}(c_2' - c_1'\beta)\mathbb{E}^*(Z_t^* v_{2t}^*)}_{(2)} + \underbrace{n^{-1/2} c_3 \mathbb{E}^*(u_t^* v_{2t}^*)}_{(3)}. \quad \text{(A.10)}$$

First, we have $(2) = 0$ because v_2^* and Z^* are independent and $\mathbb{E}^*(v_2^*) = 0$, so that $\mathbb{E}^*(Z_t^* v_{2t}^*) = 0$. Second, by the i.i.d. assumption under $\mathbb{P}^*$, we have $\mathbb{E}^*(Z_t^* v_{1t}^*) = \frac{1}{n}\sum_{t=1}^n Z_t \tilde{v}_{1t} = \left(\frac{Z'\tilde{v}_1}{n}\right)$ and $\mathbb{E}^*(u_t^* v_{2t}^*) = \frac{1}{n}\sum_{t=1}^n \tilde{u}_t \tilde{v}_{2t} = \frac{1}{n}\sum_{t=1}^n \tilde{v}_{1t}\tilde{v}_{2t} - \beta \frac{1}{n}\sum_{t=1}^n \tilde{v}_{2t}^2 = \left(\frac{\tilde{v}_1'\tilde{v}_2}{n}\right) - \beta\left(\frac{\tilde{v}_2'\tilde{v}_2}{n}\right)$. So, we get $\mathbb{E}^*(X_{nt}) = n^{-1/2} c_1' \left(\frac{Z'\tilde{v}_1}{n}\right) + n^{-1/2} c_3 \left(\frac{\tilde{v}_1'\tilde{v}_2}{n}\right) - n^{-1/2} c_3 \beta \left(\frac{\tilde{v}_2'\tilde{v}_2}{n}\right) < \infty$ *a.s* because under the conditions of of Lemma **4.1**, it is easy to see that it is straightforward to see that $\left(\frac{Z'\tilde{v}_1}{n}\right) = n^{-1/2} Q_Z b_0 + o_p(1)$, and $\left(\frac{\tilde{v}_1'\tilde{v}_2}{n}\right) - \beta\left(\frac{\tilde{v}_2'\tilde{v}_2}{n}\right) = \underbrace{n^{-1/2} Q_Z b_0 + \sigma_{v_2 e}}_{=0 \text{ under } H_0} + o_p(1) = o_p(1)$, so that we get: $\mathbb{E}^*(X_{nt}) = n^{-1} c_1' Q_Z b_0 + o_p(n^{-\frac{1}{2}}) < \infty$. This establishes the first condition of the Liapunov Central Limit Theorem.

2. We will now prove that $\mathbb{E}^*(X_{nt}^2) < +\infty$. Again, by the i.i.d. assumption on each row of (Z^*, v_1^*, v_2^*) and because under H_0, the covariance between u^* and v_2^* is $O(n^{-1/2})$ when $b = b_0/\sqrt{n}$, we get (with a simple calculation):

$$\begin{aligned} \mathbb{E}^*(X_{nt}^2) &= n^{-1}\left\{ c_1' \left(\frac{Z'\tilde{u}\tilde{u}'Z}{n}\right) c_1 + c_2' \left(\frac{Z'\tilde{v}_2\tilde{v}_2'Z}{n}\right) c_2 + c_3^2 \left(\frac{\tilde{u}'\tilde{v}_2\tilde{v}_2'\tilde{u}}{n}\right) + O_p(n^{-1/2}) \right\} < \infty \quad a.s. \\ &= n^{-1}[\sigma_e^2 c_1' Q_Z c_1 + \sigma_{v_2}^2 c_2' Q_Z c_2 + c_3^2 \sigma_e^2 \sigma_{v_2}^2] + O_p(n^{-3/2}) < \infty \; a.s. \text{ under Lemma } \mathbf{4.1}. \end{aligned}$$

3. Finally, we need to show that $\lim_{n\to\infty}\sum_{t=1}^n \mathbb{E}^*(|X_{nt}|^{2+\delta}) = 0$ a.s. for some $\delta > 0$. Indeed, we have:

$$\begin{aligned} &\sum_{t=1}^n \mathbb{E}^*|X_{nt}|^{2+\delta} = n^{-\delta/2} n^{-1} \sum_{t=1}^n \mathbb{E}^*\left[|c_1' Z_t^* u_t^* + c_2' Z_t^* v_{2t}^* + c_3 u_t^* v_{2t}^*|^{2+\delta}\right] \\ &\le K_1 n^{-\delta/2} \mathbb{E}^*\left[|c_1' Z_t^* u_t^*|^{2+\delta} + |c_2' Z_t^* v_{2t}^*|^{2+\delta} + |c_3 u_t^* v_{2t}^*|^{2+\delta}\right] \\ &\le K_2 n^{-\delta/2}\left\{\sum_{p=1}^k |c_{1p}|^{2+\delta}\mathbb{E}^*[|Z_{pt}^* u_t^*|^{2+\delta}] + \sum_{p=1}^k |c_{2p}|^{2+\delta}\mathbb{E}^*[|Z_{pt}^* v_{2t}^*|^{2+\delta} + |c_3|^{2+\delta}\mathbb{E}^*[|u_t^* v_{2t}^*|^{2+\delta}]\right\} \\ &= K_2 n^{-\delta/2}[A_1 + A_2 + A_3],\; A_1 = \sum_{p=1}^k |c_{1p}|^{2+\delta}\mathbb{E}^*[|Z_{pt}^* u_t^*|^{2+\delta}],\; A_2 = \sum_{p=1}^k |c_{2p}|^{2+\delta}\mathbb{E}^*[|Z_{pt}^* v_{2t}^*|^{2+\delta}] \end{aligned}$$

and $\quad A_3 = |c_3|^{2+\delta}\mathbb{E}^*[|u_t^* v_{2t}^*|^{2+\delta}]$

for large enough positive constants K_1 and K_2. From Lemma **4.1**, all A_1, A_2 and A_3 are bounded $a.s$. Therefore,we have $\lim_{n\to\infty}\sum_{t=1}^{n}\mathbb{E}^*[|X_{nt}|^{2+\delta}] = 0$ a.s. Lemma **4.2** holds by the Central Limit Theorem property.

□

PROOF OF LEMMA **4.3** Again, we will only focus on DW_3^*. First, the bootstrap DW_3^* statistic is given by

$$DW_j^* \quad = \quad n(\tilde{\beta}^* - \hat{\beta}^*)^2/\hat{\omega}_3^{*^2}, \tag{A.11}$$

where $\tilde{\beta}^*$, $\hat{\beta}^*$, $\hat{\omega}_3^{*^2}$ are the bootstrap counterparts of $\tilde{\beta}$, $\hat{\beta}$, and $\hat{\omega}_3^2$, respectively. So, we have $\tilde{\beta}^* - \hat{\beta}^* = (y_2^{*'}P_{Z^*}y_2^*)^{-1}y_2^{*'}P_{Z^*}u^* - (y_2^{*'}y_2^*)^{-1}y_2^{*'}u^*$ and $\hat{\omega}_3^{*^2} = \hat{\sigma}^{*^2}[(y_2^{*'}P_{Z^*}y_2^*/n)^{-1} - (y_2^{*'}y_2^*/n)^{-1}]$. From the i.i.d property of the bootstrap sample under $\mathbb{P}^*$, we almost surely have: $\mathbb{E}^*(Z^{*'}Z^*/n) = Z'Z/n$, $\mathbb{E}^*(Z^{*'}u^*/n) = Z'\tilde{u}/n$, $\mathbb{E}^*(Z^{*'}v_2^*/n) = Z'\tilde{v}_2/n$, and $\mathbb{E}^*[(u^*:v_2^*)'(u^*:v_2^*)/n] = (\tilde{u}:\tilde{v}_2)'(\tilde{u}:\tilde{v}_2)/n$. So, by the Markov law of large numbers, we have (conditional on $\mathscr{X}_n$): $Z^{*'}Z^*/n - Z'Z/n \to 0$, $Z^{*'}u^*/n - Z'\tilde{u}/n \to 0$, $Z^{*'}v_2^*/n - Z'\tilde{v}_2/n \to 0$, and $(u^*:v_2^*)'(u^*:v_2^*)/n - (\tilde{u}:\tilde{v}_2)'(\tilde{u}:\tilde{v}_2)/n \to 0$, $a.s.$ Since $Z'Z/n \xrightarrow{p} Q_Z$, $Z'\tilde{v}_2/n \xrightarrow{p} 0$, $Z'\tilde{u}/n \xrightarrow{p} Q_Z b_0$, and $(\tilde{u}:\tilde{v}_2)'(\tilde{u}:\tilde{v}_2)/n \xrightarrow{p} \Sigma_V$ when $b = b_0/\sqrt{n}$ (b_0 is fixed) and Assumptions A-C, and H_0 are satisfied, hence, we also have $Z^{*'}Z^*/n \to Q_Z$, $Z^{*'}u^*/n \to Q_Z b_0$, $Z^{*'}v_2^*/n \to 0$, and $(u^*:v_2^*)'(u^*:v_2^*)/n \to \Sigma_V$, $a.s.$

It will be illuminating to deal separately with the following two subcases : (a) $\pi \neq 0$ is fixed (strong instrument), and (b) $\pi = \pi_0/\sqrt{n}$, where π_0 is fixed (weak instruments).

(a) First, suppose that $\pi \neq 0$ is fixed. From what precedes, it is easy to see that, given $\mathscr{X}_n$, we have $\frac{1}{n}y_2^{*'}P_{Z^*}y_2^* \to \pi'Q_Z\pi$, $\frac{1}{n}y_2^{*'}y_2^* \to \pi'Q_Z\pi + \sigma_{v_2}^2$, $\hat{\sigma}^{*^2} \to \sigma_e^2$, and $\hat{\omega}_3^{*^2} \to \sigma_e^2[(\pi'Q_Z\pi)^{-1} - (\pi'Q_Z\pi + \sigma_{v_2}^2)^{-1}]$, a.s. By observing that $\frac{1}{\sqrt{n}}y_2^{*'}P_{Z^*}u^* = \pi'(Z^{*'}u^*/\sqrt{n}) + o_p(1)$, $\frac{1}{\sqrt{n}}y_2^{*'}u^* = \pi'(Z^{*'}u^*/\sqrt{n}) + v_2^{*'}u^*/\sqrt{n} + o_p(1)$ $a.s.$, and by using Lemma **4.2**, we can show as in Lemma **3.1**-(b) that $\sqrt{n}\frac{(\tilde{\beta}^* - \hat{\beta}^*)}{\hat{\omega}_3^*}|\mathscr{X}_n \to N(\bar{\tau}_{b_0},\, 1)$ $a.s.$, where $\bar{\tau}_{b_0} = \frac{1}{\sigma_e}[(\pi'Q_Z\pi)^{-1} - (\pi'Q_Z\pi + \sigma_{v_2}^2)^{-1}]^{1/2}\pi'Q_Z b_0$. Thus we get

$$DW_3^*|\mathscr{X}_n \to \chi^2(1; \| \bar{\tau}_{b_0} \|^2) \;\; a.s. \tag{A.12}$$

from (A.11).

(b) The proof for $\pi = \pi_0/\sqrt{n}$, where π_0 is fixed follows the same steps as above, therefore, it is omitted.

□

PROOF OF COROLLARY **4.4** Corollary **4.4** follows directly from Lemmas **3.1**-(a), **3.2**-(b) and Lemma **4.3**. Therefore, the proof is omitted. □

PROOF OF THEOREM **4.5** From Corollary **4.4**, we have $|\, \mathbb{P}(T_l > x) - \mathbb{P}^*(T_l^* > x) \,| = o_p(1)$ uniformly over $x \in \mathbb{R}$, for all $l = 2,3,4$ and all values of π. Thus $|\, \mathbb{P}(T_l > c^*_{T_l,\alpha}) - \mathbb{P}^*(T_l^* > c^*_{T_l,\alpha}) \,| = |\, \mathbb{P}(T_l > c^*_{T_l,\alpha}) - \alpha \,| = o_p(1)$. Similarly, we $|\, \mathbb{P}(DW_j > c^*_{DW_j,\alpha}) - \alpha \,| = o_p(1)$. Theorem **4.5** follows from an expected value argument.

□

References

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

Anderson, T. W., 1951. Estimating linear restrictions on regression coefficients for multivariate normal distributions. Annals of Mathematical Statistics 22, 327–351.

Anderson, T. W., 1971. The Statistical Analysis of Time Series. Wiley, New York.

Anderson, T. W., Kadane, J. B., 1977. A comment on the test of overidentifying restrictions. Econometrica 45, 1027–1031.

Anderson, T. W., Rubin, H., 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. Annals of Mathematical Statistics 20, 46–63.

Anderson, T. W., Rubin, H., 1950. The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations. Annals of Mathematical Statistics 21, 570–582.

Anderson, T. W., Rubin, H., 1976. Estimation of linear functional relationships: Approximate distributions and connections with simultaneous equations in econoetrics. Journal of the Royal Statistical Society, Series B 38, 1–26.

Andrews, D. W. K., 2002. High-order improvement of a computationally attractive k-step bootstrap for extremum estimators. Econometrica 70(1), 119–162.

Andrews, D. W. K., Stock, J. H., 2007. Inference with weak instruments. In: R. Blundell, W. Newey , T. Pearson, eds, Advances in Economics and Econometrics, Theory and Applications, 9th Congress of the Econometric Society Vol. 3. Cambridge University Press, Cambridge, U.K., chapter 6.

Ashley, R., 2009. Assessing the credibilility of instrumental variables inference with imperfect instruments via sensitivity analysis. Journal of Applied Econometrics 24(2), 325–337.

Basmann, R. L., 1960. On the asymptotic distributions of generalized classical linear estimators. Econometrica 28, 97–107.

Baum, C., Schaffer, M., Stillman, S., 2003. Instrumental variables and GMM: Estimation and testing. Stata Journal 3(1), 1–30.

Bazzi, S., Clemens, A. M., 2009. Blunt instruments: A cautionary note on establishing the causes of economic growth. Technical report, Center for Global Development N0. 171.

Bazzi, S., Clemens, A. M., 2013. Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth. American Economic Journal: Macroeconomics 5(3), 152–186.

Berkowitz, D., Caner, M., Fang, Y., 2008. Are nearly exogenous instruments reliable?. Economics Letters 101, 20–23.

Berkowitz, D., Caner, M., Fang, Y., 2012. The validity of instruments revisited. Journal of Econometrics 166, 255–266.

Bound, J., Jaeger, D. A., Baker, R. M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. Journal of the American Statistical Association 90, 443–450.

Chmelarova, V., Hill, R., 2010. The Hausman pretest estimator. Economics Letters 108, 96–99.

Davidson, R., Mackinnon, J., 1993. Econometric Theory and Methods. Oxford University Press, New York, New York.

Doko Tchatoka, F., 2013. On bootstrap validity for specification tests with weak instruments. Technical report, School of Economics and Finance, University of Tasmania Hobart, Australia.

Doko Tchatoka, F., 2014. Subset hypotheses testing and instrument exclusion in the linear IV regression. Econometric Theory forthcoming.

Doko Tchatoka, F., Dufour, J.-M., 2008. Instrument endogeneity and identification-robust tests: some analytical results. Journal of Statistical Planning and Inference 138(9), 2649–2661.

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

Doko Tchatoka, F., Dufour, J.-M., 2014. Identification-robust inference for endogeneity parameters in linear structural models. The Econometrics Journal 17, 165–187.

Dufour, J.-M., 1979. Methods for Specification Errors Analysis with Macroeconomic Applications PhD thesis University of Chicago. 257 + XIV pages.

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., 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.

Dufour, J.-M., Hsiao, C., 2008. Identification. In: L. E. Blume, S. N. Durlauf, eds, The New Palgrave Dictionary of Economics second edn. Palgrave Macmillan, Basingstoke, Hampshire, England. forthcoming.

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.

Engle, R. F., Hendry, D. F., Richard, J.-F., 1982. Exogeneity. Econometrica 51, 277–304.

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.

Guggenberger, P., 2011. On the asymptotic size distortion of tests when instruments locally violate the exogeneity assumption. Econometric Theory forthcoming.

Guggenberger, P., Kleibergen, F., Mavroeidis, S., Chen, L., 2012. On the asymptotic sizes of subset anderson-rubin and lagrange multiplier tests in linear instrumental variables regression. Econometrica 80(6), 2649–2666.

Hahn, J., Ham, J., Moon, H. R., 2010. The Hausman test and weak instruments. Journal of Econometrics 160, 289–299.

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

Hansen, L. P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

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

Hausman, J., Hahn, J., 2005. Estimation with valid and invalid instruments. Annales d'Économie et de Statistique 79–80, 25–57.

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.

Holly, A., 1983a. Tests d'exogénéité dans un modèle à equations simultanées: Énoncé de résultats théoriques en information limitée et illustrations à des tests de dépendance de la politique monétaire en régime de changes fixes. Cahiers du Séminaire d'Économétrie 25, 49–69.

Holly, A., 1983b. Une présentation unifiée des tests d'exogénéity dans les modèles à equations simultanées. Annales de l'INSEE 50, 3–24.

Holly, A., Monfort, A., 1983. Some useful equivalence properties of Hausman's test. Economics Letters 20, 39–43.

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

Hwang, H.-S., 1985. The equivalence of Hausman and Lagrange multiplier tests of independence between disturbance and a subset of stochastic regressors. Economics Letters 17, 83–86.

Imbens, G. W., 2003. Sensitivity to exogeneity assumptions in program evaluation. American Economic Review 93(2), 126–132.

Imbens, G. W., Kolesár, M., Chetty, R., Friedman, J., Glaeser, E., 20011. Inference and identification with many invalid instruments. Technical report, Department of Economics, Havard University Boston, MA.

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.

Kiviet, J. F., 2013. Identification and inference in a simultaneous equation under alternative information sets and sampling schemes. The Econometrics Journal 16, S24–S59.

Kiviet, J. F., Niemczyk, J., 2007. The asymptotic and finite-sample distributions of OLS and simple IV in simultaneous equations. Computational Statistics and Data Analysis 51, 3296–3318.

Kiviet, J. F., Niemczyk, J., 2012. Comparing the asymptotic and empirical (un)conditional distributions of OLS and IV in a linear static simultaneous equation. Computational Statistics and Data Analysis 56, 3567–3586.

Kiviet, J. F., Pleus, M., 2012. The performance of tests on endogeneity of subsets of explanatory variables scanned by simulation. Technical report, Amsterdam School of Economics Amsterdam, The Netherlands.

Kleibergen, F., 2002. Pivotal statistics for testing structural parameters in instrumental variables regression. Econometrica 70(5), 1781–1803.

Kleibergen, F., 2004. Testing subsets of structural coefficients in the IV regression model. Review of Economics and Statistics 86, 418–423.

Kleibergen, F., Mavroeidis, S., 2011. Inference on subsets of parameters in GMM without assuming identification. Technical report, Department of Economics, Brown University, Providence, Rhode Island.

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

Meepagala, G., 1992. On the finite sample performance of exogeneity tests of Revankar, Revankar and Hartley and Wu-Hausman. Econometric Reviews 11, 337–353.

Mikusheva, A., 2013. Survey on statistical inferences in weakly-identified instrumental variable models. Applied Econometrics 29(1), 117–131.

Moreira, M. J., 2003. A conditional likelihood ratio test for structural models. Econometrica 71(4), 1027–1048.

Muirhead, R. J., 2005. Aspects of Multivariate Statistical Theory. John Wiley & Sons, Inc., Hoboken, New Jersey.

Murray, P. M., 2006. Avoiding invalid instruments and coping with weak instruments. The Journal of Economic Perspectives 20(4), 111–132.

Nagar, A. L., 1959. The bias and moment matrix of the genaralizea *k*-class estimators of the parameters in simultaneous equations. Econometrica 27, 575–595.

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

Nakamura, A., Nakamura, M., 1985. On the performance of tests by Wu and by Hausman for detecting the ordinary least squares bias problem. Journal of Econometrics 29, 213–227.

Newey, W. K., 1985a. Generalized method of moments specification 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.

Richardson, D. H., 1968. The exact distribution of a structural coefficient estimator. Journal of the American Statistical Association 63, 1214–1226.

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

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

Sargan, J., 1958. The estimation of economic relationships using instrumental variables. Econometrica 26(3), 393–415.

Sargan, J., 1983. Identification and lack of identification. Econometrica 51, 1605–1633.

Sawa, T., 1969. The exact sampling distribution of ordinary least squares and two-stage least squares estimators. Journal of the American Statistical Association 64, 923–937.

Small, D. S., 2007. Sensitivity analisis for instrumental variables regression with overidentifying restrictions. Annals of Mathematical Statistics 102, 1049–1058.

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., 1985. Wald tests for the independence of stochastic variables and disturbance of a single linear stochastic simultaneous equation. Economics Letters 17, 87–90.

Smith, R. J., Pesaran, M., 1990. A unified 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. Erratum, Econometrica, Vol. 50, No. 4 (Jul., 1982), p. 1087.

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

Stock, J. H., Trebbi, F., 2003. Who invented instrumental variable regression?. Journal Economic Perspectives 17(3), 177–194.

Stock, J. H., Wright, J. H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20(4), 518–529.

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

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

Wong, K., 1997. Effect on inference of pretesting the exogeneity of a regressor. Economics Letters 56, 267–271.

Wright, P. G., 1928. The Tariff on Animal and Vegetable Oils. Macmillan, New York.

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.