



Munich Personal RePEc Archive

On the Validity of Durbin-Wu-Hausman Tests for Assessing Partial Exogeneity Hypotheses with Possibly Weak Instruments

Doko Tchatoka, Firmin

School of Economics and Finance, University of Tasmania

20 July 2012

Online at <https://mpra.ub.uni-muenchen.de/40184/>
MPRA Paper No. 40184, posted 20 Jul 2012 10:58 UTC

On the Validity of Durbin-Wu-Hausman Tests for Assessing Partial Exogeneity Hypotheses with Possibly Weak Instruments

Firmin Doko Tchatoka

School of Economics and Finance

University of Tasmania, Private Bag 85, Hobart TAS 7001

Tel: +613 6226 7226; Fax:+61 3 6226 7587; e-mail: Firmin.dokotchatoka@utas.edu.au

July 6, 2012

ABSTRACT

We investigate the validity of the standard specification tests for assessing the exogeneity of subvectors in the linear IV regression. Our results show that ignoring the endogeneity of the regressors whose exogeneity is not being tested leads to invalid tests (level is not controlled). When the fitted values from the first stage regression of these regressors are used as instruments under the partial null hypothesis of interest, as suggested Hausman and Taylor (1980, 1981), some versions of these tests are invalid when identification is weak and the number of instruments is moderate. However, all tests are overly conservative and have no power when the number of instruments increases, even for moderate identification strength.

Key words: Partial exogeneity; sized distortions; weak identification.

JEL classification: C12; C15; C3.

1. Introduction

Exogeneity tests of the type proposed by Durbin (1954), Wu (1973), and Hausman (1978), henceforth DWH tests, are widely used in applied work to decide whether ordinary least squares (OLS) or instrumental variable (IV) methods is appropriate. It is now well known that if the null hypothesis of interest is specified on the whole set of supposedly endogenous regressors, DWH tests are valid even when instruments are weak [see Staiger and Stock (1997), Hahn et al. (2010), Guggenberger (2010), and Doko Tchatoka and Dufour (2011a, 2011b)]. However, their validity when testing for partial exogeneity is unclear, especially when identification is weak.

Many economic questions often involve more than one supposedly endogenous variable, but researcher may want to challenge the exogeneity of only a subset of them. This may be particularly motivated by some economic theory. Furthermore, efficient estimation of model parameters requires to use available instruments only for regressors that are endogenous. Thus, assessing individual exogeneity of supposedly endogenous regressors is an important issue.

In this paper, we focus on the linear IV regression and question the validity of two testing practices, based on DWH type statistics, for assessing partial exogeneity hypotheses. The first is the use of Sargan-Hansen C-type tests ignoring the endogeneity of the regressors whose exogeneity is not being tested [see Chaudhuri and Rose (2009)]. The second is the extension of DWH tests to partial exogeneity hypotheses, proposed by Hausman and Taylor (1980, 1981). These procedures use the fitted values from the first stage regression of the regressors whose exogeneity is not being investigated as instruments under the partial null hypothesis of interest. Both practices are widely used in applied work.

In both cases, we examine the size and power of the corresponding tests through a Monte Carlo experiment, allowing for the presence of weak instruments. We find that the use of either C-tests or Hausman and Taylor (1980, 1981) extension may lead to misleading conclusions even when identification strength is moderate.

The remainder of the paper is organized as follows. Section 2 formulates the model and presents the statistical problem of interest. Section 3 analyze the behavior of the tests through a Monte Carlo experiment. Conclusions are drawn in Section 4.

2. Model, statistical problem and test statistics

We consider the following linear IV regression model

$$y = Y\beta + X\theta + u, \quad (2.1)$$

$$Y = Z\Pi + v, \quad X = Z\Gamma + \xi \quad (2.2)$$

where $y \in \mathbb{R}^n$ is a vector of observations on a dependent variable, $Y \in \mathbb{R}^{n \times m_y}$ and $X \in \mathbb{R}^{n \times m_x}$ ($m_y + m_x = m \geq 1$) are two matrices of (possibly) endogenous explanatory variables, $Z \in \mathbb{R}^{n \times l}$ is a matrix of exogenous instruments, $u = (u_1, \dots, u_n)' \in \mathbb{R}^n$ is the vector of structural disturbances, $v \in \mathbb{R}^{n \times m_y}$ and $\xi \in \mathbb{R}^{n \times m_x}$ are matrices of reduced form disturbances, $\beta \in \mathbb{R}^{m_y}$ and $\theta \in \mathbb{R}^{m_x}$ are unknown structural parameter vectors, while $\Pi \in \mathbb{R}^{l \times m_y}$ and $\Gamma \in \mathbb{R}^{l \times m_x}$ are unknown reduced form coefficient matrices. Model (2.1)-(2.2) can be modified to include exogenous variables Z_1 . If so, our results do not alter qualitatively if we replace the variables that are currently in (2.1)-(2.2) by the residuals that result from their projection onto Z_1 . We shall assume that the instrument matrix Z has full-column rank l with probability one and $l \geq m$.

The usual necessary and sufficient condition for identification of model (2.1)-(2.2) is $\text{rank}([\Pi, \Gamma]) = m$. If $[\Pi, \Gamma]$ is close not to have full rank, $(\beta', \theta)'$ is not identifiable but some of its linear combinations are ill-determined by the data, a situation often called “weak identification” in this type of setup [see Andrews and Stock (2006), Dufour (2003)]. When $[\Pi, \Gamma] = 0$, the instruments Z are irrelevant and $(\theta', \beta)'$ is completely unidentified.

We will study in turn the problem of testing the partial exogeneity of Y , *i.e.* the hypothesis

$$H_0^p : \text{cov}(Y, u) = \sigma_{vu} = 0 \quad (2.3)$$

where the regressors X whose exogeneity is not being tested may be endogenous. Hausman and Taylor (1980) [see also Hausman and Taylor (1981, p.1389, footnote 10)] propose to assess H_0^p via the standard DWH specification tests upon comparing the two stage least squares estimator $(\hat{\beta}'_{0,2SLS}, \hat{\theta}'_{0,2SLS})'$ obtained using $[Y, \hat{X}]$ as instruments, to those $(\hat{\beta}'_{1,2SLS}, \hat{\theta}'_{1,2SLS})'$ obtained from $[\hat{Y}, \hat{X}]$, where $[\hat{Y}, \hat{X}] = \mathbf{P}_Z[Y, X]$ and $\mathbf{P}_Z = Z(Z'Z)^{-1}Z'$.

This is mathematically equivalent¹ to test the exogeneity of Y in model

$$y = Y\beta + \hat{X}\theta + e, \quad (2.4)$$

$$Y = Z\Pi + v. \quad (2.5)$$

Which means that the corresponding DWH statistics can be written as a quadratic form in the subvector of differences of the 2SLS estimates of β alone. An important question is whether the tests obtained in such a way are still valid, especially when identification is weak. The problem stems from the fact that when instruments are weak, \hat{X} is not independent of u even in large-sample. So, using \hat{X} as instruments under H_0^p may be problematic.

From (2.4)-(2.5), we can express $\hat{\beta}_{0,2SLS}$ and $\hat{\beta}_{1,2SLS}$ as: $\hat{\beta}_{0,2SLS} = (Y'\mathbf{M}_{\hat{X}}Y)^{-1}Y'\mathbf{M}_{\hat{X}}y$, $\hat{\beta}_{1,2SLS} = (\hat{Y}'\mathbf{M}_{\hat{X}}\hat{Y})^{-1}\hat{Y}'\mathbf{M}_{\hat{X}}y$, where $\mathbf{M}_{\hat{X}} = \mathbf{I}_n - \mathbf{P}_{\hat{X}}$. We consider four alternative DWH type statistics written in the following unified formulation:

$$\mathscr{W}_j = \kappa_j(\hat{\beta}_{1,2SLS} - \hat{\beta}_{0,2SLS})'\hat{\Sigma}_j^{-1}(\hat{\beta}_{1,2SLS} - \hat{\beta}_{0,2SLS}), j = 1, 2, 3, 4 \quad (2.6)$$

$$\begin{aligned} \hat{\Sigma}_1 &= \tilde{\sigma}_2^2\hat{\Delta}, \hat{\Delta} = \hat{\Omega}_{IV}^{-1} - \hat{\Omega}_{LS}^{-1}, \hat{\Sigma}_2 = \tilde{\sigma}^2\hat{\Omega}_{IV}^{-1} - \hat{\sigma}^2\hat{\Omega}_{LS}^{-1}, \hat{\Sigma}_3 = \tilde{\sigma}^2\hat{\Delta}, \hat{\Sigma}_4 = \hat{\sigma}^2\hat{\Delta}, \\ \hat{\Omega}_{IV} &= \hat{Y}'\mathbf{M}_{\hat{X}}\hat{Y}/n, \hat{\Omega}_{LS} = Y'\mathbf{M}_{\hat{X}}Y/n, \tilde{\sigma}^2 = (y - Y\hat{\beta}_{1,2SLS})'\mathbf{M}_{\hat{X}}(y - Y\hat{\beta}_{1,2SLS})/n, \\ \tilde{\sigma}_2^2 &= \hat{\sigma}^2 - (\hat{\beta}_{1,2SLS} - \hat{\beta}_{0,2SLS})'\hat{\Delta}^{-1}(\hat{\beta}_{1,2SLS} - \hat{\beta}_{0,2SLS}), \kappa_1 = (n - 2m_y)/m_y \\ \hat{\sigma}^2 &= (y - Y\hat{\beta}_{0,2SLS})'\mathbf{M}_{\hat{X}}(y - Y\hat{\beta}_{0,2SLS})/n, \kappa_i = n, j = 2, 3, 4. \end{aligned}$$

The statistics in (2.6) differ only through the variance estimators of the errors in (2.4)-(2.5) and the scaling factors κ_j . $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ are the usual IV-based estimators (without correction for degrees of freedom), while $\tilde{\sigma}_2^2$ can be interpreted as an alternative IV-based scaling factor. The statistic \mathscr{W}_1 is an extension of Wu (1973) T_2 -statistic. \mathscr{W}_j ($j \geq 2$) are analogues to alternative Hausman (1978) type-statistics studied by Staiger and Stock (1997), Guggenberger (2010), and Doko Tchatoka and Dufour (2011a, 2011b). The Sargan-Hansen C-tests ignoring the endogeneity of X consist of replacing \hat{X} by X in (2.6), so that $\hat{\beta}_{0,2SLS}$ collapses to the OLS estimator of β in (2.1).

¹See Baum et al. (2003) for further details.

Following Hausman and Taylor (1980, 1981), H_0^p is rejected if the corresponding statistic is greater than the desired critical value of a $\chi^2(m_y)$ -distribution. This however requires \hat{X} be independent of the structural error, which is not guaranteed when instruments are weak.

We shall now analyze the size and power of the tests through a Monte Carlo experiment, in both strong and weak identification setups.

3. Simulation results

In each of the following experiments, data are generated from model (2.1)-(2.2) with three endogenous regressors ($Y : n \times 2$, $W : n \times 1$) by setting

$$\begin{aligned} u_i &= (1 + \rho_{v_1}^2 + \rho_{v_2}^2 + \rho_\xi^2)^{-1/2}(\varepsilon_{1i} + \rho_{v_1}\varepsilon_{2i} + \rho_{v_2}\varepsilon_{3i} + \rho_\xi\varepsilon_{4i}), \\ v_i &= [(1 + \rho_{v_1}^2)^{-1/2}(\rho_{v_1}\varepsilon_{1i} + \varepsilon_{2i}), (1 + \rho_{v_2}^2)^{-1/2}(\rho_{v_2}\varepsilon_{1i} + \varepsilon_{3i})], \\ \xi_i &= (1 + \rho_\xi^2)^{-1/2}(\rho_\xi\varepsilon_{1i} + \varepsilon_{4i}), (\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}, \varepsilon_{4i})' \sim \mathbf{IIDN}(0, \mathbf{I}_4) \end{aligned} \quad (3.7)$$

for all $i = 1, \dots, n$, $\rho_{v_1} \in \{-.9, -.2, 0, .3, .8\}$, $\rho_{v_2} = \rho_{v_1}/\sqrt{3}$ and $\rho_\xi = 0.4$. The results are qualitatively the same for alternative choices of ρ_ξ and ρ_{v_2} . The exogeneity of Y is then expressed as $H_0^p : \rho_{v_1} = 0$. Z contains $l \in \{3, 20\}$ instruments each generated $\mathbf{IIDN}(0, 1)$ and is kept fix within experiment. The true value of $(\beta', \theta)'$ is $(2, -3, 1/2)'$ and $[\mathbf{\Pi}, \mathbf{\Gamma}]$ is chosen as: $[\mathbf{\Pi}, \mathbf{\Gamma}] = \sqrt{\frac{\mu^2}{n\|\mathbf{ZC}\|}}\mathbf{\Pi}_0$, where $\mathbf{\Pi}_0$ is obtained by taking the first three columns of the identity matrix of dimension l , \mathbf{C} is an $l \times 1$ vector of ones, and $\mu^2 \in \{0, 10, 1000, 10000\}$ is the concentration parameter that characterizes the strength of the instruments. The values $\mu^2 \leq 613$ correspond to weak instruments while $\mu^2 > 613$ is for strong instruments [see Hansen et al. (2008)]. In particular, when $\mu^2 = 0$, the instruments are irrelevant and $(\beta', \theta)'$ is completely unidentified. The simulations are run with sample sizes $n = 100$ and 500 . The number of replications is $N = 10000$ and the empirical rejections of the statistics are computed using the 95% critical value value of $\chi^2(2)$.

Table 1 presents the results. In the first column, we report the statistics while the second column reports the number of instruments. In the other columns, for each value of endogeneity ρ_{v_1} and instrument quality μ^2 , the rejection frequencies are reported.

The first part of the table deals with the setup where the endogeneity of X is ig-

nored. As seen, all tests are seriously size distorted when identification is strong [column $\rho_{v_1} = 0, \mu^2 \in \{1000, 10000\}$]. The size distortions persists even for $n = 500$. When instruments are irrelevant or weak [column $\rho_{v_1} = 0, \mu^2 \in \{0, 10\}$], \mathcal{W}_1 and \mathcal{W}_4 are still size distorted but \mathcal{W}_2 and \mathcal{W}_3 are overly conservative. In all cases, the maximal size distortion of the tests approaches 40% with 3 instruments and around 50% with 20 instruments.

The second part of the table corresponds to the setup by Hausman and Taylor (1980,1981), where $[Y, \hat{X}]$ is used as instruments under H_0^p . Two observations emerge in this case depending on the sample size. Firstly, when $n = 100$, all tests have correct level when identification is very strong [column $\rho_{v_1} = 0, \mu^2 = 10000$], but are conservative for moderate identification strength [column $\rho_{v_1} = 0, \mu^2 = 1000$]. However, when identification is weak, \mathcal{W}_1 and \mathcal{W}_4 are size distorted when $l = 3$ but overly conservative when $l = 20$. At the same time, \mathcal{W}_2 and \mathcal{W}_3 are overly conservative in all cases. Secondly, when $n = 500$, all tests are conservative when identification is strong or moderate. With weak instruments, \mathcal{W}_1 and \mathcal{W}_4 are still size distorted when $l = 3$, with maximal distortion around 12%. This suggests that the null asymptotic distributions of the statistics is far from a $\chi^2(2)$ -distribution, even with moderate identification strength. Moreover, we can see that even when $\mu^2 = 1000$ (moderate identification), the tests have no power when $l = 20$ and $\rho_{v_1} = -.2; .3$. The sample realizations of the statistics are often close to zero even for a relatively high endogeneity. Overall, the use of DWH tests to assess partial exogeneity hypotheses, as suggested Hausman and Taylor (1980,1981), may be misleading if identification is not very strong. The use of C-type tests, as described here, is inappropriate and should be avoided.

Table 1. Size and power at nominal level 5%

Rejections when the endogeneity of X is ignored																					
		$n = 100$																			
Statistics	$k_2 \downarrow \mu^2 \rightarrow$	$\rho_1 = -.9$				$\rho_1 = -.2$				$\rho_1 = 0$				$\rho_1 = .3$				$\rho_1 = .8$			
		0	10	1000	10000	0	10	1000	10000	0	10	1000	10000	0	10	1000	10000	0	10	1000	10000
\mathcal{W}_1	3	31.1	65.8	89.6	87.4	30.0	42.9	79.5	96.7	29.6	38.6	37.6	39.3	29.9	46.1	94.2	99.7	31.6	62.7	93.1	99.4
\mathcal{W}_2	3	1.5	4.9	100.0	100.0	0.8	1.7	74.7	96.0	0.7	1.7	31.2	35.5	1.0	2.2	92.4	99.7	1.6	6.8	100.0	100.0
\mathcal{W}_3	3	1.8	5.9	100.0	100.0	1.0	2.2	76.8	96.5	0.9	2.2	33.7	38.0	1.3	2.6	93.1	99.7	2.0	8.1	100.0	100.0
\mathcal{W}_4	3	40.0	85.5	100.0	100.0	38.1	53.3	79.1	96.6	37.4	49.0	37.1	38.4	37.6	58.5	94.2	99.7	40.3	96.3	100.0	100.0
\mathcal{W}_1	20	9.0	12.7	99.3	100.0	7.2	7.8	38.9	80.2	7.0	7.1	17.1	13.7	7.4	8.1	54.1	95.9	8.7	19.0	99.5	100.0
\mathcal{W}_2	20	3.4	4.7	98.8	100.0	2.0	2.1	29.7	76.3	1.7	1.9	11.3	11.4	2.1	2.4	44.4	94.7	3.1	8.8	99.2	100.0
\mathcal{W}_3	20	4.2	5.9	99.0	100.0	2.7	2.8	32.6	78.7	2.5	2.4	13.2	12.7	2.8	3.2	48.1	95.5	3.9	10.5	99.3	100.0
\mathcal{W}_4	20	8.6	12.1	99.3	100.0	6.9	7.5	38.0	79.4	6.6	6.8	16.6	13.2	7.1	7.7	53.3	95.7	8.4	18.3	99.5	100.0
		$n = 500$																			
\mathcal{W}_1	3	35.6	63.3	99.3	99.8	34.2	40.6	80.3	99.8	33.3	39.3	50.2	32.7	33.6	44.3	95.5	100.0	35.0	80.0	93.1	100.0
\mathcal{W}_2	3	1.4	3.1	99.4	100.0	0.7	1.4	73.0	99.8	0.7	1.1	39.8	31.3	1.0	1.4	92.6	100.0	1.5	4.5	99.5	100.0
\mathcal{W}_3	3	1.5	3.2	99.4	100.0	0.7	1.5	73.4	99.8	0.7	1.1	40.4	31.6	1.0	1.4	92.8	100.0	1.5	4.6	99.5	100.0
\mathcal{W}_4	3	39.5	70.4	100.0	100.0	37.7	44.9	80.3	99.8	36.7	43.3	50.1	32.6	37.1	48.8	95.6	100.0	38.9	90.9	100.0	100.0
\mathcal{W}_1	20	8.0	10.1	95.7	100.0	6.6	6.8	30.7	91.2	6	6.7	20.8	29.6	6.9	7.4	48.9	99.7	7.8	15.0	97.8	100.0
\mathcal{W}_2	20	3.0	3.7	93.9	100.0	1.7	1.9	21.5	90.3	1.4	1.7	13.2	27.2	1.8	1.8	38.4	99.6	2.8	6.7	97.0	100.0
\mathcal{W}_3	20	3.1	3.9	93.9	100.0	1.7	2.0	22.1	90.5	1.6	1.8	13.6	27.8	2.0	1.9	39.1	99.6	2.9	6.9	97.1	100.0
\mathcal{W}_4	20	7.9	10.0	95.7	100.0	6.5	6.7	30.5	91.2	5.9	6.6	20.7	29.5	6.8	7.3	48.8	99.7	7.7	14.8	97.8	100.0
Rejections when $[Y, \hat{X}]$ is used as instruments under H_0^p																					
		$n = 100$																			
Statistics	$k_2 \downarrow \mu^2 \rightarrow$	$\rho_1 = -.9$				$\rho_1 = -.2$				$\rho_1 = 0$				$\rho_1 = .3$				$\rho_1 = .8$			
		0	10	1000	10000	0	10	1000	10000	0	10	1000	10000	0	10	1000	10000	0	10	1000	10000
\mathcal{W}_1	3	7.3	18.3	99.9	100.0	8.1	11.1	39.4	75.4	8.5	11.8	3.0	5.2	7.4	11.2	81.5	98.1	3.5	17.8	100.0	100.0
\mathcal{W}_2	3	0.3	1.3	100.0	100.0	0.3	0.8	29.0	71.3	0.3	0.7	1.7	3.8	0.3	0.7	72.2	97.3	0.2	1.0	100.0	100.0
\mathcal{W}_3	3	0.3	1.7	100.0	100.0	0.4	0.8	33.0	74.1	0.3	0.8	2.2	4.7	0.4	0.9	76.2	97.9	0.2	1.3	100.0	100.0
\mathcal{W}_4	3	9.1	23.1	100.0	100.0	10.8	14.2	38.3	74.4	10.9	15.0	2.9	4.8	9.5	14.8	80.6	98.0	4.3	21.9	100.0	100.0
\mathcal{W}_1	20	17.8	20.2	99.7	100.0	0.3	0.4	13.9	56.9	0.0	0.1	1.5	5.2	0.6	0.9	25.4	90.6	13.3	15.7	99.2	100.0
\mathcal{W}_2	20	6.9	8.1	99.1	100.0	0.1	0.1	7.9	51.4	0	0	0.7	3.71	0.2	0.2	15.9	88.0	4.9	5.7	97.8	100.0
\mathcal{W}_3	20	8.7	10.5	99.3	100.0	0.1	0.1	9.7	54.8	0	0	0.9	4.7	0.3	0.3	19.1	89.6	6.4	7.5	98.3	100.0
\mathcal{W}_4	20	16.8	19.4	99.6	100.0	0.2	0.4	13.3	55.8	0	0.1	1.3	4.9	0.6	0.8	24.5	90.1	12.6	14.8	99.2	100.0
		$n = 500$																			
\mathcal{W}_1	3	7.5	14.2	100.0	100.0	9.0	10.6	18.6	98.5	9.39	10.7	1.6	1.2	8.0	10.9	58.0	100.0	3.7	11.2	100.0	100.0
\mathcal{W}_2	3	0.3	0.7	98.1	100.0	0.2	0.5	8.6	98.2	0.3	0.3	0.7	1.0	0.3	0.5	35.4	100.0	0.2	0.5	98.4	100.0
\mathcal{W}_3	3	0.3	0.7	98.3	100.0	0.2	0.6	9.0	98.4	0.3	0.3	0.8	1.1	0.3	0.5	36.9	100.0	0.2	0.5	98.5	100.0
\mathcal{W}_4	3	8.5	15.8	100.0	100.0	10.1	11.9	18.4	98.5	10.5	12.1	1.6	1.1	9.0	12.2	57.7	100.0	4.2	12.4	100.0	100.0
\mathcal{W}_1	20	5.7	6.9	93.1	100.0	0.0	0.0	0.3	69.3	0	0	0	0.5	0.0	0.0	2.7	99.1	2.3	2.5	89.5	100.0
\mathcal{W}_2	20	1.4	1.8	80.3	100.0	0.0	0.0	0.1	64.8	0	0	0	0.4	0.0	0.0	0.8	98.8	0.5	0.6	75.0	100.0
\mathcal{W}_3	20	1.5	1.9	81.0	100.0	0.0	0.0	0.1	65.6	0	0	0	0.4	0.0	0.0	0.9	98.8	0.6	0.7	75.9	100.0
\mathcal{W}_4	20	5.5	6.8	93.0	100.0	0.0	0.0	0.3	69.0	0	0	0	0.5	0.0	0.0	2.6	99.1	2.3	2.4	89.4	100.0

4. Concluding remarks

In this paper, we question the validity of the use of standard DWH specification tests for assessing partial exogeneity hypotheses. Our results show that these tests are invalid (level is not controlled) if the endogeneity of the regressors whose exogeneity is not being tested is ignored within inference. When the fitted values from the first stage regression of these regressors are used as instruments under the subset null hypothesis of interest, we find that some versions of the tests are invalid when identification is weak and the number of instruments is relatively moderate. However, all tests are overly conservative and have no power when the number of instruments increases, even for moderate identification strength. So, applying them to assess partial exogeneity hypotheses, as it is often the case in applied work, may be misleading. This underscores the importance to develop tests that do not exhibit these problems when testing for partial exogeneity.

Acknowledgments

The author thanks Mardi Dungey (University of Tasmania), Jean-Marie Dufour (McGill University) and Russell Davidson (McGill University) for helpful comments and suggestions.

References

- Andrews, D. W. K. , Stock , 2006. 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.
- Baun, C. F., Schaffer, M. E. , Stillman, S. , 2003. Instrumental variable and GMM: Estimation and testing. Technical report, Department of Economics, Boston College, USA Boston, USA.
- Chaudhuri, S. , Rose, E. , 2009. Estimating the veteran effect with endogenous schooling when instruments are potentially weak. Technical report, IZA Discussion Papers No. 4203.

- 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., 2003. Identification, weak instruments and statistical inference in econometrics. *Canadian Journal of Economics* 36(4), 767–808.
- Durbin, J., 1954. Errors in variables. *Review of the International Statistical Institute* 22, 23–32.
- Guggenberger, P., 2010. The impact of a Hausman pretest on the size of the hypothesis tests. *Econometric Theory* 156, 337–343.
- 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.
- Hausman, J., 1978. Specification tests in econometrics. *Econometrica* 46, 1251–1272.
- Hausman, J., Taylor, W., 1981. Panel data and unobservable individual effects. *Econometrica* 49(6), 1377–1398.
- Hausman, J., Taylor, W. E., 1980. Comparing specification tests and classical tests. Technical report, Department of Economics, MIT Boston, MA.
- Staiger, D., Stock, J. H., 1997. Instrumental variables regression with weak instruments. *Econometrica* 65(3), 557–586.
- Wu, D.-M., 1973. Alternative tests of independence between stochastic regressors and disturbances. *Econometrica* 41, 733–750.