# Exogeneity tests and estimation in IV regressions*

Firmin Doko Tchatoka[†] Jean-Marie Dufour [‡]
University of Tasmania McGill University

First version: May 2007
Revised: October 2007, November 2008, December 2009, March 2010
This version: January 2012
Compiled: January 26, 2012, 9:44

*The authors thank Atsushi Inoue, Marine Carrasco, Jan Kiviet and Benoit Perron for several useful comments. This work was supported by the William Dow Chair in Political Economy (McGill University), the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Bank of Canada (Research Fellowship), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Institut de finance mathématique de Montréal (IFM2), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, the Fonds de recherche sur la société et la culture (Québec), and the Fonds de recherche sur la nature et les technologies (Québec), and a Killam Fellowship (Canada Council for the Arts).

† 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. Homepage: http://www.fdokotchatoka.com

‡ William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

ABSTRACT

Weak identification is likely to be prevalent in many economic models. When instruments are weak, the limiting distributions of standard test statistics - like Student, Wald, likelihood ratio and Lagrange multiplier criteria in structural models - have non-standard distributions and often depend heavily on nuisance parameters. Inference procedures robust to weak instruments have been developed. These robust procedures however test hypotheses that are specified on structural parameters. Even though robust procedures solve statistical difficulties related to identification issues, applied researchers may want to first pre-test the exogeneity of some regressors before inference on the parameters of interest. In linear IV regression, Durbin-Wu-Hausman (DWH) tests are often used as pre-tests for exogeneity. Unfortunately, these tests rely on the assumption that model parameters are identified by the available instruments. When identification is deficient or weak, the properties of DWH tests need to be investigated. Early references that study the effects of weak instruments on Hausman-type tests are not well documented and usually focus on testing. Not much is known about pre-test estimators based on DWH tests when IV are weak. In this paper, we provide a large-sample analysis of the distribution of DWH and RH tests under both the null hypothesis (level) and the alternative hypothesis, with or without identification. We show that under the null hypothesis, usual chi-square critical values are applicable irrespective of the presence of weak instruments, in the sense that the asymptotic critical values obtained under the identification assumption provide bounds when identification fails. We characterize a necessary and sufficient condition for DWH and RH tests (with fixed level) to be consistent under the alternative of endogeneity. The latter condition automatically holds when the rank condition for identification holds: DWH tests are consistent when identification holds. The consistency condition also holds in a wide range of cases where identification fails. Moreover, we study the properties of pre-test estimators where OLS or IV is used depending on the outcome of DWH exogeneity tests. We present theoretical arguments suggesting that OLS may be preferable to IV in many cases where regressor endogeneity may be an issue. We present simulation evidence indicating that: (1) over a wide range cases, including weak instruments and moderate endogeneity, OLS performs better than 2SLS [finding similar to Kiviet and Niemczyk (2007)]; (2) pretest-estimators based on exogeneity have an excellent overall performance. Hence, the recommendation of Guggenberger (2008) to abandon the practice of pretesting may go too far. We illustrate our theoretical results through two empirical applications: the relation between trade and economic growth and the widely studied problem of returns to education. We find that exogeneity tests cannot reject the exogeneity of schooling, *i.e.* the IV are possibly weak in this model [Bound (1995)]. However, "trade share" is endogenous, suggesting that the IV are not too poor as showed by Dufour and Taamouti (2007).

**Key words**: Exogeneity tests; weak instruments; pretest-estimators; bias; mean squares errors.

JEL classification: C3; C12; C15; C52.

## 1. Introduction

The literature on weak instruments in linear structural models focuses on proposing statistical procedures which are robust to instrument quality, see Anderson and Rubin (1949, AR-test), Dufour (1997, 2003), Staiger and Stock (1997), Wang and Zivot (1998), Kleibergen (2002, K-test), Moreira (2003, CLR-test), Dufour (2005, 2006), Dufour and Jasiak (2001), Stock, Wright and Yogo (2002), Hall, Rudebusch and Wilcox (1996), Hall and Peixe (2003), Donald and Newey (2001), Doko and Dufour (2008). Weak instrument robust statistics however, test hypotheses that are specified on the parameters of interest. Although robust procedures prevent statistical difficulties related to identification, applied researchers may need to check whether some regressors are exogenous before running inference on the parameters of interest (pretesting). Exogeneity tests of the type proposed by Durbin (1954), Wu (1973), Hausman (1978), Revankar and Hartley (1973) are commonly used for this purpose. Unfortunately, such tests rely on the assumption that model parameters are identified by the available instruments. When identification is weak, the properties (size and power) of exogeneity tests need to be investigated. The literature related to weak instrument problems on exogeneity tests is not well documented. Early references include Guggenberger (2008) and **?**.

Guggenberger (2008) investigates the asymptotic size properties of a two-stage test, where in the first stage a Hausman test is undertaken as a pretest for exogeneity of a regressor. His major finding is that the two-stage test based on DWH-type test have arbitrary size even in large samples. In fact, when the endogeneity between the structural and reduced form errors is local to zero of order $T^{-1/2}$, where $T$ denotes the sample size, the Hausman pretest statistic converges to a noncentral chi-squared distribution. The non centrality parameter is small when the strength of the instruments is small. In this situation, the Hausman pre-test has low power against local deviations of the pretest null hypothesis and consequently, with high probability, OLS-based inference is done in the second stage. However, the second stage OLS based t-statistic often takes on very large values under such local deviations. The latter causes size distortions in the two-stage test. Hahn, Ham and Moon (2008) consider the problem of testing the exogeneity of a subset of excluded instruments. They divide the excluded instruments from the structural equation into two components. The first component is weak but exogenous, while the second is strong but potentially invalid. They then test the validity of the strong component using a modified Hausman-type test. The test statistic proposed is valid despite the presence of the weak component.

However, neither Guggenberger (2008) nor Hahn et al. (2008) provide a formal characterization of DWH-type tests in presence of weak instruments. Furthermore, the issues related to estimation are not addressed by these papers. For example, how do pre-test estimators based on exogeneity tests behave when identification is deficient or weak? In particular, do alternative pre-test estimators based on exogeneity tests better perform (in term of bias and mean square error) than usual IV estimators when instruments are weak?

Doko and Dufour (2010) provide a finite-sample characterization of the distribution of DWH-tests under the null hypothesis (level) and the alternative hypotheses (power). However, the issues related to estimation and the large-sample behaviour of the tests are not addressed.

In this paper, we consider the problem of testing the exogeneity of included regressors in the structural equation. This problem is quite different and more complex than testing orthogonality restrictions of excluded instruments, as done by Hahn et al. (2008). We focus on large-sample and study the behaviour of DWH- and RH-type tests including when identification is deficient or weak (weak instruments). Furthermore, we analyze the properties (bias and mean squares errors) of pre-test estimators based on exogeneity tests.

First, we characterize the asymptotic distribution of DWH and RH tests under the null hypothesis (level) and the alternative hypothesis (power). We show that DWH- and RH-tests are asymptotically robust to weak instruments (level is controlled) and we provide a necessary and sufficient condition under which the tests have no power [similar to Doko and Dufour (2010) and Guggenberger (2008)]. We find that exogeneity tests have no power when all instruments are weak. Moreover, power may exist as soon as we have one strong instrument (partial identification).

Second, we characterize the asymptotic bias and mean square error of OLS, 2SLS and pre-test estimators based on DWH and RH tests. We find that: (1) when identification is deficient or weak (weak instruments) and endogeneity is local to zero [*i.e.* the endogeneity between the structural and reduced form errors converges to zero at rate ($T^{-\frac{1}{2}}$) as the sample size grows], OLS performs (in terms of bias and mean square error) better than 2SLS [finding similar to Kiviet and Niemczyk (2007)]; (2) pretest-estimators based on exogeneity tests have an excellent overall performance compared with OLS and 2SLS estimators. Therefore, the recommendation of Guggenberger (2008) to abandon the practice of pretesting may go too far.

We present two Monte Carlo experiments which confirm our theoretical results. The first examines the properties (size and power) of DWH and RH exogeneity tests. The second studies the bias and mean square error of OLS, 2SLS and pre-test estimators based on exogeneity tests. Our results indicate that: (1) over a wide range cases, including weak instruments and moderate endogeneity, ordinary least squares estimator (OLS) performs better than usual 2SLS estimator; (2) pre-test estimators based on exogeneity tests have an excellent overall performance, hence more preferable than OLS and IV estimators.

We illustrate our theoretical results through two empirical applications: the relation between trade and economic growth [see, Dufour and Taamouti (2006), Irwin and Tervio (2002), Frankel and Romer (1999), Harrison (1996), Mankiw and al. (1992)] and the widely studied problem of returns to education [Dufour and Taamouti (2006), Angrist and Krueger (1991), Angrist and Krueger (1995), Angrist and al. (1999), Mankiw and al. (1992)]. The results indicate that exogeneity tests cannot reject the exogeneity of schooling, which suggest that instruments are possibly weak in this model [Bound, D. and Baker (1995)]. However, “trade share ”is endogenous, *i.e.,* instruments are not too poor as showed in the literature [Dufour and Taamouti (2006)].

The paper is organized as follows. Section 2 formulates the model studied. Section 3 studies the asymptotic behaviour of the tests when identification is strong or deficient (lack of identification). Section 4 examines their behaviour when identification is weak (weak IV). Section 5 presents the pre-test estimators based on exogeneity tests and characterizes their asymptotic behaviour, including when identification is deficient or weak. Section 6 presents two Monte Carlo experiments (i) the properties (size and power) of exogeneity; and (ii) the performance (bias and mean squares errors–MSE) of pre-test estimators. Section 7 illustrates our theoretical results through two important applications. We conclude in Section 8 and proofs are presented in the Appendix.

## 2. Framework

We consider the linear structural model:

$$y = Y\beta + Z_1\gamma + u, \tag{2.1}$$
$$Y = Z_1\Pi_1 + Z_2\Pi_2 + V, \tag{2.2}$$

where $y \in \mathbb{R}^T$ is a dependent variable, $Y \in \mathbb{R}^{T\times G}$ is a matrix of (possibly) endogenous explanatory variables $(G \geq 1)$ $Z_1 \in \mathbb{R}^{T\times k_1}$ is a matrix of exogenous variables, $Z_2 \in \mathbb{R}^{T\times k_2}$ is a matrix of IVs, $u = (u_1, \ldots, u_T)' \in \mathbb{R}^T$ and $V = [v_1, \ldots, v_T]' \in \mathbb{R}^{T\times G}$ are disturbances, $\beta \in \mathbb{R}^G$, $\gamma \in \mathbb{R}^{k_1}$, $\Pi_1 \in \mathbb{R}^{k_1\times G}$ and $\Pi_2 \in \mathbb{R}^{k_2\times G}$ unknown coefficients. Let $Z = [Z_1 : Z_2]$ and $k = k_1 + k_2$. We assume that the "instrument matrix" $Z$ has full-column rank and $k_2 \geq G$. The usual necessary and sufficient condition for identification of this model is $\mathrm{rank}(\Pi_2) = G$. If $\mathrm{rank}(\Pi_2) < G$, $\beta$ is not identified and the instruments are weak. However, some components of $\beta$ may be identified (partial identification) even if this rank condition fails. We also suppose that $u$ can be regressed on $V$ yielding the following equation:

$$u = Va + \varepsilon \tag{2.3}$$

where $a \in \mathbb{R}^G$ is a vector of unknown coefficients, $\varepsilon$ has mean zero, variance $\sigma_\varepsilon^2$ and uncorrelated with $V$.

Let

$$M = M_Z = I - Z(Z'Z)^{-1}Z', \quad Z = [Z_1, Z_2], \quad M_1 = M_{Z_1} = I - Z_1(Z_1'Z_1)^{-1}Z_1'. \tag{2.4}$$

Then, $M_1 - M$ can be expressed as

$$M_1 - M = M_1Z_2(Z_2'M_1Z_2)^{-1}Z_2'M_1 = \bar{Z}_2(\bar{Z}_2'\bar{Z}_2)^{-1}\bar{Z}_2', \tag{2.5}$$

where $\bar{Z}_2 = M_1Z_2 \perp Z_1$. Let $\bar{Z} = [Z_1, \bar{Z}_2]$. If we replace $Z$ by $\bar{Z}$ in (2.27) - (2.29), then the statistics $\mathscr{H}_i$ $(i = 1, 2, 3)$, $\mathscr{T}_l$ $(l = 1, 2, 3, 4)$ and $\mathscr{RH}$ do not change. Therefore, the orthogalization between $Z_1$ and $\bar{Z}_2$ has no impact on our results. To simplify the notations, $\bar{Z}_2$ will be used instead of $Z_2$ [see for example, equation (2.21)].

We make the following generic assumptions on the asymptotic behaviour of model variables [where $B > 0$ for a matrix $B$ means that $B$ is positive definite (p.d.), and $\to$ refers to limits as $T \to \infty$]:

$$\frac{1}{T}\begin{bmatrix} V & \varepsilon \end{bmatrix}'\begin{bmatrix} V & \varepsilon \end{bmatrix} \xrightarrow{p} \begin{bmatrix} \Sigma_V & 0' \\ 0 & \sigma_\varepsilon^2 \end{bmatrix} > 0\,, \tag{2.6}$$

$$\frac{1}{T}Z'\begin{bmatrix} V & \varepsilon \end{bmatrix} \xrightarrow{p} 0\,, \tag{2.7}$$

$$\frac{1}{T}Z'Z \xrightarrow{p} \Sigma_Z = \begin{bmatrix} \Sigma_{Z_1} & \Sigma_{Z_2Z_1}' \\ \Sigma_{Z_2Z_1} & \Sigma_{Z_2} \end{bmatrix} > 0\,, \tag{2.8}$$

$$\frac{1}{\sqrt{T}}V'\varepsilon \xrightarrow{L} S_{V\varepsilon}\,,\, \frac{1}{\sqrt{T}}Z'[u,\, V,\, \varepsilon] \xrightarrow{L} [S_u,\, S_V,\, S_\varepsilon]\,, \tag{2.9}$$

$$\text{vec}[S_u,\, S_V,\, S_\varepsilon,\, S_{V\varepsilon}] \sim \mathrm{N}\,[0\,,\, \Sigma_S]\ ,S_\varepsilon \text{ and } S_V \text{ are uncorrelated,} \tag{2.10}$$

$$S_u = \begin{bmatrix} S_{1u} \\ S_{2u} \end{bmatrix}, \quad S_V = \begin{bmatrix} S_{1V} \\ S_{2V} \end{bmatrix}, \quad S_\varepsilon = \begin{bmatrix} S_{1\varepsilon} \\ S_{2\varepsilon} \end{bmatrix}, \tag{2.11}$$

$$S_{1u} \sim \mathrm{N}\left[0,\, \sigma_u^2 \Sigma_{Z_1}\right], \quad S_{2u} \sim \mathrm{N}\left[0,\, \sigma_u^2 \Sigma_{Z_2}\right], \tag{2.12}$$

$$S_{1\varepsilon} \sim \mathrm{N}\left[0,\, \sigma_\varepsilon^2 \Sigma_{Z_1}\right], \quad S_{2\varepsilon} \sim \mathrm{N}\left[0,\, \sigma_\varepsilon^2 \Sigma_{Z_2}\right], \tag{2.13}$$

$S_{iu}$ is a $k_i \times 1$ random vector, $S_{iV}$ is a $k_i \times G$ random matrix matrix $(i = 1, 2)$, $\Sigma_V$ is $G \times G$ positive definite matrix, and $\sigma_u^2 > 0$.

From the above assumptions, we have

$$\frac{1}{T}Z'u \xrightarrow{p} 0\,, \quad \frac{1}{T}\begin{bmatrix} u & V \end{bmatrix}'\begin{bmatrix} u & V \end{bmatrix} \xrightarrow{p} \Sigma = \begin{bmatrix} \sigma_u^2 & \delta' \\ \delta & \Sigma_V \end{bmatrix} > 0\,, \tag{2.14}$$

where

$$\delta = \Sigma_V a\,,\, \sigma_u^2 = a'\Sigma_V a + \sigma_\varepsilon^2\,,\, S_u = S_V a + S_\varepsilon = S_V(\Sigma_V^{-1}\delta) + S_\varepsilon\,. \tag{2.15}$$

Furthermore,

$$\frac{1}{T}\bar{Z}'\begin{bmatrix} u & V & \varepsilon \end{bmatrix} \xrightarrow{p} 0, \quad \frac{1}{T}\bar{Z}'\bar{Z} \xrightarrow{p} \Sigma_{\bar{Z}} = \begin{bmatrix} \Sigma_{Z_1} & 0 \\ 0 & \Sigma_{\bar{Z}_2} \end{bmatrix} > 0\,, \tag{2.16}$$

$$\frac{1}{\sqrt{T}}\bar{Z}'[u, V, \varepsilon] \xrightarrow{L} [\bar{S}_u,\, \bar{S}_V,\, \bar{S}_\varepsilon]\,, \tag{2.17}$$

$$\text{vec}[\bar{S}_u,\, \bar{S}_V,\, \bar{S}_\varepsilon,\, S_{V\varepsilon}] \sim N\,[0\,,\, \Sigma_{\bar{S}}]\ ,\bar{S}_\varepsilon \text{ and } \bar{S}_V \text{ are uncorrelated,} \tag{2.18}$$

$$\bar{S}_u = \begin{bmatrix} S_{1u} \\ \bar{S}_{2u} \end{bmatrix}, \quad \bar{S}_V = \begin{bmatrix} S_{1V} \\ \bar{S}_{2V} \end{bmatrix}, \quad S_\varepsilon = \begin{bmatrix} S_{1\varepsilon} \\ \bar{S}_{2\varepsilon} \end{bmatrix}, \tag{2.19}$$

$$\bar{S}_{2u} \sim \mathrm{N}\left[0,\, \sigma_u^2 \Sigma_{\bar{Z}_2}\right], \quad \bar{S}_{2\varepsilon} \sim \mathrm{N}\left[0,\, \sigma_\varepsilon^2 \Sigma_{\bar{Z}_2}\right], \tag{2.20}$$

where

$$\Sigma_{\bar{Z}_2} = \Sigma_{Z_2} - \Sigma_{Z_2 Z_1}\Sigma_{Z_1}^{-1}\Sigma_{Z_2 Z_1}'. \tag{2.21}$$

Under assumptions (2.6) - (2.7),

$$\operatorname*{plim}_{T\to\infty} \hat{\beta} = \beta + (\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 + \Sigma_V)^{-1}\delta \tag{2.22}$$

and $\hat{\beta}$ is consistent if and only if $\delta = 0$, irrespective of the rank of $\Pi_2$. In particular, under local alternative considered by Guggenberger (2008) [$\delta = \delta_0/\sqrt{T} \to 0$ as $T \to \infty$], $\hat{\beta}$ is consistent.

However,

$$\tilde{\beta} = \beta + [Y'(M_1 - M)Y]^{-1}Y'(M_1 - M)u = \beta + \left[\frac{Y'(M_1 - M)Y}{T}\right]^{-1}\frac{Y'(M_1 - M)u}{T}\,, \tag{2.23}$$

so, provided that the identification condition $\mathrm{rank}(\Pi_2) = G$ holds,

$$\frac{Y'(M_1 - M)Y}{T} \xrightarrow{p} \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 > 0,\ \frac{Y'(M_1 - M)u}{T} \xrightarrow{p} 0, \tag{2.24}$$

and

$$\underset{T\to\infty}{\mathrm{plim}}\,\tilde{\beta} = \beta\,. \tag{2.25}$$

Nevertheless, $\tilde{\beta}$ does not generally converge to $\beta$ when $\mathrm{rank}(\Pi_2) < G$.

This paper focuses on both testing and estimation. First, we investigate the large-sample properties of DWH and RH exogeneity tests, including when identification is deficient or weak (weak instruments). Second, we study the performance (bias and mean squares errors- MSE) of pre-test estimators based on DWH and RH exogeneity tests, allowing for the presence of weak instruments.

From (2.14) - (2.15), the exogeneity assumption of $Y$ can be expressed as

$$H_0 : \delta = 0 \quad \Leftrightarrow \quad H_a : a = 0\,. \tag{2.26}$$

We consider the Durbin-Wu-Hausman (DWH) test statistics, namely three versions of Hausman-type statistics $[\mathscr{H}_i, i = 1, 2, 3]$, the four statistics proposed by Wu (1973) $[\mathscr{T}_l, l = 1, 2, 3, 4]$ and Revankar and Hartley (1973, RH) test statistic. These statistics are defined by equations (2.27) - (2.29) below:

$$\mathscr{T}_l \;=\; \kappa_l(\tilde{\beta} - \hat{\beta})'\tilde{\Sigma}_l^{-1}(\tilde{\beta} - \hat{\beta})\,, \quad l = 1, 2, 3, 4; \tag{2.27}$$

$$\mathscr{H}_i \;=\; T(\tilde{\beta} - \hat{\beta})'\hat{\Sigma}_i^{-1}(\tilde{\beta} - \hat{\beta})\,, \quad i = 1, 2, 3\,, \tag{2.28}$$

$$\mathscr{R}\mathscr{H} \;=\; \kappa_R y'\hat{\Sigma}_R y\,, \tag{2.29}$$

where $\hat{\beta} = (Y'M_1Y)^{-1}Y'M_1y$ is the ordinary least squares (OLS) estimator of $\beta$, $\tilde{\beta} = [Y'(M_1 - M)Y]^{-1}Y'(M_1 - M)y$ is the two-stage least squares (2SLS) estimator of $\beta$, $\kappa_1 = (k_2 - G)/G$, $\kappa_2 = (T - k_1 - 2G)/G$, $\kappa_3 = \kappa_4 = T - k_1 - G$, $\kappa_R = (T - k_1 - k_2 - G)/k_2$, and

$$\tilde{\Sigma}_1 = \tilde{\sigma}_1^2\hat{\Delta}\,, \quad \tilde{\Sigma}_2 = \tilde{\sigma}_2^2\hat{\Delta}\,, \quad \tilde{\Sigma}_3 = \tilde{\sigma}^2\hat{\Delta}\,, \quad \tilde{\Sigma}_4 = \hat{\sigma}^2\hat{\Delta}\,, \tag{2.30}$$

$$\hat{\Sigma}_1 = \tilde{\sigma}^2\hat{\Omega}_{IV}^{-1} - \hat{\sigma}^2\hat{\Omega}_{LS}^{-1}\,, \quad \hat{\Sigma}_2 = \tilde{\sigma}^2\hat{\Delta}\,, \quad \hat{\Sigma}_3 = \hat{\sigma}^2\hat{\Delta}\,, \tag{2.31}$$

$$\hat{\Sigma}_R = \frac{1}{\hat{\sigma}_R^2}D_1Z_2(Z_2'D_1Z_2)^{-1}Z_2'D_1\,, \tag{2.32}$$

$$\hat{\Omega}_{IV} = \frac{1}{T}Y'(M_1 - M)Y\,, \quad \hat{\Omega}_{LS} = \frac{1}{T}Y'M_1Y\,, \tag{2.33}$$

$$\hat{\Delta} = \hat{\Omega}_{IV}^{-1} - \hat{\Omega}_{LS}^{-1},\ D_1 = \frac{1}{T} M_1 M_{M_1 Y}\,, \tag{2.34}$$
$$\tilde{\sigma}^2 = (y - Y\tilde{\beta})' M_1 (y - Y\tilde{\beta})/T\,, \quad \hat{\sigma}^2 = (y - Y\hat{\beta})' M_1 (y - Y\hat{\beta})/T\,, \tag{2.35}$$
$$\tilde{\sigma}_1^2 = (y - Y\tilde{\beta})'(M_1 - M)(y - Y\tilde{\beta})/T = \tilde{\sigma}^2 - \tilde{\sigma}_e^2\,, \tag{2.36}$$
$$\tilde{\sigma}_2^2 = \hat{\sigma}^2 - (\tilde{\beta} - \hat{\beta})' \hat{\Delta}^{-1} (\tilde{\beta} - \hat{\beta}) = \hat{\sigma}^2 - \tilde{\sigma}^2 (\tilde{\beta} - \hat{\beta})' \hat{\Sigma}_2^{-1} (\tilde{\beta} - \hat{\beta})\,, \tag{2.37}$$
$$\tilde{\sigma}_e^2 = (y - Y\tilde{\beta})' M (y - Y\tilde{\beta})/T,\ \hat{\sigma}_R^2 = y M_{\bar{X}} y'/T\,, \tag{2.38}$$
$$M_{M_1 Y} = I - M_1 Y (Y' M_1 Y)^{-1} Y' M_1\,, \tag{2.39}$$

Note that $\hat{\sigma}^2$ is the OLS-based estimator of $\sigma_u^2$, $\tilde{\sigma}^2$ is the usual 2SLS-based estimator of $\sigma_u^2$ (both without correction for degrees of freedom), while $\tilde{\sigma}_1^2$, $\tilde{\sigma}_2^2$ and $\hat{\sigma}_R^2$ may be interpreted as alternative IV-based scaling factors.

The link between Wu $\mathscr{T}$-tests and Hausman $\mathscr{H}$-tests and the regression formula of these tests has been given in Doko and Dufour (2010). For example, we can observe that $\tilde{\Sigma}_3 = \hat{\Sigma}_2$ and $\tilde{\Sigma}_4 = \hat{\Sigma}_3$, so $\mathscr{T}_3 = (\kappa_3/T)\mathscr{H}_2$ and $\mathscr{T}_4 = (\kappa_4/T)\mathscr{H}_3$. Since $\kappa_3/T = \kappa_4/T \to 1$ as $T \to +\infty$, $\mathscr{T}_3$ is asymptotically equivalent with $\mathscr{H}_2$, and $\mathscr{T}_4$ is asymptotically equivalent with $\mathscr{H}_3$.

Finite-sample distributions for all exogeneity test statistics with possibly weak IVs and non Gaussian errors are available in Doko and Dufour (2010).

We distinguish two setups: (1) $\Pi_2 = \Pi_2^0$ is fixed; and (2) $\Pi_2 = \Pi_2^0/\sqrt{T}$, where $\Pi_2^0 = 0$ is allowed (weak instruments). Section 3 below characterizes the limiting distributions of the statistics under the null hypothesis ($\delta = 0$) and the alternative hypothesis ($\delta \neq 0$) when $\Pi_2$ is fixed (*i.e.* does not depend on the sample size).

## 3. Asymptotic behaviour of exogeneity tests

In this section, we characterize the asymptotic behaviour of the statistics under the null ($\delta = 0$) and the alternative hypotheses ($\delta \neq 0$) when parameters are fixed, so they do not depend on the sample size $T$. We distinguish two cases for the reduced form parameters $\Pi_2$ : (i) $\Pi_2 = \Pi_2^0$, with $\mathrm{rank}(\Pi_2^0) = G$ (strong identification); and (ii) $\Pi_2 = \Pi_2^0$, with $\mathrm{rank}(\Pi_2^0) < G$ (partial identification). To recover partial identification setup, it will be useful to parameterize the model as in Choi and Phillips (1992):

$$\begin{aligned} y &= Y_1\beta_1 + Y_2\beta_2 + Z_1\gamma + u, && (3.1)\\ Y_1 &= Z_1\Pi_{11} + Z_2\Pi_{21} + V_{21}, && (3.2)\\ Y_2 &= Z_1\Pi_{12} + V_{22}, && (3.3) \end{aligned}$$

where

$$\begin{aligned} \Pi_{11} &= \Pi_1\mathscr{S}_1,\ \Pi_{12} = \Pi_1\mathscr{S}_2,\ \Pi_{21} = \Pi_2\mathscr{S}_1, && (3.4)\\ \Pi_{22} &= \Pi_2\mathscr{S}_2 = 0,\ \beta_1 = \mathscr{S}_1'\beta,\ \beta_2 = \mathscr{S}_2'\beta, && (3.5)\\ Y_1 &= Y\mathscr{S}_1,\ Y_2 = Y\mathscr{S}_2,\ V_{21} = V\mathscr{S}_1,\ V_{22} = V\mathscr{S}_2 && (3.6) \end{aligned}$$

and $\mathscr{S} = [\mathscr{S}_1, \mathscr{S}_2] \in \mathscr{O}(G)$, $\mathscr{O}(G)$ denotes the orthogonal group of $G \times G$ matrices, $\mathscr{S}_2 : G \times G_2$ spans the null space of $\Pi_2$, $\mathscr{S}_1 : G \times G_1$, $\beta_1 : G_1 \times 1$ and $\beta_2 : G_2 \times 1$. The necessary and sufficient condition for identification of $\beta_1$ is

$$\operatorname{rank}(\Pi_{21}) = G_1, \tag{3.7}$$

where $\Pi_{21}$ is a $k_2 \times G_1$. This can be seen easily by considering the reduced form for model (3.1)-(3.6)

$$y = Z_1\pi_1 + Z_2\pi_2 + v \tag{3.8}$$

where $\pi_1 = \Pi_{11}\beta_1 + \Pi_{12}\beta_2 + \gamma$, $\pi_2 = \Pi_{21}\beta_1$, and $v = u + V_{21}\beta_1 + V_{12}\beta_2$. So, $\beta_1$ is identified if and only if $\operatorname{rank}(\Pi_{21}) = G_1$. It is important to observe that $\beta_1$ and $\beta_2$ are linear combinations of the original coefficient $\beta$. The original coefficient $\beta$ is recovered by the equation

$$\beta = \mathscr{S}_1\beta_1 + \mathscr{S}_2\beta_2. \tag{3.9}$$

Equation (3.9) can then be used to find the effect of partial identification on the entire vector $\beta$. Of course, if $\operatorname{rank}(\Pi_2) = G$ (strong identification), we have $\mathscr{S}_2\beta_2 = 0$ and $\mathscr{S} = \mathscr{S}_1 = I_G$. Also, if $\operatorname{rank}(\Pi_2) = 0$ (complete non identification), we have $\mathscr{S}_1\beta_1 = 0$ and $\mathscr{S} = \mathscr{S}_2 = I_G$. So, the above parametrization includes strong identification and complete non identification setups as special cases.

We assume that $\beta_1$ is identified but $\beta_2$ may not (partial identification), *i.e.*

$$\operatorname{rank}(\Pi_{21}) = G_1,\ \operatorname{rank}(\Pi_{12}) \leq G_2. \tag{3.10}$$

In particular, if $\operatorname{rank}(\Pi_{12}) = 0$, $\beta_2$ is not identified at all. Note that assumption (3.10) does not constitute a restriction of the model. If assumption (3.10) fails, either the model is identified or absolutely not. Both setups are special cases of (3.1)-(3.6) and will be recovered by our results.

From the above parametrization, the 2SLS estimator of $\beta_1$ and $\beta_2$ are defined by

$$\tilde{\beta}_1 = (Y_1'EY_1)^{-1}Y_1'Ey,\ \tilde{\beta}_2 = (Y_2'JY_2)^{-1}Y_2'Jy, \tag{3.11}$$

where

$$\begin{aligned} E &= M_1 - M - (M_1 - M)Y_2[Y_2'(M_1 - M)Y_2]^{-1}Y_2'(M_1 - M), \\ J &= M_1 - M - (M_1 - M)Y_1[Y_1'(M_1 - M)Y_1]^{-1}Y_1'(M_1 - M). \end{aligned} \tag{3.12}$$

Throughout the paper, the following definitions and notations will be used:

$$\bar{\sigma}_u^2 = \sigma_u^2 + \bar{S}_{2u}'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})\Psi_V^{-1}\Sigma_V\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}, \tag{3.13}$$

$$\Sigma_A \equiv \Sigma_A(\bar{S}_{2V}) = \Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}, \tag{3.14}$$

$$\Sigma_{\bar{Z}_2}^{*} = \Sigma_{\bar{Z}_2} - \Sigma_{\bar{Z}_2}\Pi_2^0(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}, \tag{3.15}$$

$$\Psi_V = (\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V}), \tag{3.16}$$

$$\Delta_V = (\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})', \; \Delta_V^* = I_{k_2} - \Sigma_{\bar{Z}_2}^{-1/2}\Delta_V\Sigma_{\bar{Z}_2}^{-1/2}, \tag{3.17}$$

$$\sigma_{1*}^2 = \sigma_{2*}^2 = \tilde{\sigma}_*^2, \; \sigma_{3*}^2 = \sigma_\varepsilon^2, \; \tilde{\sigma}_{2*}^2 = \tilde{\sigma}_{4*}^2 = \sigma_\varepsilon^2, \; \tilde{\sigma}_{3*}^2 = \tilde{\sigma}_*^2, \tag{3.18}$$

$$\tilde{\sigma}_{1*}^2 = (\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a + \Sigma_{\bar{Z}_2}^{-1/2}S_{2\varepsilon})'\Delta_V^*(\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a + \Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2\varepsilon}), \tag{3.19}$$

$$\begin{aligned}\tilde{\sigma}_*^2 &= \sigma_u^2 - 2\delta'\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u} \\ &+ \bar{S}_{2u}'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})\Psi_V^{-1}\Sigma_V\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u},\end{aligned} \tag{3.20}$$

$$\mu_V = \frac{1}{\sigma_\varepsilon^2}a'\Pi_0'\Delta_V\Pi_0 a = \delta'\Sigma_V^{-1}\Pi_0'\Delta_V\Pi_0\Sigma_V^{-1}\delta, \tag{3.21}$$

$$\lambda_V = \frac{1}{\sigma_\varepsilon^2}a'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1/2}\Delta_V^*\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a = \frac{1}{\sigma_\varepsilon^2}a'\bar{S}_{2V}'(\Sigma_{\bar{Z}_2}^{-1} - \Sigma_{\bar{Z}_2}^{-1}\Delta_V\Sigma_{\bar{Z}_2}^{-1})\bar{S}_{2V}a, \tag{3.22}$$

$$\Sigma_A^0 = \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}(\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V})^{-1}\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}, \tag{3.23}$$

$$\tilde{\sigma}_u^2 = \sigma_u^2 + \mathscr{N}_B'\mathscr{S}_2'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\mathscr{N}_B\mathscr{S}_2, \tag{3.24}$$

$$\sigma_{0*}^2 = \sigma_\varepsilon^2 + \bar{S}_{2\varepsilon}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}(\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V})^{-1}\Sigma_V(\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V})^{-1}\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon} \geq \sigma_\varepsilon^2. \tag{3.25}$$

Finally, for a random variable $\zeta$ whose distribution depends on the sample size $T,$ the notation $\zeta \xrightarrow{L} +\infty$ means that $P[\zeta > x] \to 1$ as $T \to \infty$, for any $x$. We will now characterize the behaviour of the tests under the null hypothesis $H_0$ (section 3.1) and the alternative (section 3.2).

### 3.1. Asymptotic distributions under the null hypothesis

This subsection describes the asymptotic behaviour of DWH and RH tests under the null hypothesis $\delta = 0,$ including when identification is deficient. Theorem **3.1** below shows that all exogeneity tests are valid (level is controlled) even if parameters are not identifiable.

**Theorem 3.1** ASYMPTOTIC DISTRIBUTIONS UNDER $H_0$. *Suppose the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold, and let* $\delta = 0$. *If rank*$(\Pi_2^0) = G,$ *then*

$$\mathscr{H}_i \xrightarrow{L} \chi^2(G), \; i = 1, 2, 3\,, \tag{3.26}$$

$$\mathscr{T}_1 \xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \quad \mathscr{T}_l \xrightarrow{L} \chi^2(G), \; l = 3, 4\,, \tag{3.27}$$

$$\mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2)\,. \tag{3.28}$$

*If rank*$(\Pi_2^0) \leq G,$ *then*

$$\begin{aligned}\mathscr{H}_i &\xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \chi^2(G), \; i = 1, 2, \\ \mathscr{H}_3 &\xrightarrow{L} \chi^2(G),\end{aligned} \tag{3.29}$$

$$
\begin{aligned}
\mathscr{T}_1 &\xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \\
\mathscr{T}_4 &\xrightarrow{L} \chi^2(G),\, \mathscr{T}_3 \xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \chi^2(G),
\end{aligned} \tag{3.30}
$$

$$
\mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2), \tag{3.31}
$$

*where*

$$
\begin{aligned}
\mathscr{N}_B &= \mathscr{S}_2'a + \mathscr{B}^{-1}\mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2\varepsilon} \\
\mathscr{N}_B|_{\bar{S}_{2V}} &\sim N\left[\mathscr{S}_2'a,\, \sigma_\varepsilon^2\mathscr{B}^{-1}\right], \\
\mathscr{B} &= \mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2V}\mathscr{S}_2, \\
\mathscr{A} &= \mathscr{S}_2\mathscr{B}\mathscr{S}_2' = \bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2V},
\end{aligned}
$$

$\tilde{\sigma}_u^2$ *is defined by* (3.24), $a = \Sigma_V^{-1}\delta$ *and* $\mathscr{S}_2$ *is defined in* (3.6) - (3.9).

In the above theorem, since $\delta = 0$ if and only if $a = 0$, we first note that $\mathscr{N}_B|_{\bar{S}_{2V}} \sim \mathrm{N}\left[0,\, \sigma_\varepsilon^2\mathscr{B}^{-1}\right]$. Second, when identification is strong, the asymptotic null distribution of all exogeneity tests is free of nuisance parameters (as expected). When identification fails, the asymptotic null distribution of $\mathscr{T}_1$, $\mathscr{T}_2$, $\mathscr{T}_4$ and $\mathscr{H}_3$ is still pivotal. However, the null distribution of $\mathscr{T}_3$, $\mathscr{H}_1$ and $\mathscr{H}_2$ is asymptotically bounded by a central chi-square with $G$ degrees of freedom. Overall, usual chi-square critical values are applicable irrespective of the presence weak instruments, in the sense that the asymptotic critical values obtained under the identification assumption provide bounds when identification fails [similar to Doko and Dufour (2010)]. We now study the properties of the tests under the alternative hypothesis $\delta \neq 0$.

## 3.2. Asymptotic power

We distinguish two cases for the characterization of the power of the tests. (i) The parameter representing the level of endogeneity $\delta$ is fixed and different from zero; (ii) the endogeneity is local to zero, *i.e,* $\delta$ converges to zero at rate $T^{-\frac{1}{2}}$ as the sample size increases [$\delta = \delta_0/\sqrt{T}$, $\delta_0$ is given]. Theorem **3.2** below presents the results for $\delta$ fixed.

**Theorem 3.2** NECESSARY AND SUFFICIENT CONDITION FOR CONSISTENCY. *Suppose the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold. If* $\Pi_2 = \Pi_2^0$ *is fixed, the necessary and sufficient conditions under which DWH and RH exogeneity tests are consistent is* $\Pi_2^0 a \neq 0$, *where* $a = \Sigma_V^{-1}\delta$. *More precisely,*

$$
\mathscr{H}_i \xrightarrow{L} +\infty,\, \mathscr{T}_l \xrightarrow{L} +\infty, \mathscr{R}\mathscr{H} \xrightarrow{L} +\infty, \tag{3.32}
$$

*for* $i = 1, 2, 3$ *and* $l = 1, 2, 3, 4,$ *if and only if* $\Pi_2^0 a \neq 0$.

Theorem **3.2** above provides the necessary and sufficient condition for consistency of all DWH and RH exogeneity tests when $\Pi_2$ is fixed. The result shows that exogeneity tests can detect an exogeneity problem even if not all model parameters are identified, provided partial identification holds. In particular, we have the following result when model parameters are identified (strong instruments).

**Theorem 3.3** CONSISTENCY OF EXOGENEITY TESTS. *Suppose the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold. If rank*$(\Pi_2^0) = G$*, then all DWH and RH exogeneity tests are consistent.*

Clearly, exogeneity tests may be inconsistent only when identification is deficient. When identification is strong, the tests always detect an endogeneity problem. We can now show the following result concerning the asymptotic behaviour of the tests when $\Pi_2^0 a = 0$.

**Corollary 3.4** ASYMPTOTIC POWER WHEN $\Pi_2 a = 0$. *Suppose the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold and let* $\Pi_2 = \Pi_2^0$ *fixed. If* $\Pi_2 a = 0$, *and rank*$(\Pi_2^0) = G$*, then*

$$\mathscr{H}_i \xrightarrow{L} \chi^2(G),\, i = 1, 2, 3\,, \tag{3.33}$$

$$\mathscr{T}_1 \xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \quad \mathscr{T}_l \xrightarrow{L} \chi^2(G),\, l = 3, 4\,, \tag{3.34}$$

$$\mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2)\,. \tag{3.35}$$

*If* $\Pi_2 a = 0$, *and rank*$(\Pi_2^0) \leq G$*, then*

$$\mathscr{H}_i \xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \chi^2(G),\, i = 1, 2,\, \mathscr{H}_3 \xrightarrow{L} \chi^2(G), \tag{3.36}$$

$$\begin{aligned} \mathscr{T}_1 &\xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \\ \mathscr{T}_4 &\xrightarrow{L} \chi^2(G),\, \mathscr{T}_3 \xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \chi^2(G), \end{aligned} \tag{3.37}$$

$$\mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2), \tag{3.38}$$

*where* $\tilde{\sigma}_u^2$, $\mathscr{N}_B$, $\mathscr{S}_2$ *and* $\mathscr{A}$ *are defined in Theorem* **3.1**.

When rank$(\Pi_2^0) = G$, $\Pi_2^0 a = 0$ if and only if $\delta = 0$. Hence, the null hypothesis is satisfied. Since identification is strong, all DWH and RH statistics are pivotal. However, when rank$(\Pi_2^0) \leq G$, $\Pi_2^0 a = 0$ does not entails that $\delta = 0$. The results of the above corollary indicate that when identification is deficient and $\Pi_2^0 a = 0$, the asymptotic distribution of the statistics is the same under the null hypothesis ($\delta = 0$) and the alternative hypothesis ($\delta \neq 0$). Consequently, exogeneity tests have no asymptotic power in this case.

We now characterize the asymptotic distributions of the statistics tests when the endogeneity is local to zero ($\delta = \delta_0/\sqrt{T}$) and rank$(\Pi_2^0) = G$ (strong identification). The results are presented in the following theorem.

**Theorem 3.5** ASYMPTOTIC POWER WHEN ENDOGENEITY IS LOCAL TO ZERO. *Suppose that the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold, and let* $\delta = \delta_0/\sqrt{T}$. *We have:*

$$\mathscr{H}_i \xrightarrow{L} \chi^2(G, \mu_{\delta_0}), i = 1, 2, 3, \tag{3.39}$$

$$\mathscr{T}_1 \xrightarrow{L} F(G, k_2 - G; \mu_{\delta_0}), \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G, \mu_{\delta_0}), \mathscr{T}_l \xrightarrow{L} \chi^2(G, \mu_{\delta_0}), l = 3, 4, \tag{3.40}$$

$$\mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2, \nu_{\delta_0}), \tag{3.41}$$

*if rank*$(\Pi_2^0) = G$, *where*

$$\begin{aligned} \mu_{\delta_0} &= \frac{1}{\sigma_u^2}\delta_0'(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\Delta_\Pi^{-1}(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\delta_0, \\ \nu_{\delta_0} &= \frac{1}{\sigma_u^2}\delta_0'(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Sigma_{\bar{Z}_2}^{*-1}\Sigma_{\bar{Z}_2}\Pi_2^0(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\delta_0 \end{aligned} \tag{3.42}$$

*and*

$$\mathscr{H}_i \xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \chi^2(G), i = 1, 2, \mathscr{H}_3 \xrightarrow{L} \chi^2(G), \tag{3.43}$$

$$\begin{aligned} \mathscr{T}_1 &\xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \\ \mathscr{T}_4 &\xrightarrow{L} \chi^2(G), \mathscr{T}_3 \xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \chi^2(G), \end{aligned} \tag{3.44}$$

$$\mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2), \tag{3.45}$$

*where* $\tilde{\sigma}_u^2$, $\mathscr{N}_B$, $\mathscr{S}_2$ *and* $\mathscr{A}$ *are defined in Theorem* **3.1**.

First, we note that when identification is strong, all exogeneity tests have non zero power against local alternatives. However, the tests are not consistent whenever $\delta = \delta_0/\sqrt{T} \to 0$, as $T \to +\infty$. If $\delta_0 \neq 0$, the distributions of all exogeneity tests are non central chi-squares, where the non centrality parameters are given in (3.42). Second, when identification is deficient, the distribution of the tests remain the same as when $\delta = 0$. In this case, all tests have no power against local alternatives. So, OLS procedure is used with a high probability in the second stage if one uses a two-stage $t$-test based on a DWH or RH pre-tests. Unlike Guggenberger (2008), we will see in Section 5 that this is a good new in the view point of estimation. In fact, when identification is deficient and endogeneity local to zero, OLS estimator is preferable to 2SLS. Since pre-test estimators behave like OLS in this case, they are also preferable to 2SLS. Clearly, the practice of pre-testing should not be abandoned, as recommended by Guggenberger (2008).

We now focus on weak instruments setup.

# 4. Asymptotic behaviour of exogeneity tests when IV are asymptotically weak

In this section, we focus on weak instruments setup and characterize the behaviour of DWH and RH tests under the null hypothesis ($\delta = 0$) and the alternative hypothesis ($\delta \neq 0$). Weak instruments are characterized as in Staiger and Stock (1997), *i.e.* $\Pi_2 = \Pi_2^0/\sqrt{T}$ where $\Pi_2^0$ is a $k_2 \times G$ constant matrix and $\Pi_2^0 = 0$ is allowed. The subsection 4.1 studies the properties of the tests under the null hypothesis.

## 4.1. Asymptotic distributions under the null hypothesis

Following Staiger and Stock (1997), weak instruments are characterized by the local to zero condition for the reduced form matrix $\Pi_2$:

$$\Pi_2 = \Pi_2^0/\sqrt{T}, \tag{4.1}$$

where $\Pi_2^0$ is a $k_2 \times G$ constant matrix and $\Pi_2^0 = 0$ is allowed. Theorem **4.1** below shows that all exogeneity are valid when instruments are weak.

**Theorem 4.1** ASYMPTOTIC DISTRIBUTION UNDER $H_0$ WHEN ASYMPTOTICALLY IV ARE WEAK . *Suppose that the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold. If* $\Pi_2 = \Pi_2^0/\sqrt{T}$ ($\Pi_2^0 = 0$ *is allowed*)*, then under the null hypothesis* $\delta = 0$*, all DWH and RH tests are valid (level is controlled). In particular, we have*

$$\mathscr{H}_i \xrightarrow{L} \frac{1}{\bar{\sigma}_u^2}\bar{S}_{2u}'\Sigma_A\bar{S}_{2u} \leq \chi^2(G),\, i = 1,\, 2,\, \mathscr{H}_3 \xrightarrow{L} \chi^2(G), \tag{4.2}$$

$$\mathscr{T}_1 \xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \quad \mathscr{T}_4 \xrightarrow{L} \chi^2(G),\, \mathscr{T}_3 \xrightarrow{L} \frac{1}{\bar{\sigma}_u^2}\bar{S}_{2u}'\Sigma_A\bar{S}_{2u} \leq \chi^2(G), \tag{4.3}$$

$$\mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2), \tag{4.4}$$

$\bar{S}_{2u}$ *is defined in* (2.17) - (2.20), $\bar{\sigma}_u^2$ *and* $\Sigma_A$ *are defined in* (3.13) - (3.25).

We observe that when identification is weak (weak IVs), the statistics $\mathscr{T}_1$, $\mathscr{T}_2$, $\mathscr{T}_4$ and $\mathscr{H}_3$ are asymptotically pivotal under the null hypothesis ($\delta = 0$). However, the asymptotic distributions of $\mathscr{T}_3,\ \mathscr{H}_1$ and $\mathscr{H}_2$ depend on model parameters, but are bounded by a central chi square with $G$ degrees of freedom. Hence, $\mathscr{T}_3,\ \mathscr{H}_1$ and $\mathscr{H}_2$ are conservative [similar to Doko and Dufour (2010)]. We now study the properties of the tests under the alternative hypothesis $\delta \neq 0$.

### 4.2. Asymptotic power

We will now examine the properties of exogeneity tests under the alternative hypothesis $\delta \neq 0$. The following theorem presents the results.

**Theorem 4.2** ASYMPTOTIC POWER WHEN INSTRUMENTS ARE ASYMPTOTICALLY WEAK. *Suppose that the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold. If* $\Pi_2 = \Pi_2^0/\sqrt{T}$ $(\Pi_2^0 = 0$ *is allowed*$)$, *then, for* $i = 1, 2, 3$ *and* $l = 1, 2, 3, 4,$ *we have*

$$\mathscr{H}_i \xrightarrow{L} \frac{1}{\sigma_{i*}^2}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon}),\ i = 1, 2, 3, \tag{4.5}$$

$$\mathscr{T}_l \xrightarrow{L} \frac{\bar{\kappa}_l}{\tilde{\sigma}_{l*}^2}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon}),\ l = 1, 2, 3, 4, \tag{4.6}$$

$$\mathscr{R}\mathscr{H} \quad \xrightarrow{L} \quad \frac{1}{k_2\sigma_\varepsilon^2}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0 a)'\Sigma_{\bar{Z}_2}^{-1}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0 a) \sim \frac{1}{k_2}\chi^2(k_2, \mu_R), \tag{4.7}$$

*where* $\mu_R = a'\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0 a,\ \bar{\kappa}_1 = (k_2 - G)/G,\ \bar{\kappa}_2 = 1/G$ *and* $\bar{\kappa}_3 = \bar{\kappa}_4 = 1.$ *Furthermore, we have,*

$$\mathscr{H}_i|\bar{S}_{2V} \xrightarrow{L} \frac{1}{\sigma_{i*}^2}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})|\bar{S}_{2V} \leq \chi^2(G, \mu_V),\ i = 1, 2, \tag{4.8}$$

$$\mathscr{H}_3|\bar{S}_{2V} \xrightarrow{L} \chi^2(G, \mu_V),\ \mathscr{T}_1|\bar{S}_{2V} \xrightarrow{L} F(G, k_2 - G; \mu_V, \lambda_V), \tag{4.9}$$

$$\mathscr{T}_2|\bar{S}_{2V} \xrightarrow{L} \frac{1}{G}\chi^2(G, \mu_V), \quad \mathscr{T}_4|\bar{S}_{2V} \xrightarrow{L} \chi^2(G, \mu_V), \tag{4.10}$$

$$\mathscr{T}_3|\bar{S}_{2V} \xrightarrow{L} \frac{\bar{\kappa}_3}{\tilde{\sigma}_{3*}^2}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})|\bar{S}_{2V} \leq \chi^2(G, \mu_V), \tag{4.11}$$

$\bar{S}_{2\varepsilon},\ \bar{S}_{2V}$ *are defined in* (2.17) - (2.20), $\sigma_{i*}^2,\ i = 1, 2, 3\ \ \tilde{\sigma}_{l*}^2,\ l = 1, 2, 3, 4,$ *and* $\Delta_V,\ \mu_V,\ \lambda_V,$ *are defined in* (3.13) - (3.25).

From the above theorem, we note that when identification is weak, exogeneity tests do not converge under the alternative hypothesis $\delta \neq 0$. The asymptotic distribution of the statistics converge to finite non-degenerate distributions. Furthermore, the conditional limiting distributions of $\mathscr{H}_3$, $\mathscr{T}_2$, $\mathscr{T}_4$ and $\mathscr{R}\mathscr{H}$ given $\bar{S}_{2V}$ are noncentral chi-square distributions while $\mathscr{T}_1$ follows a double non-central $F$-distribution. However, $\mathscr{H}_1,\ \mathscr{H}_2$, and $\mathscr{T}_3$ are bounded upward by a non central chi square distribution with $G$ degrees of freedom and non centrality parameter $\mu_V$. This suggests that exogeneity tests can have non zero power even in presence of weak identification, provided the non central parameters in the above theorem are different from zero. So, we can then characterize in

Theorem **4.3** below, the necessary and sufficient condition under which exogeneity tests have no power when identification is weak.

**Theorem 4.3** NECESSARY AND SUFFICIENT CONDITIONS FOR NO POWER. *Under the assumption of Theorem* **4.2**, *the power of DWH and RH tests does not exceed the nominal levels if and only if* $\Pi_2^0 a = 0$. *More precisely, we have under* $\Pi_2^0 a = 0$

$$\mathscr{H}_i \xrightarrow{L} \frac{1}{\sigma_{0*}^2} \bar{S}_{2\varepsilon}' \Sigma_A^0 \bar{S}_{2\varepsilon} \leq \chi^2(G),\, i = 1, 2,\; \mathscr{H}_3 \xrightarrow{L} \chi^2(G)\,, \tag{4.12}$$

$$\mathscr{T}_1 \xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \quad \mathscr{T}_4 \xrightarrow{L} \chi^2(G), \tag{4.13}$$

$$\mathscr{T}_3 \xrightarrow{L} \frac{1}{\sigma_{0*}^2} \bar{S}_{2\varepsilon}' \Sigma_A^0 \bar{S}_{2\varepsilon} \leq \chi^2(G),\, \mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2), \tag{4.14}$$

*where* $\sigma_{0*}^2,\ \Sigma_A^0$ *are defined in* (3.13) - (3.25) *and* $\bar{S}_{2\varepsilon}$ *in* (2.17) - (2.20).

Observe that when $\Pi_2^0 a = 0,$ the non centrality parameters in Theorem **4.2** vanish so that the statistics $\mathscr{H}_3,\ \mathscr{T}_2,\ \mathscr{T}_4$ and $\mathscr{R}\mathscr{H}$ have central chi-square limiting distributions while $\mathscr{T}_1$ is asymptotically distributed as a Fisher with $(k_2 - G, G)$ degrees of freedom. Furthermore, $\mathscr{H}_1,\ \mathscr{H}_2$ and $\mathscr{T}_3$ are bounded by a central chi-square distribution with $G$ degrees of freedom. Therefore, the asymptotic power of $\mathscr{H}_3,\ \mathscr{T}_2,\ \mathscr{T}_4,\ \mathscr{T}_1$ and $\mathscr{R}\mathscr{H}$ equals the nominal levels while those of $\mathscr{H}_1,\ \mathscr{H}_2$ and $\mathscr{T}_3$ cannot exceed the nominal level [similar to Doko and Dufour (2010)]. Section 5 below studies the asymptotic behaviour of pre-test estimators based on DWH and RH tests.

## 5. Pre-test estimators based on exogeneity tests

An important and pratical problem in econometrics consists in using DWH-and RH-type tests to pretest the exogeneity of regressors to decide whether one should apply ordinary least squares or instrumental variables methods for satistical inference. Although this practice seems to be prevalent in applied research, some authors, including Guggenberger (2008), have shown that the two-stage $t$-test procedure based on DWH-and RH-tests is unreliable from the viewpoint of size control when endogeneity is local to zero of order $T^{-1/2}$ and the instruments are weak. In both cases, exogeneity tests are inconsistent and the two-stage $t$-test procedure may be arbitrary size distorted. This is showed by Guggenberger (2008), using some configurations of model parameters. Guggenberger (2008) suggests to use a 2SLS based $t$-test when instruments are strong and the identification-robust procedures [Anderson and Rubin (1949, AR-test), Kleibergen (2002, K-test), Moreira (2003, CLR-test), projection-based techniques, see Dufour (1997, 2003), Dufour (2005, 2006), split-sample methods, see Dufour and Jasiak (2001)] when there are weak. This suggests that the practice of pretesting of the regressors should be abandoned. However, it is not clear how behave the pre-test estimators when instruments are weak. The framework of Guggenberger (2008) focuses in testing and does not deal with estimation. The main objective of this section is to study the behaviour of pre-test estimators based on exogeneity tests, including when identification is deficient or weak (weak instruments).

We consider eight pre-test estimators associated to DWH and RH pre-tests defined by equations (5.1) - (5.3) below:

$$\begin{aligned}
\hat{\beta}_{Hi} &= \hat{\beta}\mathbb{1}[\mathscr{H}_i \le c_{\mathscr{H}_i,1-\xi}] + \tilde{\beta}\mathbb{1}[\mathscr{H}_i > c_{\mathscr{H}_i,1-\xi}], i = 1, 2, 3, && (5.1)\\
\hat{\beta}_{Tl} &= \hat{\beta}\mathbb{1}[\mathscr{T}_l \le c_{\mathscr{T}_l,1-\xi}] + \tilde{\beta}\mathbb{1}[\mathscr{T}_i > c_{\mathscr{T}_l,1-\xi}], l = 1, 2, 3, 4, && (5.2)\\
\hat{\beta}_{RH} &= \hat{\beta}\mathbb{1}[\mathscr{RH} \le c_{\mathscr{RH},1-\xi}] + \tilde{\beta}\mathbb{1}[\mathscr{RH} > c_{\mathscr{RH},1-\xi}], && (5.3)\\
& && (5.4)
\end{aligned}$$

where $\hat{\beta}$ and $\tilde{\beta}$ are the OLS and 2SLS estimators of $\beta$, $\mathbb{1}[.]$ is the indicator function and $c_{\mathscr{H}_i,1-\xi}$, $i = 1, 2, 3$, $c_{\mathscr{T}_l,1-\xi}$, $l = 1, 2, 3, 4$, and $c_{\mathscr{RH},1-\xi}$ are the usual $1-\xi$ quantile of the standard distributions of DWH and RH statistics respectively. It is important to observe that the pre-test estimators defined by (5.1)-(5.3) are convex combinations of OLS and 2SLS estimators. The weight allocated to each estimator is determined by the outcome of the underline pre-test in the first stage.

Lemma **5.1** below characterizes the probability limit of OLS and 2SLS estimators when $\Pi_2$ is fixed.

**Lemma 5.1** LIMIT VALUES OF OLS AND 2SLS ESTIMATORS. *Suppose the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold. If $\Pi_2$ is fixed, then*

$$\begin{aligned}
\underset{T\to\infty}{\text{plim}}\,\hat{\beta} &= \bar{\beta}_{LS} = \beta + (\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 + \Sigma_V)^{-1}\delta, && (5.5)\\
\underset{T\to\infty}{\text{plim}}\,\tilde{\beta}_1 &= \beta_1, \underset{T\to\infty}{\text{plim}}\,\tilde{\beta}_2 = \beta_2 + \mathscr{N}_B, && (5.6)\\
\underset{T\to\infty}{\text{plim}}\,\tilde{\beta} &= \bar{\beta}_{IV} = \beta + \mathscr{S}_2\mathscr{N}_B && (5.7)
\end{aligned}$$

*where*

$$\mathscr{N}_B = \begin{cases} 0 \text{ if } rank(\Pi_2) = G, \\ \\ \mathscr{S}_2'a + \mathscr{B}^{-1}\mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2\varepsilon} \text{ if } rank(\Pi_2) \le G\,, \end{cases} \tag{5.8}$$

$\mathscr{N}_B|_{\bar{S}_{2V}} \sim N\left[\mathscr{S}_2'a,\, \sigma_\varepsilon^2\mathscr{B}^{-1}\right]$, $\mathscr{B} = \mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2V}\mathscr{S}_2$, *$\mathscr{S}_2$ is defined in* (3.6) - (3.9) *and $a = \Sigma_V^{-1}\delta$.*

We make the following observations concerning Lemma **5.1**.

(i) From (5.7)-(5.8), we have

$$\underset{T\to\infty}{\text{plim}}\,\tilde{\beta} = \beta + a + \mathscr{S}_2\mathscr{N}_B^0 = \beta^* + \mathscr{S}_2\mathscr{N}_B^0 \tag{5.9}$$

where $\beta^* = \beta + a$ and $\quad \mathcal{N}_B^0 = \begin{cases} 0 \text{ if } \operatorname{rank}(\Pi_2) = G, \\ \\ \mathscr{B}^{-1}\mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2\varepsilon} \text{ if } \operatorname{rank}(\Pi_2) \leq G. \end{cases}$

Furthermore, by using the generalization of matrix inversion lemma [see Tylavsky and Sohie (1986, Equation (1d))], we have

$$(\Sigma_V + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2)^{-1} \;=\; \Sigma_V^{-1} - \Sigma_V^{-1}(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1} \tag{5.10}$$

so that

$$\begin{aligned}(\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 + \Sigma_V)^{-1}\delta &= \Sigma_V^{-1}\delta - \Sigma_V^{-1}(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1}\delta \\ &= a - \Sigma_V^{-1}(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 a.\end{aligned}$$

Thus (5.5) can be written as

$$\operatorname*{plim}_{T\to\infty}\hat{\beta} \;=\; \beta^* - \Sigma_V^{-1}(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 a. \tag{5.11}$$

(ii) If $\Pi_2 a = 0$, we have

$$\operatorname*{plim}_{T\to\infty}\hat{\beta} \;=\; \beta^* = \beta + a,\ \operatorname*{plim}_{T\to\infty}\tilde{\beta} = \beta^* + \mathscr{S}_2\mathcal{N}_B^0, \tag{5.12}$$

$$AMSE(\hat{\beta}) \;=\; \operatorname*{plim}_{T\to\infty}[MSE(\hat{\beta})] = \|a\| = \delta'\Sigma_V^{-2}\delta, \tag{5.13}$$

$$\begin{aligned}AMSE(\hat{\beta}) &= \operatorname*{plim}_{T\to\infty}[MSE(\tilde{\beta})] = \|a + \mathscr{S}_2\mathcal{N}_B^0\| = \|a\| + \|\mathscr{S}_2\mathcal{N}_B^0\| + 2a'\mathscr{S}_2\mathcal{N}_B^0 \\ &= \delta'\Sigma_V^{-2}\delta + \mathcal{N}_B^{0'}\mathscr{S}_2'\mathscr{S}_2\mathcal{N}_B^0 + 2a'\mathscr{S}_2\mathcal{N}_B^0,\end{aligned} \tag{5.14}$$

where $AMSE(\hat{\theta})$ is the asymptotic mean square error of $\hat{\theta} \in \{\hat{\beta}, \tilde{\beta}\}$. Hence OLS is always consistent under the hypothesis of exogeneity ($\delta = 0$), but 2SLS may not provided identification is deficient [$\operatorname{rank}(\Pi_2) < G$].

Suppose that $\operatorname{rank}(\Pi_2) = G$. Then, we have $\Pi_2 a = 0$ if and only if $a = 0$. By using (5.12), we get

$$\operatorname*{plim}_{T\to\infty}\hat{\beta} \;=\; \beta,\ \operatorname*{plim}_{T\to\infty}\tilde{\beta} = \beta, \tag{5.15}$$

$$AMSE(\hat{\beta}) \;=\; AMSE(\tilde{\beta}) = 0. \tag{5.16}$$

Both OLS and 2SLS estimators are consistent if strong is identification (as expected).

Suppose now that $\operatorname{rank}(\Pi_2) < G$ (*i.e.* identification is deficient or weak). Since $\Pi_2 a = 0 \not\Rightarrow a = 0$, if endogeneity is present ($a \neq 0$), OLS converges to a pseudo value $\beta^* = \beta + a$ while 2SLS converges to $\beta^*$ plus a non degenerate random variable. More interestingly, the pseudo value $\beta^*$ is *observationally equivalent* to the true value $\beta$. To see this latter point, consider equations (2.1)-(2.3). From (2.2) and (2.3), we have $V = Y - Z_1\Pi_1 - Z_2\Pi_2$ and $u = Va + \varepsilon$. Substituting these expressions in (2.1) gives

$$y = Y\beta^* + Z_2\Pi_2 a + Z_1\gamma^* + \varepsilon, \tag{5.17}$$

where $\beta^* = \beta + a$ and $\gamma^* = \gamma + \Pi_1 a$. If $\Pi_2 a = 0$, ( 5.17) becomes

$$y = Y\beta^* + Z_1\gamma^* + \varepsilon, \tag{5.18}$$

and $\hat{\beta}^* = \hat{\beta}$. Clearly, the pseudo value $\beta^*$ is observationally equivalent to the true value $\beta$. This means that when identification fails, unlike 2SLS estimator, the inconsistency of OLS estimator is not too problematic as one should think. Now, define

$$AMSE_{OLS}(\beta^*) = \operatorname*{plim}_{T\to\infty}\|\hat{\beta} - \beta^*\| \quad \text{and} \quad AMSE_{IV}(\beta^*) = \operatorname*{plim}_{T\to\infty}\|\tilde{\beta} - \beta^*\|. \tag{5.19}$$

If $\Pi_2 a = 0$, then we have

$$AMSE_{OLS}(\beta^*) = 0,\ AMSE_{IV}(\beta^*) = \|\mathscr{S}_2\mathscr{N}_B^0\| > 0. \tag{5.20}$$

Hence, OLS is preferable to 2SLS if identification is deficient. Of course, (5.17)-(5.20) remain valid if $\Pi_2 = 0$ (complete non identification of $\beta$).

(iii) If $\Pi_2 a \neq 0$, then both OLS and 2SLS estimators are biased and their respective asymptotic biases and mean square errors (centered at $\beta^*$) are given by

$$\begin{aligned} \operatorname*{plim}_{T\to\infty}(\hat{\beta} - \beta^*) &= -\Sigma_V^{-1}(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 a, \\ AMSE_{OLS}(\beta^*) &= a'\mathscr{C}a, \end{aligned} \tag{5.21}$$

where $\mathscr{C} = \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Sigma_V^{-2}(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2$ and

$$\operatorname*{plim}_{T\to\infty}(\tilde{\beta} - \beta^*) = \mathscr{S}_2\mathscr{N}_B^0,\ AMSE_{OLS}(\beta^*) = \|\mathscr{S}_2\mathscr{N}_B^0\|. \tag{5.22}$$

So, unlike 2SLS, we observe that the asymptotic bias and mean square error of OLS, centered at $\beta^*$, depend on the degree of endogeneity $a$. Furthermore, since $\mathscr{C} \geq 0$, $AMSE_{OLS}(\beta^*)$ is a nondecreasing and unbounded function of $a$. This suggests that if $\Pi_2 a \neq 0$ and endogeneity is large, 2SLS is preferable to OLS.

(iv) Finally, we note that $\tilde{\beta}_1$ is still consistent even if identification is deficient or weak, while $\tilde{\beta}_2$ is consistent only when IV are strong. Hence, the inconsistency of $\tilde{\beta}$ comes from $\tilde{\beta}_2$.

We can state a similar lemma concerning the behaviour of OLS and 2SLS estimators when instruments are asymptotically weak [$\Pi_2 = \Pi_2^0/\sqrt{T}$]. The results are presented in Lemma **5.2** below.

**Lemma 5.2** LIMIT VALUES OF OLS AND 2SLS ESTIMATORS WHEN INSTRUMENTS ARE

ASYMPTOTICALLY WEAK. *Suppose that the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold. If* $\Pi_2 = \Pi_2^0/\sqrt{T}$, *where* $\Pi_2^0$ *is a* $k_2 \times G$ *constant matrix* ($\Pi_2^0 = 0$ *is allowed*), *then*

$$\operatorname*{plim}_{T\to\infty} \hat{\beta} = \beta^*, \tag{5.23}$$

$$\operatorname*{plim}_{T\to\infty} \tilde{\beta} = \beta^* + \mathscr{N}_{\Psi}^{W}, \tag{5.24}$$

*where* $\Psi_V = (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V}')\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V}')$, $\mathscr{N}_{\Psi}^{W} = \Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0 a)$, $\mathscr{N}_{\Psi}^{W}|_{\bar{S}_{2V}} \sim N[-\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\Sigma_{\bar{Z}_2}\Pi_2^0 a,\ \sigma_\varepsilon^2\Psi_V^{-1}]$ *and* $\beta^* = \beta + a$.

So, we see that the observations in Lemma **5.1**-(ii) still hold.

We can now prove the above results on the behaviour of pre-test estimators defined in (5.1)-(5.3).

**Theorem 5.3** LIMIT VALUES OF PRE-TEST ESTIMATORS. *Suppose the assumptions* (2.1) - (2.3) *and* (2.6) - (2.13) *hold. If* $\Pi_2$ *is fixed, then*

$$\begin{aligned}\operatorname*{plim}_{T\to\infty}(\hat{\beta}_{\mathscr{W}} - \beta^*) &= -p_{\mathscr{W}}\Sigma_V^{-1}(I + \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2\Sigma_V^{-1})^{-1}\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 a + \\ &\quad (1 - p_{\mathscr{W}})\mathscr{S}_2\mathscr{N}_B^0,\end{aligned} \tag{5.25}$$

*where* $\mathscr{S}_2\mathscr{N}_B^0$ *is defined by* (5.8). *If* $\Pi_2 = \Pi_2^0/\sqrt{T}$, *then*

$$\operatorname*{plim}_{T\to\infty}(\hat{\beta}_{\mathscr{W}} - \beta^*) = (1 - p_{\mathscr{W}})\mathscr{N}_{\Psi}^{W}, \tag{5.26}$$

*where* $\mathscr{N}_{\Psi}^{W}$ *is defined in Lemma* **5.2** *and*

$$p_{\mathscr{W}} = \lim_{T\to\infty} P[\mathscr{W} \le c_{\mathscr{W},1-\xi}] \tag{5.27}$$

*and* $\mathscr{W} \in \{Hi, Tl, RH\}$, $i = 1, 2, 3$, $l = 1, 2, 3, 4$.

We make the following remarks:

(i) when $\Pi_2$ is fixed, if further $\Pi_2 a = 0$, we have

$$\operatorname*{plim}_{T\to\infty}(\hat{\beta}_{\mathscr{W}} - \beta^*) = (1 - p_{\mathscr{W}})\mathscr{S}_2\mathscr{N}_B^0 \le \mathscr{S}_2\mathscr{N}_B^0, \tag{5.28}$$

$$AMSE_{\mathscr{W}}(\beta^*) = (1 - p_{\mathscr{W}})^2 AMSE_{IV}(\beta^*) \le AMSE_{IV}(\beta^*). \tag{5.29}$$

In particular, when identification is deficient, the two-stage estimator is preferable to 2SLS. If $\Pi_2 a \neq 0$, we have $p_{\mathscr{W}} = 0$ (consistency of DWH and RH tests) and

$$\operatorname*{plim}_{T\to\infty}(\hat{\beta}_{\mathscr{W}} - \beta^*) = \mathscr{S}_2\mathscr{N}_B^0, \tag{5.30}$$

$$AMSE_{\mathscr{W}}(\beta^*) = AMSE_{IV}(\beta^*). \tag{5.31}$$

So, pre-test estimators based on exogeneity tests behave like 2SLS. Since 2SLS is preferable to OLS when $\Pi_2 a \neq 0$ and endogeneity is large, pre-test estimators estimators are also preferable to OLS in this cases;

(ii) if $\Pi_2 = \Pi_2^0/\sqrt{T}$ (instruments are asymptotically weak), the results are similar to $\Pi_2 a = 0$. Thus, pre-test estimators based on exogeneity tests are preferable to 2SLS.

Overall, pre-test estimators based on exogeneity have an excellent performance compared to OLS and 2SLS estimators.

Section 6 below presents the Monte Carlo experiment.

# 6. Monte Carlo experiment

In this section, we perform two Monte Carlo experiments. The first experiment study the effects of weak IVs on DWH and RH tests. In this experiment, we consider three setup: (1) Strong identification of model parameters; (2) partial identification; and (3) weak identification. The second experiment analyzes the performance (bias and MSE) of the pre-test estimators based on DWH and RH exogeneity tests. The framework of this experiment is similar to Guggenberger (2008).

## 6.1. Size and power of DWH and RH tests

Consider the two endogenous variables model described by the following data generating process:

$$y = Y_1\beta_1 + Y_2\beta_2 + u, \quad (Y_1, Y_2) = (Z_2\Pi_{21}, Z_2\Pi_{22}) + (V_1, V_2), \tag{6.32}$$

where $Z_2$ is a $T \times k_2$ matrix of instruments such that $Z_{2t}$ follow $i.i.d$ $N(0, I_{k_2})$ for $t = 1, \ldots, T$, $\Pi_{21}$ and $\Pi_{22}$ are vectors of dimension $k_2$. We assume that

$$u = Va + \varepsilon = V_1 a_1 + V_2 a_2 + \varepsilon, \tag{6.33}$$

where $a_1$ and $a_2$ are $2 \times 1$ vectors and $\varepsilon$ is independent with $V = (V_1, V_2)$, $V_1$ and $V_2$ are $T \times 1$ vectors. Through this experiment, $V$ and $\varepsilon$ are drawn as

$$(V_{1t}, V_{2t})' \overset{i.i.d}{\sim} N\left(0, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \quad \text{and} \quad \varepsilon_t \overset{i.i.d}{\sim} N(0, 1), \quad \text{for all} \quad t = 1, \ldots, T. \tag{6.34}$$

The above setup allows us to take into account situations where $\beta = (\beta_1, \beta_2)'$ is partially identified. In particular, if $\Pi_{21} = 0$ and $\det(\Pi_{22}'\Pi_{22}) \neq 0$, the instruments $Z_2$ cannot identify $\beta_1$. However, $\beta_2$ is identified. We define

$$\Pi_{21} = \eta_1 C_0, \; \Pi_{22} = \eta_2 C_1, \tag{6.35}$$

where $\eta_1$ and $\eta_2$ take the value 0 (design of complete non identification), .01 (design of weak identification) or .5 (design of strong identification), $[C_0, C_1]$ is a $k_2 \times 2$ matrix obtained by taking the first two columns of the identity matrix of order $k_2$. The number of instruments $k_2$ varies in $\{5, 10, 20\}$ and the true value of $\beta$ is set at $\beta_0 = (2, 5)'$. It is worthwhile to note that when $\eta_1$ and $\eta_2$

belong to $\{0, .01\}$, the instruments $Z_2$ are weak and both ordinary least squares and two stage least squares estimators of $\beta$ in (6.32) are biased and inconsistent unless $a_1 = a_2 = 0$. The simulations are run the sample $T = 500$, and the number of replications is $N = 10,000$. The endogeneity $a$ is chosen such that

$$a = (a_1, a_2)' \in \{(-20, 0)', (-5, 5)', (0, 0)', (.5, .2)', (100, 100)'\}. \tag{6.36}$$

From the above setup, the exogeneity hypothesis for $Y$ is expressed as

$$H_0 : a = (a_1, a_2)' = (0, 0)'. \tag{6.37}$$

The nominal level of the tests is 5%. For each value of the vector $a$, we compute the empirical rejection probability of exogeneity test statistics. When $a = 0$, the rejection frequencies are the empirical levels of the tests. However, if $a \neq 0$, the rejection frequencies represent the power of the tests.

The results are presented in Table 1 below. In the first column of the table, we report the statistics while in the second column, we report the values of $k_2$ (number of excluded instruments). Finally in the other columns, we report for each value of the endogeneity $a$ and instrument qualities $\eta_1$ and $\eta_2$, the rejection frequencies at nominal level 5%.

First, we note that all exogeneity tests are valid whether the instruments are strong or weak. In particular, $\mathscr{T}_1$, $\mathscr{T}_2$, $\mathscr{T}_4$, $\mathscr{H}_3$ and $RH$ control the level while $\mathscr{T}_3$, $\mathscr{H}_1$ and $\mathscr{H}_2$ are conservative when IVs are weak. However, $\mathscr{T}_3$, $H_1$ and $\mathscr{H}_2$ do not exhibit this problem when identification is strong [see the column $(a_1, a_2)' = (0, 0)'$ in Table 1 below].

Second, all exogeneity tests have a low power when both $\beta_1$ and $\beta_2$ are not identified even in large-sample Nevertheless, when at least one component of $\beta$ is identified [Table 1 (continued)], all exogeneity tests exhibit power.

Table 1. Power of exogeneity tests at nominal level 5%; $G = 2$, $T = 500$

| | $k_2$ | $(a_1,a_2)' = (-20,0)'$ | | | $(a_1,a_2)' = (-5,5)'$ | | | $(a_1,a_2)' = (0,0)'$ | | | $(a_1,a_2)' = (.5,.2)'$ | | | $(a_1,a_2)' = (100,100)'$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\eta_1 = 0$ $\eta_2 = 0$ | $\eta_1 = .01$ $\eta_2 = 0$ | $\eta_1 = .5$ $\eta_2 = 0$ | $\eta_1 = 0$ $\eta_2 = 0$ | $\eta_1 = .01$ $\eta_2 = 0$ | $\eta_1 = .5$ $\eta_2 = 0$ | $\eta_1 = 0$ $\eta_2 = 0$ | $\eta_1 = .01$ $\eta_2 = 0$ | $\eta_1 = .5$ $\eta_2 = 0$ | $\eta_1 = 0$ $\eta_2 = 0$ | $\eta_1 = .01$ $\eta_2 = 0$ | $\eta_1 = .5$ $\eta_2 = 0$ | $\eta_1 = 0$ $\eta_2 = 0$ | $\eta_1 = .01$ $\eta_2 = 0$ | $\eta_1 = .5$ $\eta_2 = 0$ |
| $\mathscr{T}_1$ | 5 | 5.24 | 6.23 | 100 | 5.12 | 5.35 | 100 | 5.06 | 4.76 | 4.73 | 4.8 | 4.98 | 94.91 | 4.91 | 5.96 | 100 |
| $\mathscr{T}_2$ | 5 | 4.66 | 91.92 | 100 | 5.11 | 27.86 | 100 | 5.11 | 4.91 | 4.43 | 5.35 | 5.09 | 100 | 4.92 | 98.13 | 100 |
| $\mathscr{T}_3$ | 5 | 0.02 | 13.61 | 100 | 0.04 | 0.51 | 99.98 | 0 | 0 | 0.99 | 0.02 | 0.03 | 99.45 | 0.01 | 19.26 | 99.99 |
| $\mathscr{T}_4$ | 5 | 4.64 | 91.89 | 100 | 5.06 | 27.79 | 100 | 5.03 | 4.89 | 4.38 | 5.29 | 5.09 | 100 | 4.88 | 98.13 | 100 |
| $\mathscr{H}_1$ | 5 | 0.02 | 13.26 | 99.93 | 0.04 | 0.45 | 99.86 | 0 | 0 | 0.64 | 0.02 | 0.03 | 98.25 | 0.01 | 18.87 | 99.88 |
| $\mathscr{H}_2$ | 5 | 0.02 | 13.72 | 100 | 0.05 | 0.53 | 99.98 | 0 | 0 | 1.01 | 0.02 | 0.03 | 99.46 | 0.01 | 19.39 | 99.99 |
| $\mathscr{H}_3$ | 5 | 4.68 | 91.94 | 100 | 5.14 | 27.96 | 100 | 5.12 | 4.94 | 4.44 | 5.39 | 5.12 | 100 | 4.98 | 98.13 | 100 |
| $\mathscr{RH}$ | 5 | 4.76 | 100 | 100 | 5.04 | 45.45 | 100 | 5.02 | 5.02 | 4.74 | 5.05 | 5.59 | 100 | 5.34 | 100 | 100 |
| $\mathscr{T}_1$ | 10 | 5.26 | 6.71 | 100 | 5.46 | 6.32 | 100 | 5 | 5.37 | 4.96 | 5.16 | 5.15 | 100 | 5.23 | 7.52 | 100 |
| $\mathscr{T}_2$ | 10 | 4.63 | 86.64 | 100 | 4.75 | 30.49 | 100 | 4.84 | 5.6 | 4.91 | 4.74 | 5.53 | 100 | 4.91 | 95.81 | 100 |
| $\mathscr{T}_3$ | 10 | 0.16 | 46.63 | 100 | 0.17 | 4.49 | 100 | 0.14 | 0.2 | 1.7 | 0.12 | 0.24 | 100 | 0.19 | 64.18 | 100 |
| $\mathscr{T}_4$ | 10 | 4.62 | 86.63 | 100 | 4.7 | 30.45 | 100 | 4.84 | 5.57 | 4.9 | 4.68 | 5.48 | 100 | 4.91 | 95.81 | 100 |
| $\mathscr{H}_1$ | 10 | 0.15 | 45.96 | 100 | 0.17 | 4.26 | 100 | 0.14 | 0.2 | 0.92 | 0.12 | 0.23 | 99.99 | 0.19 | 63.68 | 100 |
| $\mathscr{H}_2$ | 10 | 0.16 | 46.97 | 100 | 0.17 | 4.62 | 100 | 0.15 | 0.2 | 1.72 | 0.15 | 0.25 | 100 | 0.19 | 64.5 | 100 |
| $\mathscr{H}_3$ | 10 | 4.68 | 86.67 | 100 | 4.77 | 30.55 | 100 | 4.87 | 5.65 | 4.93 | 4.78 | 5.56 | 100 | 4.96 | 95.83 | 100 |
| $\mathscr{RH}$ | 10 | 4.7 | 100 | 100 | 4.5 | 67.61 | 100 | 5.01 | 5.44 | 4.89 | 4.78 | 5.69 | 100 | 4.85 | 100 | 100 |
| $\mathscr{T}_1$ | 20 | 5.07 | 10.67 | 100 | 5.27 | 8.1 | 100 | 4.84 | 5.15 | 5.03 | 4.82 | 5.45 | 100 | 4.99 | 11 | 100 |
| $\mathscr{T}_2$ | 20 | 5.07 | 86.47 | 100 | 5.17 | 31.8 | 100 | 4.79 | 5.3 | 5.07 | 5.16 | 5.51 | 100 | 4.87 | 93.16 | 100 |
| $\mathscr{T}_3$ | 20 | 1.2 | 79.4 | 100 | 1.38 | 17.44 | 100 | 1.1 | 1.46 | 2.87 | 1.22 | 1.52 | 100 | 1.28 | 89.05 | 100 |
| $\mathscr{T}_4$ | 20 | 5.03 | 86.43 | 100 | 5.13 | 31.71 | 100 | 4.78 | 5.23 | 5.06 | 5.14 | 5.46 | 100 | 4.87 | 93.16 | 100 |
| $\mathscr{H}_1$ | 20 | 1.16 | 79.11 | 100 | 1.28 | 17.08 | 100 | 1.03 | 1.42 | 1.44 | 1.11 | 1.43 | 100 | 1.2 | 88.91 | 100 |
| $\mathscr{H}_2$ | 20 | 1.21 | 79.52 | 100 | 1.43 | 17.58 | 100 | 1.13 | 1.48 | 2.91 | 1.26 | 1.56 | 100 | 1.32 | 89.1 | 100 |
| $\mathscr{H}_3$ | 20 | 5.08 | 86.49 | 100 | 5.22 | 31.83 | 100 | 4.83 | 5.33 | 5.13 | 5.17 | 5.54 | 100 | 4.88 | 93.16 | 100 |
| $\mathscr{RH}$ | 20 | 5.27 | 100 | 100 | 5.06 | 86.37 | 100 | 5.01 | 5.07 | 4.99 | 4.97 | 5.84 | 100 | 5.26 | 100 | 100 |

Table 1 (continued). Power of exogeneity tests at nominal level 5%; $G = 2$, $T = 500$

| | $k_2$ | $(a_1,a_2)' = (-20,0)'$ | | | $(a_1,a_2)' = (-5,5)'$ | | | $(a_1,a_2)' = (0,0)'$ | | | $(a_1,a_2)' = (.5,.2)'$ | | | $(a_1,a_2)' = (100,100)'$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\eta_1 = 0$ $\eta_2 = .5$ | $\eta_1 = .01$ $\eta_2 = .5$ | $\eta_1 = .5$ $\eta_2 = .5$ | $\eta_1 = 0$ $\eta_2 = .5$ | $\eta_1 = .01$ $\eta_2 = .5$ | $\eta_1 = .5$ $\eta_2 = .5$ | $\eta_1 = 0$ $\eta_2 = .5$ | $\eta_1 = .01$ $\eta_2 = .5$ | $\eta_1 = .5$ $\eta_2 = .5$ | $\eta_1 = 0$ $\eta_2 = .5$ | $\eta_1 = .01$ $\eta_2 = .5$ | $\eta_1 = .5$ $\eta_2 = .5$ | $\eta_1 = 0$ $\eta_2 = .5$ | $\eta_1 = .01$ $\eta_2 = .5$ | $\eta_1 = .5$ $\eta_2 = .5$ |
| $\mathscr{T}_1$ | 5 | 4.72 | 5.64 | 99.56 | 5.1 | 5.28 | 99.49 | 5.17 | 5.04 | 5.33 | 5.25 | 4.95 | 57.68 | 5.07 | 5.33 | 99.57 |
| $\mathscr{T}_2$ | 5 | 4.59 | 90.91 | 100 | 4.96 | 64.46 | 100 | 5.26 | 4.94 | 5.02 | 5.34 | 5.79 | 99.35 | 5.32 | 94.61 | 100 |
| $\mathscr{T}_3$ | 5 | 0.82 | 27.15 | 100 | 0.91 | 9.89 | 100 | 0.84 | 0.92 | 4.31 | 0.78 | 0.99 | 99.18 | 1.04 | 30.41 | 100 |
| $\mathscr{T}_4$ | 5 | 4.55 | 90.9 | 100 | 4.91 | 64.42 | 100 | 5.25 | 4.9 | 5 | 5.33 | 5.79 | 99.35 | 5.26 | 94.59 | 100 |
| $\mathscr{H}_1$ | 5 | 0.75 | 26.34 | 100 | 0.8 | 9.38 | 100 | 0.74 | 0.8 | 4.21 | 0.63 | 0.87 | 99.16 | 0.87 | 29.62 | 100 |
| $\mathscr{H}_2$ | 5 | 0.84 | 27.37 | 100 | 0.95 | 10.1 | 100 | 0.86 | 0.94 | 4.36 | 0.81 | 1.03 | 99.21 | 1.06 | 30.64 | 100 |
| $\mathscr{H}_3$ | 5 | 4.63 | 90.94 | 100 | 5 | 64.52 | 100 | 5.29 | 4.98 | 5.04 | 5.38 | 5.82 | 99.35 | 5.35 | 94.64 | 100 |
| $\mathscr{RH}$ | 5 | 4.7 | 100 | 100 | 4.98 | 99.1 | 100 | 5.07 | 5 | 4.86 | 5.54 | 6.46 | 97.45 | 5.41 | 100 | 100 |
| $\mathscr{T}_1$ | 10 | 5.19 | 7.33 | 100 | 4.93 | 6.55 | 100 | 4.83 | 4.97 | 5.13 | 5.2 | 4.85 | 91.46 | 5.19 | 7.75 | 100 |
| $\mathscr{T}_2$ | 10 | 5.31 | 86.33 | 100 | 5.32 | 50.06 | 100 | 4.99 | 4.95 | 4.87 | 5.28 | 5.7 | 99.56 | 4.99 | 91.52 | 100 |
| $\mathscr{T}_3$ | 10 | 1.59 | 61.19 | 100 | 1.58 | 21.63 | 100 | 1.42 | 1.69 | 4.34 | 1.63 | 1.96 | 99.39 | 1.61 | 69.56 | 100 |
| $\mathscr{T}_4$ | 10 | 5.3 | 86.32 | 100 | 5.29 | 49.98 | 100 | 4.96 | 4.94 | 4.83 | 5.24 | 5.66 | 99.55 | 4.94 | 91.51 | 100 |
| $\mathscr{H}_1$ | 10 | 1.45 | 59.83 | 100 | 1.43 | 20.57 | 100 | 1.22 | 1.44 | 4.21 | 1.46 | 1.67 | 99.36 | 1.41 | 68.37 | 100 |
| $\mathscr{H}_2$ | 10 | 1.62 | 61.44 | 100 | 1.63 | 21.73 | 100 | 1.44 | 1.71 | 4.38 | 1.69 | 2.01 | 99.41 | 1.62 | 69.85 | 100 |
| $\mathscr{H}_3$ | 10 | 5.36 | 86.34 | 100 | 5.35 | 50.17 | 100 | 5.02 | 5.02 | 4.92 | 5.3 | 5.75 | 99.56 | 5.01 | 91.54 | 100 |
| $\mathscr{RH}$ | 10 | 4.44 | 100 | 100 | 5.06 | 98.91 | 100 | 5.34 | 4.9 | 4.84 | 5 | 5.95 | 95.38 | 5.01 | 100 | 100 |
| $\mathscr{T}_1$ | 20 | 5.11 | 7.85 | 100 | 4.94 | 6.22 | 100 | 4.93 | 5.05 | 5.35 | 5.1 | 5.02 | 94.32 | 5.25 | 8.1 | 100 |
| $\mathscr{T}_2$ | 20 | 5.42 | 76.16 | 100 | 4.85 | 30.65 | 100 | 4.76 | 5.25 | 5.59 | 5.02 | 4.91 | 98.81 | 5.24 | 84.89 | 100 |
| $\mathscr{T}_3$ | 20 | 2.77 | 70.09 | 100 | 2.59 | 20.47 | 100 | 2.56 | 2.77 | 4.9 | 2.67 | 2.84 | 98.6 | 2.84 | 80.64 | 100 |
| $\mathscr{T}_4$ | 20 | 5.39 | 76.12 | 100 | 4.84 | 30.55 | 100 | 4.73 | 5.2 | 5.57 | 5 | 4.89 | 98.8 | 5.2 | 84.85 | 100 |
| $\mathscr{H}_1$ | 20 | 2.65 | 69.59 | 100 | 2.4 | 19.9 | 100 | 2.36 | 2.57 | 4.76 | 2.51 | 2.61 | 98.52 | 2.68 | 80.17 | 100 |
| $\mathscr{H}_2$ | 20 | 2.85 | 70.24 | 100 | 2.64 | 20.63 | 100 | 2.58 | 2.83 | 4.93 | 2.7 | 2.88 | 98.62 | 2.85 | 80.7 | 100 |
| $\mathscr{H}_3$ | 20 | 5.43 | 76.19 | 100 | 4.88 | 30.7 | 100 | 4.78 | 5.31 | 5.6 | 5.04 | 4.92 | 98.82 | 5.25 | 84.92 | 100 |
| $\mathscr{RH}$ | 20 | 5.19 | 100 | 100 | 4.61 | 94.7 | 100 | 4.66 | 5.12 | 5.32 | 5.03 | 5.29 | 86.39 | 5.59 | 100 | 100 |

### 6.2. Performance of OLS, 2SLS and two-stage estimators

Consider now a single simultaneous equations system described by the following DGP:

$$y = Y\beta + u, \quad Y = Z_2\Pi_2 + V, \tag{6.38}$$

where $y$ and $Y$ are $T \times 1$ random vectors $(G = 1)$, $Z_2$ is a $T \times k_2$ matrix of instruments such that $Z_{2t} \overset{i.i.d}{\sim} N(0, I_{k_2})$, $t = 1, \ldots, T$, and $\Pi_2$ is a vector of dimension $k_2$ with $\Pi_2 = \sqrt{\frac{\mu^2}{T\|Z_2C\|}}C$, where $C$ is a $k_2 \times 1$ vector of ones and $\mu^2$ is a concentration parameter. As in Guggenberger (2008), we cover several values of $\mu^2$ : $\mu^2 \in \{0; 13; 200; 613; 1,000; 1,000,000\}$ where the values of $\mu^2$ less than 613 correspond to those in Hansen, Hausman and Newey (2008). In our framework, small values of $\mu^2$ (say $\mu^2 \leq 613$) depicted cases where the IV are weak so that the parameter of interest $\beta$ is not identified or weakly identified. The correlation between $u$ and $V$ is set at $\rho \in \{0, .05, .1, .5, .6, .95\}$ and the true value of $\beta$ equals 1. We take $k_2 = 5$ instruments[1], so, both 2SLS and OLS estimators have finite moments. The sample size is $T = 500$ and the number of replications is $N = 10,000$. The results are presented in Tables 2 - 3 above.

In the first column of the tables, we report the different estimators while in the second, we report the concentration parameters $\mu^2$ which represents the quality of the IV. Finally, the other columns report the correlation $\rho$ between the errors and (possibly) endogenous regressors.

Our major findings can be summarized into two points: (1) over a wide range cases, including weak IV and moderate endogeneity, OLS performs better than 2SLS [finding similar to Kiviet and Niemczyk (2007)]; (2) pretest-estimators based on exogeneity have an excellent overall performance compared with usual IV estimator. This suggests that the practice of pretesting based on exogeneity tests is not to bad (at least in the viewpoint of estimation) as claimed by Guggenberger (2008).

[1] The choices of $k_2 = 10, 20$ lead to the same conclusions.

Table 2. Relative bias of OLS and two-stage estimators compared with 2SLS for $\beta = 10$

| **Estimators** | | $\mu^2 \downarrow, \rho \rightarrow$ | 0 | .05 | .1 | .5 | .6 | .95 |
|---|---|---|---|---|---|---|---|---|
| | | 0 | -0.07 | 0.83 | 1.06 | 1.00 | 1.00 | 1.00 |
| | | 13 | 0.35 | 1.48 | 1.01 | 1.02 | 0.99 | 1.00 |
| | | 200 | 0.11 | 1.19 | 1.26 | 1.09 | 1.08 | 1.09 |
| MCO | | 613 | 0.14 | 1.14 | 1.44 | 1.26 | 1.26 | 1.27 |
| | | 1000 | -0.10 | 1.36 | 1.58 | 1.44 | 1.45 | 1.41 |
| | | 2000000 | 0.83 | -83.23 | 43.87 | 132.32 | 135.86 | 105.40 |
| | Pre-tests | 0 | -0.02 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| two-stage | | 13 | 0.38 | 1.46 | 1.01 | 1.02 | 1.00 | 1.00 |
| | | 200 | 0.15 | 1.19 | 1.25 | 1.09 | 1.08 | 1.08 |
| | $\mathscr{T}_1$ | 613 | 0.18 | 1.13 | 1.42 | 1.24 | 1.24 | 1.25 |
| | | 1000 | -0.04 | 1.34 | 1.55 | 1.41 | 1.42 | 1.38 |
| | | 2000000 | 0.83 | -72.44 | 28.59 | 1.00 | 1.00 | 1.00 |
| | | 0 | -0.02 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| | | 13 | 0.38 | 1.46 | 1.01 | 1.02 | 1.00 | 1.00 |
| | | 200 | 0.15 | 1.19 | 1.25 | 1.09 | 1.08 | 1.08 |
| | $\mathscr{T}_2$ | 613 | 0.18 | 1.13 | 1.42 | 1.24 | 1.24 | 1.24 |
| | | 1000 | -0.05 | 1.34 | 1.55 | 1.40 | 1.41 | 1.36 |
| | | 2000000 | 0.83 | -67.81 | 20.37 | 1.00 | 1.00 | 1.00 |
| | | 0 | -0.07 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| | | 13 | 0.35 | 1.48 | 1.01 | 1.02 | 0.99 | 1.00 |
| | | 200 | 0.11 | 1.19 | 1.26 | 1.09 | 1.08 | 1.09 |
| | $\mathscr{T}_3$ | 613 | 0.14 | 1.14 | 1.44 | 1.25 | 1.25 | 1.26 |
| | | 1000 | -0.09 | 1.36 | 1.58 | 1.43 | 1.44 | 1.40 |
| | | 2000000 | 0.83 | -67.81 | 20.76 | 1.00 | 1.00 | 1.00 |
| | | 0 | -0.02 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| | | 13 | 0.38 | 1.46 | 1.01 | 1.02 | 1.00 | 1.00 |
| | | 200 | 0.15 | 1.19 | 1.25 | 1.09 | 1.08 | 1.08 |
| | $\mathscr{T}_4$ | 613 | 0.18 | 1.13 | 1.42 | 1.24 | 1.24 | 1.24 |
| | | 1000 | -0.05 | 1.34 | 1.55 | 1.40 | 1.41 | 1.36 |
| | | 2000000 | 0.83 | -67.81 | 20.37 | 1.00 | 1.00 | 1.00 |

Table 2 (Continued). Relative bias of OLS and two-stage estimators compared with 2SLS for $\beta = 10$

| **Estimators** | $\mu^2 \downarrow, \rho \rightarrow$ | 0 | .05 | .1 | .5 | .6 | .95 |
|---|---|---|---|---|---|---|---|
| | 0 | -0.07 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| | 13 | 0.35 | 1.48 | 1.01 | 1.02 | 0.99 | 1.00 |
| | 200 | 0.11 | 1.19 | 1.26 | 1.09 | 1.08 | 1.09 |
| $\mathscr{H}_1$ | 613 | 0.14 | 1.14 | 1.44 | 1.25 | 1.25 | 1.26 |
| | 1000 | -0.09 | 1.36 | 1.58 | 1.43 | 1.44 | 1.40 |
| | 2000000 | 0.83 | -67.81 | 20.37 | 1.00 | 1.00 | 1.00 |
| | 0 | -0.07 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| | 13 | 0.35 | 1.48 | 1.01 | 1.02 | 0.99 | 1.00 |
| | 200 | 0.11 | 1.19 | 1.26 | 1.09 | 1.08 | 1.09 |
| $\mathscr{H}_2$ | 613 | 0.14 | 1.14 | 1.44 | 1.25 | 1.25 | 1.26 |
| | 1000 | -0.09 | 1.36 | 1.58 | 1.43 | 1.44 | 1.40 |
| | 2000000 | 0.83 | -67.81 | 20.37 | 1.00 | 1.00 | 1.00 |
| | 0 | -0.02 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| | 13 | 0.38 | 1.46 | 1.01 | 1.02 | 1.00 | 1.00 |
| | 200 | 0.15 | 1.19 | 1.25 | 1.09 | 1.08 | 1.08 |
| $\mathscr{H}_3$ | 613 | 0.18 | 1.13 | 1.42 | 1.24 | 1.24 | 1.24 |
| | 1000 | -0.09 | 1.34 | 1.55 | 1.40 | 1.41 | 1.36 |
| | 2000000 | 0.83 | -67.81 | 20.37 | 1.00 | 1.00 | 1.00 |
| | 0 | -0.02 | 0.84 | 1.06 | 1.00 | 1.00 | 1.00 |
| | 13 | 0.38 | 1.46 | 1.01 | 1.02 | 1.00 | 1.00 |
| | 200 | 0.15 | 1.19 | 1.25 | 1.09 | 1.08 | 1.08 |
| $\mathscr{RH}$ | 613 | 0.18 | 1.13 | 1.42 | 1.24 | 1.24 | 1.24 |
| | 1000 | -0.05 | 1.34 | 1.55 | 1.40 | 1.41 | 1.36 |
| | 2000000 | 0.83 | -75.52 | 30.16 | 1.00 | 1.00 | 1.00 |

Table 3. Relative MSE of OLS and two-stage estimators compared with 2SLS for $\beta = 10$

| **Estimators** | | $\mu^2 \downarrow$, $\rho \rightarrow$ | 0 | .05 | .1 | .5 | .6 | .95 |
|---|---|---|---|---|---|---|---|---|
| MCO | | 0 | 0.01 | 0.01 | 0.04 | 0.43 | 0.52 | 0.72 |
| | | 13 | 0.01 | 0.01 | 0.04 | 0.44 | 0.52 | 0.73 |
| | | 200 | 0.01 | 0.01 | 0.04 | 0.48 | 0.58 | 1.09 |
| | | 613 | 0.01 | 0.02 | 0.04 | 1.26 | 0.71 | 1.05 |
| | | 1000 | 0.01 | 0.19 | 0.05 | 0.68 | 0.83 | 1.19 |
| | | 2000000 | 0.89 | 1.01 | 1.38 | 10.87 | 13.63 | 24.24 |
| two-stage | Pre-tests | 0 | 0.01 | 0.02 | 0.04 | 0.41 | 0.49 | 0.68 |
| | | 13 | 0.01 | 0.02 | 0.04 | 0.41 | 0.50 | 0.69 |
| | | 200 | 0.01 | 0.02 | 0.04 | 0.46 | 0.56 | 1.08 |
| | $\mathscr{T}_1$ | 613 | 0.01 | 0.02 | 0.04 | 1.24 | 0.67 | 0.97 |
| | | 1000 | 0.01 | 0.20 | 0.05 | 0.63 | 0.77 | 1.09 |
| | | 2000000 | 0.85 | 0.88 | 0.91 | 1.00 | 1.00 | 1.00 |
| | | 0 | 0.01 | 0.02 | 0.04 | 0.41 | 0.49 | 0.68 |
| | | 13 | 0.01 | 0.02 | 0.04 | 0.42 | 0.50 | 0.69 |
| | | 200 | 0.01 | 0.02 | 0.04 | 0.46 | 0.55 | 1.08 |
| | $\mathscr{T}_2$ | 613 | 0.01 | 0.02 | 0.04 | 1.24 | 0.66 | 0.94 |
| | | 1000 | 0.01 | 0.20 | 0.05 | 0.62 | 0.75 | 1.02 |
| | | 2000000 | 0.84 | 0.84 | 0.81 | 1.00 | 1.00 | 1.00 |
| | | 0 | 0.01 | 0.01 | 0.04 | 0.43 | 0.51 | 0.71 |
| | | 13 | 0.01 | 0.01 | 0.04 | 0.43 | 0.52 | 0.72 |
| | | 200 | 0.01 | 0.01 | 0.04 | 0.48 | 0.58 | 1.09 |
| | $\mathscr{T}_3$ | 613 | 0.01 | 0.02 | 0.04 | 1.25 | 0.71 | 1.03 |
| | | 1000 | 0.01 | 0.02 | 0.05 | 0.67 | 0.82 | 1.15 |
| | | 2000000 | 0.84 | 0.84 | 0.81 | 1.00 | 1.00 | 1.00 |
| | | 0 | 0.01 | 0.02 | 0.04 | 0.41 | 0.49 | 0.68 |
| | | 13 | 0.01 | 0.02 | 0.04 | 0.42 | 0.50 | 0.69 |
| | | 200 | 0.01 | 0.02 | 0.04 | 0.46 | 0.58 | 0.97 |
| | $\mathscr{T}_4$ | 613 | 0.01 | 0.02 | 0.04 | 1.24 | 0.66 | 0.94 |
| | | 1000 | 0.01 | 0.20 | 0.05 | 0.62 | 0.75 | 1.02 |
| | | 2000000 | 0.84 | 0.84 | 0.81 | 1.00 | 1.00 | 1.00 |

Table 3 (Continued). Relative MSE of OLS and two-stage estimators compared with 2SLS for $\beta = 10$

| **Estimators** | $\mu^2 \downarrow,\ \rho \rightarrow$ | 0 | .05 | .1 | .5 | .6 | .95 |
|---|---|---|---|---|---|---|---|
| | 0 | 0.01 | 0.01 | 0.04 | 0.43 | 0.51 | 0.71 |
| | 13 | 0.01 | 0.01 | 0.04 | 0.43 | 0.52 | 0.72 |
| | 200 | 0.01 | 0.01 | 0.04 | 0.48 | 0.58 | 1.09 |
| $\mathscr{H}_1$ | 613 | 0.01 | 0.02 | 0.04 | 1.25 | 0.71 | 1.03 |
| | 1000 | 0.01 | 0.02 | 0.05 | 0.67 | 0.82 | 1.15 |
| | 2000000 | 0.84 | 0.84 | 0.81 | 1.00 | 1.00 | 1.00 |
| | 0 | 0.01 | 0.01 | 0.04 | 0.43 | 0.51 | 0.71 |
| | 13 | 0.01 | 0.01 | 0.04 | 0.43 | 0.52 | 0.72 |
| | 200 | 0.01 | 0.01 | 0.04 | 0.48 | 0.58 | 1.03 |
| $\mathscr{H}_2$ | 613 | 0.01 | 0.02 | 0.04 | 1.25 | 0.70 | 1.03 |
| | 1000 | 0.01 | 0.02 | 0.05 | 0.67 | 0.82 | 1.15 |
| | 2000000 | 0.84 | 0.84 | 0.81 | 1.00 | 1.00 | 1.00 |
| | 0 | 0.01 | 0.02 | 0.04 | 0.41 | 0.49 | 0.68 |
| | 13 | 0.01 | 0.02 | 0.04 | 0.42 | 0.50 | 0.69 |
| | 200 | 0.01 | 0.02 | 0.04 | 0.46 | 0.55 | 1.09 |
| $\mathscr{H}_3$ | 613 | 0.01 | 0.02 | 0.04 | 1.24 | 0.66 | 0.94 |
| | 1000 | 0.01 | 0.20 | 0.05 | 0.62 | 0.75 | 1.02 |
| | 2000000 | 0.84 | 0.84 | 0.81 | 1.00 | 1.00 | 1.00 |
| | 0 | 0.01 | 0.01 | 0.04 | 0.41 | 0.49 | 0.68 |
| | 13 | 0.01 | 0.02 | 0.04 | 0.42 | 0.50 | 0.69 |
| | 200 | 0.01 | 0.02 | 0.04 | 0.46 | 0.55 | 1.08 |
| $\mathscr{R}\mathscr{H}$ | 613 | 0.01 | 0.02 | 0.04 | 1.24 | 0.66 | 0.93 |
| | 1000 | 0.01 | 0.20 | 0.05 | 0.62 | 0.75 | 1.00 |
| | 2000000 | 0.85 | 0.90 | 0.94 | 1.00 | 1.00 | 1.00 |

## 7. Empirical illustrations

This section illustrates the behaviour of exogeneity tests through two empirical applications related to important issues in macroeconomics and labor economics literature: the relation between trade and growth [see, Dufour and Taamouti (2006), Irwin and Tervio (2002), Frankel and Romer (1999), Harrison (1996), Mankiw and al. (1992)] and the widely studied problem of returns to education [Dufour and Taamouti (2006), Angrist and Krueger (1991), Angrist and Krueger (1995), Angrist and al. (1999), Mankiw and al. (1992)].

### 7.1. Trade and growth

The trade and growth model studies the relationship between standards of living and openness. The recent studies in this issue include Irwin and Tervio (2002), Frankel and Romer (1999), Harrison (1996), Mankiw and al. (1992) and the survey of Rodrik (1995). Even if many studies conclude that openness is conductive to higher growth, there is no evidence concerning the effect of openness on income. Estimating the impact of openness on income through cross-country regression often raises the problem of finding a good proxy for openness. Frankel and Romer (1999) argue that trade share (ratio of imports or exports to GDP) which is the commonly used indicator of openness should be viewed as endogenous variable, and similarly for the other indicators such as trade policies. So, instrumental variables method should be applied for estimating the income-trade relationship. The equation studied is

$$y_i = \alpha + \beta Tr_i + \gamma_1 N_i + \gamma_2 Ar_i + u_i, \tag{7.1}$$

where $y_i$ is log of income per capita in country $i$, $Tr_i$ the trade share (measured as a ratio of imports and exports to GDP), $N_i$ the logarithm of population, and $Ar_i$ the logarithm of country area. Since the trade share $Tr_i$ may be endogenous, Frankel and Romer (1999) used an instrument constructed on the basis of geographic characteristics. The first stage equation is given by

$$Tr_i = a + bX_i + c_1 N_i + c_2 Ar_i + v_i, \tag{7.2}$$

where $X_i$ is a constructed instrument from geographic characteristics. In this paper, we use the sample of 150 countries and the data include for each country: the trade share in 1985, the area and population (1985), per capita income (1985), and the fitted trade share (instrument)[2]. In this application, we focus on testing whether trade share is exogenous in (7.1). However, it is not clear how "weak "instruments are in this model. In fact, the F-statistic in the first stage regression (7.2) is around 13 [see Frankel and Romer (1999, Table 2, p.385)], which may indicate a possible weak identification problem [ Staiger-Stock(1997)]. Dufour and Taamouti (2006) proposed to use directly identification-robust procedures to draw inference on the coefficients of model (7.1). The projection approach shows that there is a slight difference between the usual 95 % IV-type confidence sets and the 95 % AR-based confidence sets of the coefficients of the structural equation (7.1). The 95 % IV-type confidence interval for the trade share coefficient is $[-.01\,,3.95]$, while the corresponding 95 % AR-based confidence set is $[.284\,,4.652]$. However, since all the confidence sets are bounded,

[2] The data set and its sources are given in the Appendix of Frankel and Romer (1999)

we do not have a serious problem of identification in this model. We provide an alternative way to access whether the instrument used is weak by examining the behaviour of DWH and $\mathscr{RH}$ statistics. For example, if the test for exogeneity based on these statistics does not reject trade share exogeneity, this may indicate that instrument are not "very poor ". Note that the model contains only one endogenous and one excluded instrument, hence $k_2 = G$, and the statistic $T_1$ is not considered in this application because it is identically zero. Table 4 below summarizes the results. In the first column of the table, we report the statistics while in the second and third columns, we report the sample values and the sample p-value of these tests. In the other columns, we report the Monte Carlo tests p-values for two data generating process where the disturbances $u$ are drawn from normal and Cauchy distributions.

Table 4. Tests for exogeneity of trade share in trade-income relation

| Statistics | Sample value | Sample p-value (%) | MC-test p-value (normal distribution) | MC-test p-value (Cauchy distribution) |
|---|---|---|---|---|
| $\mathscr{RH}$ | 3.9221 | $4.95^{*}$ | $5.02^{*}$ | $2.74^{*}$ |
| $\mathscr{H}_1$ | 2.3883 | 12.23 | 6.15 | $2.93^{*}$ |
| $\mathscr{H}_2$ | 2.4269 | 11.93 | 6.12 | $2.94^{*}$ |
| $\mathscr{H}_3$ | 3.9505 | $4.67^{*}$ | 5.49 | $2.85^{*}$ |
| $\mathscr{T}_2$ | 3.9221 | $4.95^{*}$ | 5.49 | $2.85^{*}$ |
| $\mathscr{T}_3$ | 2.3622 | 12.43 | 6.12 | $2.94^{*}$ |
| $\mathscr{T}_4$ | 3.8451 | $4.99^{*}$ | 5.49 | $2.85^{*}$ |

Note $-*$ : $H_0$ is rejected at nominal level $\alpha = 5\%$.

First, we note from Table 4 that $\mathscr{H}_3$, $\mathscr{T}_2$, $\mathscr{T}_4$ and $\mathscr{RH}$ , reject trade share exogeneity while $\mathscr{H}_1$, $\mathscr{H}_2$, and $\mathscr{T}_3$, cannot reject the null hypothesis. When we run exact Monte Carlo tests (for Gaussian and Cauchy type errors), we see that all statistics strongly reject trade share exogeneity at level 5 %, which means that the quality of the instrument is not too poor in this model as noted by Dufour and Taamouti (2006) . Our results also underscore the difference between exact Monte Carlo exogeneity procedures and earlier procedures.

### 7.2. Education and earnings

This application considers the well known problem of estimating returns to education. The literature in this issue includes Angrist and Krueger (1991), Angrist and Krueger (1995), Angrist and al. (1999), Bound et al. (1995). The equation studies is a relationship where the log weekly earning is explained by the number of years of education and several other covariates (age, age squared, year of birth, . . .). Since education can be viewed as an endogenous variable, Angrist and Krueger (1991) used the birth quarter as an instrument. The basic idea is that individuals born in the first quarter of the year start school at an older age, and can therefore drop out after completing less schooling than individuals born near the end of the year. Consequently, individuals born at the beginning of the year are likely to earn less than those born during the rest of the year[3]. However, it is well known that the instruments used by Angrist and Krueger (1991) are weak and explains very little of the variation in education; see Bound et al. (1995). So, standard IV-based inference is quite unreliable. As showed in this paper, DWH or RH tests for the exogeneity of education will lead to accept the null hypothesis of exogeneity of this variable. The model considered is specified as:

$$y = \beta_0 + \beta_1 E + \sum_{i=1}^{k_1} \gamma_i X_i + u\,, \tag{7.3}$$

$$E = \Pi_2^0 + \sum_{i=1}^{k_2} \pi_i Z_i + \sum_{i=1}^{k_1} \phi_i X_i + v\,, \tag{7.4}$$

where $y$ is log-weekly earnings, $E$ is the number of years of education (possibly endogenous), $X$ contains the exogenous covariates (age, age squared, 10 dummies for birth of year). $Z$ contains 40 dummies obtained by interacting the quarter of birth with the year of birth. In this model, $\beta_1$ measures the return to education. The data set consists of the 5% public-use sample of the 1980 US census for men born between 1930 and 1939. The sample size is 329 509 observations. We test the exogeneity of education in this model using DWH and RH statistics. The results are summarized in Table 5. As showed in this table, all exogeneity tests cannot reject the exogeneity of "education "even at level 15%. This is true for earlier versions of the tests or the MCE-tests.

The results can be interpreted as follow: (a) either the instruments are strong and education is effectively exogenous, (b) or education is endogenous but the instruments are too poor and the tests fail to detect that education is endogenous. Moreover, it is well documented that the generated instruments obtained by interacting the quarter of birth with the year of birth are weak, see e.g., Bound et al. (1995). So, our interpretation in (b) matter with these observations.

Table 5. Tests for exogeneity of education in income-education equation.

| Statistics | Sample value | Sample p-value | MC-test p-value (normal distribution) | MC-test p-value (Cauchy distribution) |
|---|---|---|---|---|
| $\mathscr{RH}$ | .6783 | .93986 | .6590 | .9451 |
| $\mathscr{H}_1$ | 1.337 | .24757 | .2474 | .2488 |
| $\mathscr{H}_2$ | 1.337 | .24756 | .2474 | .2488 |
| $\mathscr{H}_3$ | 1.3492 | .24542 | .2474 | .2488 |
| $\mathscr{T}_1$ | 2.0406 | .16111 | .2302 | .2308 |
| $\mathscr{T}_2$ | 1.3491 | .24543 | .2474 | .2488 |
| $\mathscr{T}_3$ | 1.3369 | . 224757 | .2474 | .2488 |
| $\mathscr{T}_4$ | 1.3491 | .24543 | .2474 | .2488 |

[3]Other versions of the IV regression take as instruments interactions between the birth quarter and regional and/or birth year dummies.

## 8. Conclusion

Exogeneity tests of the type proposed by Durbin (1954), Wu (1973), Hausman (1978) (DWH) and, Revankar and Hartley (1973) (RH) are built on the prerequisite of having strong IVs. Not much is known about their behaviour of such tests when identification is weak. This paper proposes a large-sample analysis of the distribution of these tests under the null hypothesis (level) and the alternative hypothesis (power). Two main contributions is established.

First, the characterization of the large-sample distribution of the test statistics shows that DWH- and RH-type tests are typically robust to weak IV. We provide a provide a necessary and sufficient condition under which the tests have no power. In particular, the tests have no power when all IV are weak [similar to Guggenberger (2008)]. But, power does exist as soon as we have one strong IV. The conclusions of Guggenberger (2008) focus on the case where all IV are weak (a case of little practical interest).

Furthermore, we present simulation evidence indicating that: (1) Over a wide range cases, including weak IV and moderate endogeneity, OLS performs better than 2SLS [Similar to Kiviet and Niemczyk (2007)]; (2) pretest-estimators based on exogeneity tests have an excellent overall performance compared with OLS and IV estimators. We illustrate our theoretical results through two empirical applications: the returns to education and the relation between trade and economic growth. We find that exogeneity tests cannot reject the exogeneity of schooling, indicating that IVs are possibly weak in this model [Bound et al. (1995)]. However, “trade share ”is endogenous, *i.e.* IVs are not too poor [similar to Dufour and Taamouti (2006)].

APPENDIX

# A. Proofs

PROOF OF THEOREM **3.1** Assume that $\delta = 0$. Then, we have $a = \Sigma_V^{-1}\delta = 0$. We shall distinguish two cases: (A) $\Pi_2 = \Pi_2^0$ where $\Pi_2^0$ is a $k_2 \times G$ constant matrix with rank $G$; and (B) $\Pi_2 = \Pi_2^0$, rank$(\Pi_2^0) < G$.
(A) Suppose first that $\Pi_2 = \Pi_2^0$, with rank$(\Pi_2^0) = G$ (strong identification). Then, we have:

$$\hat{\Omega}_{IV} \xrightarrow{p} \Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0 \, , \, \hat{\Omega}_{LS} \xrightarrow{p} \Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V \, , \tag{A.1}$$

$$Y'u/T \xrightarrow{p} \delta = 0 \, , \, Y'M_1u/T \xrightarrow{p} \delta = 0. \tag{A.2}$$

From (A.1) - (A.2), we get

$$\begin{aligned} \hat{\sigma}^2/T &= \hat{u}'\hat{u}/T = u'u/T - (u'M_1Y/T)\hat{\Omega}_{LS}^{-1}(Y'M_1u/T) \xrightarrow{p} \sigma_u^2 \, , & \text{(A.3)} \\ \tilde{\sigma}^2/T &= u'u/T - 2(u'M_1Y/T)\hat{\Omega}_{IV}^{-1}(Y'(M_1 - M)u/T) \\ &\quad + (u'(M_1 - M)Y/T)\hat{\Omega}_{IV}^{-1}(Y'(M_1 - M)u/T) \xrightarrow{p} \sigma_u^2 \, . & \text{(A.4)} \end{aligned}$$

Moreover, we can write $\frac{Y'u}{\sqrt{T}}$ as

$$\frac{Y'u}{\sqrt{T}} = \bar{\Pi}_1' \frac{Z_1'u}{\sqrt{T}} + \Pi_2^{0\prime}\frac{\bar{Z}_2'u}{\sqrt{T}} + \frac{V'V}{\sqrt{T}}(\Sigma_V^{-1}\delta) + \frac{V'\varepsilon}{\sqrt{T}} = \bar{\Pi}_1' \frac{Z_1'u}{\sqrt{T}} + \Pi_2^{0\prime}\frac{\bar{Z}_2'u}{\sqrt{T}} + \frac{V'\varepsilon}{\sqrt{T}}, \tag{A.5}$$

where $\bar{Z}_2 = M_1Z_2$ and $\bar{\Pi}_1 = \Pi_1 + (Z_1'Z_1)^{-1}Z_1'Z_2\Pi_2$. Since $\bar{\Pi}_1 \xrightarrow{p} \bar{\Pi}_{01} = \Pi_1 + \Sigma_{Z_1}^{-1}\Sigma_{Z_1Z_2}\Pi_2^0$, it follows that

$$\frac{Y'u}{\sqrt{T}} \xrightarrow{L} \bar{\Pi}_{01}' S_{1u} + \Pi_2^{0\prime}\bar{S}_{2u} + S_{V\varepsilon} \, . \tag{A.6}$$

Thus, we get

$$\begin{aligned} \frac{Y'M_1u}{\sqrt{T}} &= \frac{Y'u}{\sqrt{T}} - \left(\frac{Y'Z_1}{T}\right)\left(\frac{Z_1'Z_1}{T}\right)^{-1}\frac{Z_1'u}{\sqrt{T}} \\ &\xrightarrow{L} (\bar{\Pi}_{01}' S_{1u} + \Pi_2^{0\prime}\bar{S}_{2u} + S_{V\varepsilon}) - \bar{\Pi}_{01}' S_{1u} = \Pi_2^{0\prime}\bar{S}_{2u} + S_{V\varepsilon} \, , \end{aligned} \tag{A.7}$$

$$\frac{1}{\sqrt{T}}Y'(M_1 - M)u = \left(\frac{Y'\bar{Z}_2}{T}\right)\left(\frac{\bar{Z}_2'\bar{Z}_2}{T}\right)^{-1}\left(\frac{\bar{Z}_2'u}{\sqrt{T}}\right) \xrightarrow{L} \Pi_2^{0\prime}\bar{S}_{2u} \, . \tag{A.8}$$

We can then observe that

$$\sqrt{T}(\tilde{\beta} - \hat{\beta}) = \hat{\Omega}_{LS}^{-1}\frac{Y'M_1u}{\sqrt{T}} - \hat{\Omega}_{IV}^{-1}\frac{Y'(M_1 - M)u}{\sqrt{T}} \xrightarrow{L} \psi_\pi \, , \tag{A.9}$$

$$\hat{\blacksquare}_i \xrightarrow{p} \sigma_u^2 \Delta_\Pi,\, \Delta_\Pi = (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0)^{-1} - (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 + \Sigma_V)^{-1},\, i = 1, 2, 3\,, \tag{A.10}$$

where

$$\psi_\pi = (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 + \Sigma_V)^{-1} (\Pi_2^{0\prime} \bar{S}_{2u} + S_{V\varepsilon}) - (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0)^{-1} \Pi_2^{0\prime} \bar{S}_{2u}, \tag{A.11}$$

so that

$$\mathscr{H}_i = \sqrt{T}(\tilde{\beta} - \hat{\beta})' \hat{\blacksquare}_i^{-1} \sqrt{T}(\tilde{\beta} - \hat{\beta}) \xrightarrow{L} \frac{1}{\sigma_u^2} \psi_\pi' \Delta_\Pi^{-1} \psi_\pi,\, i = 1, 2, 3\,. \tag{A.12}$$

Since $a = 0$, we have $\sigma_u^2 = \sigma_\varepsilon^2$, hence

$$\left[ \begin{array}{c} \Pi_2^{0\prime} \bar{S}_{2u} + S_{V\varepsilon} \\ \Pi_2^{0\prime} \bar{S}_{2u} \end{array} \right] \sim N\left[0,\, \sigma_u^2 \Sigma_0\right]. \tag{A.13}$$

where

$$\Sigma_0 = \left[ \begin{array}{cc} \Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 + \Sigma_V & \Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 \\ \Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 & \Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 \end{array} \right]. \tag{A.14}$$

This entails

$$\psi_\pi \sim N\left\{0,\, \sigma_u^2 [(\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0)^{-1} - (\Sigma_V + \Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0)^{-1}]\right\} \equiv N(0,\, \sigma_u^2 \Delta_\Pi), \tag{A.15}$$

hence

$$\mathscr{H}_i \xrightarrow{L} \chi^2(G),\, i = 1, 2, 3\,. \tag{A.16}$$

Applying the same arguments as above, we get

$$\mathscr{T}_2 \xrightarrow{L} \frac{1}{G} \chi^2(G),\; \mathscr{T}_l \xrightarrow{L} \chi^2(G),\, l \quad = \quad 3, 4, \quad \text{and} \quad \mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2} \chi^2(k_2)\,. \tag{A.17}$$

We now derive the distribution of $T_1$. We can write

$$\mathscr{T}_1 = \frac{k_2 - G}{G} \frac{T(\tilde{\beta} - \hat{\beta})' \Delta^{-1} (\tilde{\beta} - \hat{\beta})}{T \tilde{\sigma}_1^2} \tag{A.18}$$

and $T(\tilde{\beta} - \hat{\beta})' \Delta^{-1} (\tilde{\beta} - \hat{\beta}) \xrightarrow{L} \psi_\pi' \Delta_\Pi^{-1} \psi_\pi \sim \sigma_u^2 \chi^2(G)$. Furthermore, because $Z_1$ is orthogonal to $\bar{Z}_2$, we can observe that

$$T \tilde{\sigma}_1^2 = u'((M_1 - M) - P_{\tilde{Y}})u = u'(M_1 - M)u - u' P_{\tilde{Y}} u\,,$$

where $\tilde{Y} = (M_1 - M)Y$. Thus, we have

$$\begin{aligned} T \tilde{\sigma}_1^2 \quad &\xrightarrow{L} \quad \bar{S}_{2u}' \Sigma_{\bar{Z}_2}^{-1} \bar{S}_{2u} - \bar{S}_{2u}' \Pi_2^0 (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0)^{-1} \Pi_2^{0\prime} \bar{S}_{2u} \\ &= \quad \bar{S}_{2u}' \Sigma_{\bar{Z}_2}^{-1/2} [I_{k_2} - P(P'P)^{-1} P'] \Sigma_{\bar{Z}_2}^{-1/2} \bar{S}_{2u}\,, \end{aligned} \tag{A.19}$$

where $P = \Sigma_{\bar{Z}_2}^{1/2}\Pi_2^0$ and the matrix $I_k - P(P'P)^{-1}P'$ is idempotent with rank $k_2 - G$. Furthermore, $\frac{1}{\sigma_u}\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2u} \sim N[0, I_{k_2}]$, hence $T\tilde{\sigma}_1^2 \xrightarrow{L} \sigma_u^2\chi^2(k_2 - G)$. Moreover, we can write $T\tilde{\sigma}_1^2 = u'(M_1 - M)M_{\hat{Y}}(M_1 - M)u$ and $T(\tilde{\beta} - \hat{\beta})'\Delta^{-1}(\tilde{\beta} - \hat{\beta}) = u'A_Z u,$ where

$$\mathscr{A}_Z = \frac{1}{T}(M_1 Y \hat{\Omega}_{LS}^{-1} - \hat{Y}\hat{\Omega}_{IV}^{-1})\Delta^{-1}(\hat{\Omega}_{LS}^{-1}Y'M_1 - \hat{\Omega}_{IV}^{-1}\hat{Y}')$$

is symmetric, idempotent and $A_Z((M_1 - M)M_{\hat{Y}}\bar{M}_1) = ((M_1 - M)M_{\hat{Y}}(M_1 - M))A_Z = 0$. This entails that $T\tilde{\sigma}_1^2$ and $T(\tilde{\beta} - \hat{\beta})'\Delta^{-1}(\tilde{\beta} - \hat{\beta})$ are independent, hence asymptotically independent and distributed as $\chi^2$ with $k_2 - G$ and $G$ degrees of freedom respectively. Consequently,

$$\mathscr{T}_1 \xrightarrow{L} F(G, k_2 - G). \tag{A.20}$$

(B) Suppose now that $\Pi_2 = \Pi_2^0$, where $(\Pi_2^0) < G$. We shall only prove the validity of $\mathscr{H}_3$. The validity of other statistics can be proved in a similar way. We recall that

$$\mathscr{H}_3 = T(\tilde{\beta} - \hat{\beta})'\hat{\blacksquare}_3^{-1}(\tilde{\beta} - \hat{\beta}), \tag{A.21}$$

where $\hat{\blacksquare}_3^{-1} = \frac{1}{\hat{\sigma}^2}\hat{\Delta}^{-1}$ with $\hat{\Delta} = \hat{\Omega}_{IV}^{-1} - \hat{\Omega}_{LS}^{-1}$ and $\hat{\Omega}_{LS}$ and $\hat{\Omega}_{IV}$ are defined in (2.33).

Using (3.9), we can now write equation (A.21) as

$$\begin{aligned}
\mathscr{H}_3 &= [(\hat{\beta} - \beta) - \mathscr{S}_1(\tilde{\beta}_1 - \beta_1) - \mathscr{S}_2(\tilde{\beta}_2 - \beta_2)]'\hat{\blacksquare}_3^{*^{-1}}[(\hat{\beta} - \beta) - \mathscr{S}_1(\tilde{\beta}_1 - \beta_1) - \mathscr{S}_2(\tilde{\beta}_2 - \beta_2)], && \text{(A.22)}\\
&= (\tilde{\beta}_2 - \beta_2)'\mathscr{S}_2'\hat{\blacksquare}_3^{*^{-1}}\mathscr{S}_2(\tilde{\beta}_2 - \beta_2) - 2[(\hat{\beta} - \beta) - \mathscr{S}_1(\tilde{\beta}_1 - \beta_1)]'\hat{\blacksquare}_3^{*^{-1}}\mathscr{S}_2(\tilde{\beta}_2 - \beta_2) \\
&\quad + [(\hat{\beta} - \beta) - \mathscr{S}_1(\tilde{\beta}_1 - \beta_1)]'\hat{\blacksquare}_3^{*^{-1}}[(\hat{\beta} - \beta) - \mathscr{S}_1(\tilde{\beta}_1 - \beta_1)] && \text{(A.23)}
\end{aligned}$$

where $\hat{\blacksquare}_3^* = \hat{\sigma}^2[(Y'(M_1 - M)Y)^{-1} - \frac{1}{T}(Y'M_1Y/T)^{-1}]$. We first find the limit of $\hat{\blacksquare}_3^*$. Since $(Y'M_1Y/T)^{-1} \xrightarrow{p} (\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}$, hence, we have $\frac{1}{T}(Y'M_1Y/T)^{-1} \xrightarrow{p} 0$. It is also easy to see that $\hat{\sigma}^2 \xrightarrow{p} \sigma_u^2$. We now focus on $[Y'(M_1 - M)Y]^{-1}$. We have

$$\begin{aligned}
\mathscr{S}'Y'(M_1 - M)Y\mathscr{S} &= \begin{bmatrix} Y_1' \\ Y_2' \end{bmatrix}(M_1 - M)\begin{bmatrix} Y_1 & Y_2 \end{bmatrix} \\
&= \begin{bmatrix} Y_1'(M_1 - M)Y_1 & Y_1'(M_1 - M)Y_2 \\ Y_2'(M_1 - M)Y_1 & Y_2'(M_1 - M)Y_2 \end{bmatrix},
\end{aligned} \tag{A.24}$$

So, the partitioned inverse of $\mathscr{S}'Y'(M_1 - M)Y\mathscr{S}$ can be written as

$$\mathscr{S}'[Y'(M_1 - M)Y]^{-1}\mathscr{S} = \begin{bmatrix} P_{11} & P_{21}' \\ P_{21} & P_{22} \end{bmatrix}, \tag{A.25}$$

where

$$P_{11} = [Y_1'(M_1 - M)Y_1 - Y_1'(M_1 - M)Y_2(Y_2'(M_1 - M)Y_2)^{-1}Y_2'(M_1 - M)Y_1]^{-1}, \tag{A.26}$$

$$
\begin{aligned}
P_{21} &= -(Y_2'(M_1-M)Y_2)^{-1}Y_2'(M_1-M)Y_1P_{11}, &(A.27)\\
P_{22} &= (Y_2'(M_1-M)Y_2)^{-1}+(Y_2'(M_1-M)Y_2)^{-1}Y_2'(M_1-M)Y_1P_{11}\\
&\quad \times Y_1'(M_1-M)Y_2(Y_2'(M_1-M)Y_2)^{-1}. &(A.28)
\end{aligned}
$$

However, we have

$$
\begin{aligned}
Y_1'(M_1-M)Y_1/T &\xrightarrow{p} \Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21},\, Y_2'(M_1-M)Y_1/T \xrightarrow{p} 0, &(A.29)\\
Y_2'(M_1-M)Y_2 &\xrightarrow{L} \mathscr{S}_2'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}\mathscr{S}_2,\; Y_2'(M_1-M)Y_1/\sqrt{T} \xrightarrow{L} \mathscr{S}_2'\bar{S}_{2V}'\Pi_{21}. &(A.30)
\end{aligned}
$$

So, we get

$$
\begin{aligned}
TP_{11} &\xrightarrow{L} \bar{P}_{11} = [\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21}-\Pi_{21}\bar{S}_{2V}\mathscr{S}_2(\mathscr{S}_2'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}\mathscr{S}_2)^{-1}\mathscr{S}_2'\bar{S}_{2V}'\Pi_{21}]^{-1}, &(A.31)\\
T^{1/2}P_{21} &\xrightarrow{L} \bar{P}_{21} = -(\mathscr{S}_2'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}\mathscr{S}_2)^{-1}\mathscr{S}_2'\bar{S}_{2V}'\Pi_{21}\bar{P}_{11}, &(A.32)\\
P_{22} &\xrightarrow{L} \bar{P}_{22} = (\mathscr{S}_2'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}\mathscr{S}_2)^{-1}+(\mathscr{S}_2'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}\mathscr{S}_2)^{-1}\\
&\quad \times\mathscr{S}_2'\bar{S}_{2V}'\Pi_{21}\bar{P}_{11}\Pi_{21}'\bar{S}_{2V}\mathscr{S}_2(\mathscr{S}_2'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}\mathscr{S}_2)^{-1}. &(A.33)
\end{aligned}
$$

Hence, we have

$$
[Y'(M_1-M)Y]^{-1} \xrightarrow{L} \mathscr{S}\begin{bmatrix} 0 & 0\\ 0 & \bar{P}_{22}\end{bmatrix}\mathscr{S}' = \mathscr{S}_2\bar{P}_{22}\mathscr{S}_2'. \tag{A.34}
$$

Furthermore, under $\delta=0$, we have $\hat{\beta}-\beta \xrightarrow{p} 0$, and using (3.11), we can show that

$$
\begin{aligned}
\tilde{\beta}_1-\beta_1 &\xrightarrow{p} 0,\, Y_2'(M_1-M)u \xrightarrow{L} \mathscr{S}_2'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u} &(A.35)\\
\frac{Y_1'(M_1-M)u}{\sqrt{T}} &\xrightarrow{L} \Pi_{21}'\bar{S}_{2u}. &(A.36)
\end{aligned}
$$

So, we have

$$
\tilde{\beta}_2-\beta_2 \xrightarrow{L} \mathscr{B}^{-1}\mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1}-\Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2u} \equiv \mathscr{N}_B, \tag{A.37}
$$

where $\mathscr{B}=\mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1}-\Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2V}\mathscr{S}_2$. Moreover, because from (2.3) we have $\bar{S}_{2u}=\bar{S}_{2V}a+\bar{S}_{2\varepsilon}$, we easily get

$$
\mathscr{N}_B=\mathscr{S}_2'a+\mathscr{B}^{-1}\mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1}-\Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2\varepsilon}. \tag{A.38}
$$

Under $H_0:\delta=0$, $\mathscr{N}_B|_{\bar{S}_{2V}} \sim N[0,\sigma_\varepsilon^2\mathscr{B}^{-1}]$. Hence, we have $\mathscr{S}_2\mathscr{N}_B|_{\bar{S}_{2V}} \sim N[0,\sigma_\varepsilon^2\mathscr{A}^{-1}]$, where $\mathscr{A}^{-1}=\mathscr{S}_2\mathscr{B}^{-1}\mathscr{S}_2' : G\times G$. From (A.37), we get

$$
\mathscr{A} = \bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1}-\Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2V}.
$$

Now, using Anderson (2003, Theorem A.3.3 and Theorem A.3.4), we can write $P_{22}$ as

$$\begin{aligned} P_{22} &= [Y_2'(M_1-M)Y_2 - Y_2'(M_1-M)Y_1(Y_1'(M_1-M)Y_1)^{-1}Y_1'(M_1-M)Y_2]^{-1} \\ &= (Y_2'JY_2)^{-1} \xrightarrow{L} \mathscr{B}^{-1} = \bar{P}_{22}. \end{aligned} \tag{A.39}$$

And by noting that $(\mathscr{S}_2\bar{P}_{22}\mathscr{S}_2')^{-1} = (\mathscr{S}_2\mathscr{B}^{-1}\mathscr{S}_2')^{-1} = \mathscr{A}$, we have

$$\mathscr{H}_3 \xrightarrow{L} \frac{1}{\sigma_\varepsilon^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B. \tag{A.40}$$

Since $\mathscr{S}_2\mathscr{N}_B|_{\bar{S}_{2V}} \sim N[0,\, \sigma_\varepsilon^2\mathscr{A}^{-1}]$, we get

$$\mathscr{H}_3 \xrightarrow{L} \frac{1}{\sigma_\varepsilon^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B|_{\bar{S}_{2V}} \sim \chi^2(G). \tag{A.41}$$

Because the conditional null distribution does not depend neither on $\bar{S}_{2V}$, we have

$$\mathscr{H}_3 \xrightarrow{L} \chi^2(G), \tag{A.42}$$

and $\mathscr{H}_3$ still is valid even if identification is deficient.

We will now focus on $\mathscr{H}_1$ and $\mathscr{H}_2$. First, we note that

$$\mathscr{H}_2 = \frac{\hat{\sigma}^2}{\tilde{\sigma}^2}\mathscr{H}_3, \tag{A.43}$$

where $\hat{\sigma}^2 \xrightarrow{p} \sigma_\varepsilon^2$ and

$$\begin{aligned} \tilde{\sigma}^2 &= \frac{u'u}{T} - 2\frac{u'M_1Y}{T}(\tilde{\beta}-\beta) + (\tilde{\beta}-\beta)'\hat{\Omega}_{LS}(\tilde{\beta}-\beta) \\ &\xrightarrow{L} \tilde{\sigma}_u^2 = \sigma_u^2 + \mathscr{N}_B'\mathscr{S}_2'(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\mathscr{N}_B\mathscr{S}_2 \geq \sigma_\varepsilon^2. \end{aligned} \tag{A.44}$$

Hence

$$\mathscr{H}_2 \xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \frac{1}{\sigma_\varepsilon^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \sim \chi^2(G). \tag{A.45}$$

Second, using (2.31), we can easily show that $T\hat{\Sigma}_1 \xrightarrow{p} \frac{1}{\tilde{\sigma}_u^2}\mathscr{A}$, so that

$$\mathscr{H}_1 \xrightarrow{L} \frac{1}{\tilde{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \leq \frac{1}{\sigma_\varepsilon^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \sim \chi^2(G). \tag{A.46}$$

By using the relations between $\mathscr{T}_l$ and $\mathscr{H}_i$, we get the results for $\mathscr{T}_2$, $\mathscr{T}_3$ and $\mathscr{T}_4$. Finally, by following the same steps as for $\mathscr{H}_3$, we get the results for $\mathscr{T}_1$ and $\mathscr{RH}$. Clearly, all exogeneity tests are valid even if identification is deficient.

□

PROOF OF THEOREM **3.2** (A) Suppose that $\Pi_2 = \Pi_2^0$ with $\mathrm{rank}(\Pi_2^0) = G$. From the proof of Theorem **3.1**, we have

$$\hat{\Omega}_{IV} \xrightarrow{p} \Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0,\, \hat{\Omega}_{LS} \xrightarrow{p} \Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 + \Sigma_V\,, \tag{A.47}$$

$$\begin{aligned}
\frac{Y' M_1 u}{T} &\xrightarrow{p} \delta,\, \frac{Y'(M_1 - M)u}{T} \xrightarrow{p} 0\,, && \text{(A.48)}\\
\hat{\sigma}^2 &= u'u/T - (u' M_1 Y/T) \hat{\Omega}_{LS}^{-1} (Y' M_1 u/T) \\
&\xrightarrow{p} \sigma_u^2 - \delta' (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 + \Sigma_V)^{-1} \delta = \tilde{\sigma}_u^2\,, && \text{(A.49)}\\
\tilde{\sigma}^2 &= u'u/T - 2(u' M_1 Y/T) \hat{\Omega}_{IV}^{-1} (Y'(M_1 - M)u/T) + \\
&\quad (u'(M_1 - M)Y/T) \hat{\Omega}_{IV}^{-1} (Y'(M_1 - M)u/T) \xrightarrow{p} \sigma_u^2, && \text{(A.50)}
\end{aligned}$$

so that we get

$$(\tilde{\beta} - \hat{\beta}) = \hat{\Omega}_{LS}^{-1} (Y' M_1 u/T) - \hat{\Omega}_{IV}^{-1} (Y'(M_1 - M)u/T) \xrightarrow{p} (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V)^{-1} \delta\,, \tag{A.51}$$

$$\begin{aligned}
\hat{\blacksquare}_i &\xrightarrow{p} \sigma_i^2 \Delta_\Pi,\, \Delta_\Pi = (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0)^{-1} - (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 + \Sigma_V)^{-1},\, i = 2, 3, && \text{(A.52)}\\
\hat{\blacksquare}_1 &\xrightarrow{p} \Sigma_{1\Pi},\, \Sigma_{1\Pi} = \sigma_u^2 (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0)^{-1} - \tilde{\sigma}_u^2 (\Pi_2^{0\prime} \Sigma_{\bar{Z}_2} \Pi_2^0 + \Sigma_V)^{-1}, && \text{(A.53)}
\end{aligned}$$

where $\sigma_2^2 = \sigma_u^2$ and $\sigma_3^2 = \tilde{\sigma}_u^2$. Let first focus on $\mathscr{H}_i$, $i = 1, 2, 3$. We recall that $\mathscr{H}_i$ is defined as

$$\mathscr{H}_i = T(\tilde{\beta} - \hat{\beta})' \hat{\blacksquare}_i^{-1} (\tilde{\beta} - \hat{\beta}) \tag{A.54}$$

and from (A.51)-(A.53), we have

$$\begin{aligned}
(\tilde{\beta} - \hat{\beta})' \hat{\blacksquare}_i^{-1} (\tilde{\beta} - \hat{\beta}) &\xrightarrow{p} \frac{1}{\sigma_i^2} \delta' (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V)^{-1} \Delta_\Pi^{-1} (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V)^{-1} \delta,\, i = 2, 3,\\
(\tilde{\beta} - \hat{\beta})' \hat{\blacksquare}_1^{-1} (\tilde{\beta} - \hat{\beta}) &\xrightarrow{p} \delta' (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V)^{-1} \Sigma_{1\Pi}^{-1} (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V)^{-1} \delta\,. && \text{(A.55)}
\end{aligned}$$

Using Doko and Dufour (2010, Lemma A.1), we have

$$\Delta_\Pi^{-1} = (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0) \Sigma_V^{-1} (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V), \tag{A.56}$$

hence

$$\begin{aligned}
&\delta' (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V)^{-1} \Delta_\Pi^{-1} (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0 + \Sigma_V)^{-1} \delta\\
= \;& a' \Pi_2^{0\prime} \Sigma_Z \Pi_2^0 [(\Pi_2^{0\prime} \Sigma_Z \Pi_2^0) \Sigma_V^{-1} (\Pi_2^{0\prime} \Sigma_Z \Pi_2^0) + \Pi_2^{0\prime} \Sigma_Z \Pi_2^0]^{-1} \Pi_2^{0\prime} \Sigma_Z \Pi_2^0 a. && \text{(A.57)}
\end{aligned}$$

If $\Pi_2^0 a \neq 0$, then the RHS of (A.57) is positive and we have $H_i \xrightarrow{L} +\infty$ for $i = 2,\ 3$. The same decomposition applies to $\hat{\blacksquare}_1^{-1}$ and $H_1 \xrightarrow{L} +\infty$. By the same way, we also get $T_l \xrightarrow{L} +\infty$ for $l = 1, 2, 3, 4$

and $RH \xrightarrow{L} +\infty$.

Now, suppose that $\Pi_2^0 a = 0$, *i.e.* $a = 0$, because $\text{rank}(\Pi_2^0) = G$. This entails that $\delta = 0$ (remember that $a = \Sigma_V^{-1}\delta$). So, the null hypothesis of exogeneity is satisfied and all test statistics converge to non degenerate random variables as given in Theorem **3.1**.

Overall, the tests $H_i$, $T_l$ and $RH$ are consistent if and only if $\Pi_2^0 a \neq 0$.

(B) Suppose now that $\Pi_2 = \Pi_2^0$ with $\text{rank}(\Pi_2^0) \leq G$. We shall only focus on $\mathscr{H}_3$. The proof is similar for the other statistics. Let write $\mathscr{H}_3$ as

$$\begin{aligned} \mathscr{H}_3 &= T[(\tilde{\beta}-\beta)-(\hat{\beta}-\beta)]'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}[(\tilde{\beta}-\beta)-(\hat{\beta}-\beta)]/\hat{\sigma}^2 \\ &= T(\tilde{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\tilde{\beta}-\beta)/\hat{\sigma}^2 + T(\hat{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\hat{\beta}-\beta)/\hat{\sigma}^2 \\ &\quad -2T(\hat{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\tilde{\beta}-\beta)/\hat{\sigma}^2. \end{aligned} \tag{A.58}$$

We now study the asymptotic behaviour of the three terms in (A.58). First, we have

$$\begin{aligned} \hat{\beta}-\beta &= (Y'M_1Y/T)^{-1}(Y'M_1u/T) \xrightarrow{p} (\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta, \\ \hat{\sigma}^2 &\xrightarrow{p} \bar{\sigma}_u^2 = \sigma_u^2 - \delta'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta, \end{aligned} \tag{A.59}$$

and from (A.35)-(A.40), we have

$$T(\tilde{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\tilde{\beta}-\beta)/\hat{\sigma}^2 \xrightarrow{L} \frac{1}{\bar{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B, \tag{A.60}$$

$$2(\hat{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\tilde{\beta}-\beta)/\hat{\sigma}^2 \xrightarrow{L} \frac{1}{\bar{\sigma}_u^2}\delta'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\mathscr{A}\mathscr{S}_2\mathscr{N}_B. \tag{A.61}$$

Moreover, using (A.59) and the equality

$$\begin{aligned} (\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1} &= \hat{\Omega}_{IV}(\hat{\Omega}_{LS}-\hat{\Omega}_{IV})^{-1}\hat{\Omega}_{LS} \\ &\xrightarrow{p} (\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0)\Sigma_V^{-1}(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V), \end{aligned} \tag{A.62}$$

if $\Pi_2^0 a \neq 0$, we get

$$\begin{aligned} (\hat{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\hat{\beta}-\beta)/\hat{\sigma}^2 &\xrightarrow{p} \delta'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0)\Sigma_V^{-1}\delta/\bar{\sigma}_u^2 \\ &= a[(\Pi_2^{0\prime}\Sigma_Z\Pi_2^0)\Sigma_V^{-1}+I_G]^{-1}\Pi_2^{0\prime}\Sigma_Z\Pi_2^0 a/\bar{\sigma}_u^2, \end{aligned} \tag{A.63}$$

However, we have

$$a[(\Pi_2^{0\prime}\Sigma_Z\Pi_2^0)\Sigma_V^{-1}+I_G]^{-1}\Pi_2^{0\prime}\Sigma_Z\Pi_2^0 a/\bar{\sigma}_u^2 > 0$$

if and only $a \notin Ker\{[(\Pi_2^{0\prime}\Sigma_Z\Pi_2^0)\Sigma_V^{-1}+I_G]^{-1}\Pi_2^{0\prime}\Sigma_Z\Pi_2^0\}$, $Ker(L)$ denotes the null space spanned by the columns of the matrix $L$. Because $[(\Pi_2^{0\prime}\Sigma_Z\Pi_2^0)\Sigma_V^{-1}+I_G]^{-1}$ and $\Sigma_Z$ are nonsingular, we then have

$$Ker\{[(\Pi_2^{0\prime}\Sigma_Z\Pi_2^0)\Sigma_V^{-1}+I_G]^{-1}\Pi_2^{0\prime}\Sigma_Z\Pi_2^0\} = Ker(\Pi_2^{0\prime}\Sigma_Z\Pi_2^0) = Ker(\Pi_2^0).$$

So, the last term in (A.63) is positive if and only if $a \notin Ker(\Pi_2^0)$, *i.e.*, $\Pi_2^0 a \neq 0$. In this case, we get

$$T(\hat{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\hat{\beta}-\beta)/\hat{\sigma}^2 \xrightarrow{p} +\infty, \tag{A.64}$$

which entails that $\mathscr{H}_3 \xrightarrow{L} +\infty$.

Suppose now that $\Pi_2^0 a = 0$. We have

$$T(\hat{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\hat{\beta}-\beta)/\hat{\sigma}^2 = (\hat{\beta}-\beta)'[(T\hat{\Omega}_{IV})^{-1}-\frac{1}{T}\hat{\Omega}_{LS}^{-1})^{-1}]^{-1}(\hat{\beta}-\beta). \tag{A.65}$$

(A.66)

Since $\hat{\beta}-\beta \xrightarrow{p} (\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta$ and $[(T\hat{\Omega}_{IV})^{-1}-\frac{1}{T}\hat{\Omega}_{LS}^{-1})^{-1}]^{-1} \xrightarrow{L} \mathscr{A}$, where $\mathscr{A} = (\mathscr{S}_2\bar{P}_{22}\mathscr{S}_2')^{-1} = \mathscr{S}_2\mathscr{B}^{-1}\mathscr{S}_2'$, we get

$$\begin{aligned} T(\hat{\beta}-\beta)'(\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1})^{-1}(\hat{\beta}-\beta)/\hat{\sigma}^2 \quad &\xrightarrow{L} \quad \frac{1}{\bar{\sigma}_u^2}\delta'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\mathscr{A}\times \\ &\quad\quad (\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta. \end{aligned} \tag{A.67}$$

Thus from (A.60)-(A.61), we find

$$\begin{aligned} \mathscr{H}_3 \quad \xrightarrow{L} \quad & \frac{1}{\bar{\sigma}_u^2}\delta'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\mathscr{A}(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta+\frac{1}{\bar{\sigma}_u^2}\mathscr{N}_B'\mathscr{S}_2'\mathscr{A}\mathscr{S}_2\mathscr{N}_B \\ & -\frac{2}{\bar{\sigma}_u^2}\delta'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\mathscr{A}\mathscr{S}_2\mathscr{N}_B. \end{aligned} \tag{A.68}$$

With a little manipulation, we get

$$\mathscr{H}_3 \quad \xrightarrow{L} \quad \frac{1}{\bar{\sigma}_u^2}\left[\mathscr{S}_2\mathscr{N}_B-(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta\right]'\mathscr{A}\left[\mathscr{S}_2\mathscr{N}_B-(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta\right]. \tag{A.69}$$

From (A.38), we have

$$\begin{aligned} \mathscr{S}_2\mathscr{N}_B|_{\bar{S}_{2V}\mathscr{S}_2} \quad &\sim \quad \mathrm{N}[a,\, \sigma_\varepsilon^2\mathscr{S}_2\mathscr{B}^{-1}\mathscr{S}_2'] \equiv \mathrm{N}[\Sigma_V^{-1}\delta,\, \sigma_\varepsilon^2\mathscr{S}_2\mathscr{B}^{-1}\mathscr{S}_2'], \\ \mathscr{S}_2\mathscr{B}^{-1}\mathscr{S}_2' \quad &= \quad \left\{\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1}-\Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2V}\right\}^{-1} = \mathscr{A}^{-1}. \end{aligned} \tag{A.70}$$

It follows that

$$\mathscr{H}_3 \quad \xrightarrow{L} \quad \frac{\sigma_\varepsilon^2}{\bar{\sigma}_u^2}\chi^2(G;\mu_A), \tag{A.71}$$

where $\mu_A = \frac{1}{\sigma_\varepsilon^2}\delta'[\Sigma_V^{-1} - (\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}]\mathscr{A}^{-1}[\Sigma_V^{-1} - (\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}]\delta$. From Doko and Dufour (2010, Lemma A.1), we have

$$\Sigma_V^{-1} - (\Pi_2^{0'}\Sigma_Z\Pi_2^0 + \Sigma_V)^{-1} = (\Pi_2^{0'}\Sigma_Z\Pi_2^0 + \Sigma_V)^{-1}(\Pi_2^{0'}\Sigma_Z\Pi_2^0)\Sigma_V^{-1}. \tag{A.72}$$

Since $\Pi_2^0 a = 0$, this entails that

$$\begin{aligned}
\mu_A &= \frac{1}{\sigma_\varepsilon^2}\delta'(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0)\Sigma_V^{-1}\mathscr{A}^{-1}(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\times \\
&\quad \Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 a = 0, \\
\bar{\sigma}_u^2 &= \sigma_u^2 - \delta'(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\delta = \sigma_\varepsilon^2 + \delta'[\Sigma_V^{-1} - (\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}]\delta \\
&= \sigma_\varepsilon^2 + \delta'(\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V)^{-1}\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 a = \sigma_\varepsilon^2,
\end{aligned}$$

where $a = \Sigma_V^{-1}\delta$ and $\sigma_u^2 = \sigma_\varepsilon^2 + \delta'\Sigma_V^{-1}\delta$. Hence $\mathscr{H}_3 \xrightarrow{L} \chi^2(G)$. And $\mathscr{H}_3$ is not consistent. A similar result holds for the other statistics.

Overall, exogeneity tests are consistent if and only if $\Pi_2^0 a \neq 0$.

□

PROOF OF THEOREM **3.3** For any $a \neq 0$, we have $\text{rank}(\Pi_2) = G$ if and only if $\Pi_2 a \neq 0$ if and only if DWH and RH tests are consistent.

□

PROOF OF COROLLARY **3.4** The proof follows directly from those of Theorem **3.1** and Theorem **3.2**. □

PROOF OF THEOREM **3.5** Suppose that $\delta = \delta_0/\sqrt{T}$ and $\Pi_2 = \Pi_2^0$ is fixed. (A) If $\text{rank}(\Pi_2^0) = G$, From (A.1)-(A.8), we have

$$\hat{\Omega}_{IV} \xrightarrow{p} \Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0,\ \hat{\Omega}_{LS} \xrightarrow{p} \Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0 + \Sigma_V\,, \tag{A.73}$$

$$\frac{Y'M_1u}{T} \xrightarrow{p} 0,\ \frac{Y'(M_1 - M)u}{T} \xrightarrow{p} 0\,, \tag{A.74}$$

$$\hat{\sigma}^2 = u'u/T - (u'M_1Y/T)\hat{\Omega}_{LS}^{-1}(Y'M_1u/T) \xrightarrow{p} \sigma_u^2\,, \tag{A.75}$$

$$\tilde{\sigma}^2 = u'u/T - 2(u'M_1Y/T)\hat{\Omega}_{IV}^{-1}(Y'(M_1 - M)u/T) + \tag{A.76}$$

$$+(u'(M_1 - M)Y/T)\hat{\Omega}_{IV}^{-1}(Y'(M_1 - M)u/T) \xrightarrow{p} \sigma_u^2\,. \tag{A.77}$$

$$\frac{Y'M_1u}{\sqrt{T}} \xrightarrow{L} \Pi_2^{0'}\bar{S}_{2u} + S_{V\varepsilon} + \delta_0,\ \frac{Y'(M_1 - M)u}{\sqrt{T}} \xrightarrow{L} \Pi_2^{0'}\bar{S}_{2u}. \tag{A.78}$$

Since we have

$$\mathscr{H}_i = \sqrt{T}(\tilde{\beta} - \hat{\beta})'\hat{\blacksquare}_i^{-1}\sqrt{T}(\tilde{\beta} - \hat{\beta}), \tag{A.79}$$

and

$$
\begin{aligned}
\sqrt{T}(\tilde{\beta}-\hat{\beta}) &= \hat{\Omega}_{LS}^{-1}(Y'M_1u/\sqrt{T}) - \hat{\Omega}_{IV}^{-1}(Y'(M_1-M)u/\sqrt{T}) \\
&\xrightarrow{L} (\Pi_2^{0\prime}\Sigma_Z\Pi_2^0+\Sigma_V)^{-1}(\Pi_2^{0\prime}\bar{S}_{2u}+S_{V\varepsilon}+\delta_0) - (\Pi_2^{0\prime}\Sigma_Z\Pi_2^0)^{-1}\Pi_2^{0\prime}\bar{S}_{2u}\,, \qquad (A.80)
\end{aligned}
$$

$$
\hat{\blacksquare}_i \xrightarrow{p} \sigma_u^2\Delta_\Pi,\, \Delta_\Pi = (\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0)^{-1} - (\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}, \quad i=1,2,3. \qquad (A.81)
$$

So, following (A.13)-(A.15), we find

$$
\mathscr{H}_i \xrightarrow{L} \chi^2(G,\mu_{\delta_0}), \qquad (A.82)
$$

where $\mu_{\delta_0} = \frac{1}{\sigma_u^2}\delta_0'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\Delta_\Pi^{-1}(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\delta_0$. By the same way, we can show that

$$
\mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G,\mu_{\delta_0}),\ \mathscr{T}_l \xrightarrow{L} \chi^2(G,\mu_{\delta_0}),\ l=3,4,\ \mathscr{T}_1 \xrightarrow{L} F(G,k_2-G,\mu_{\delta_0}). \qquad (A.83)
$$

For the statistic $\mathscr{RH}$, its denominator $\frac{1}{T}u'M_{\bar{X}}u$ converges to $\sigma_u^2$. Its numerator is

$$
\frac{1}{k_2}u'(M_{X_1}-M_{\bar{X}})u = \frac{1}{k_2}\frac{u'M_{X_1}\bar{Z}_2}{\sqrt{T}}\left(\frac{\bar{Z}_2'M_{X_1}\bar{Z}_2}{T}\right)^{-1}\frac{\bar{Z}_2'M_{X_1}u}{\sqrt{T}}\,. \qquad (A.84)
$$

Moreover, we have

$$
\begin{aligned}
\frac{\bar{Z}_2'M_{X_1}\bar{Z}_2}{T} &\xrightarrow{p} \Sigma_{\bar{Z}_2}^{*} = \Sigma_{\bar{Z}_2} - \Sigma_{\bar{Z}_2}\Pi_2^0(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}^{-1}\Pi_2^0+\Sigma_V)^{-1}\Pi_2^{0\prime}\Sigma_{\bar{Z}_2} \\
\frac{\bar{Z}_2'M_{X_1}u}{\sqrt{T}} &= \frac{\bar{Z}_2'u}{\sqrt{T}} - \frac{\bar{Z}_2'M_1Y}{T}\left(\frac{Y'M_1Y}{T}\right)^{-1}\frac{Y'M_1u}{\sqrt{T}} \\
&\xrightarrow{L} \bar{S}_{2u} - \Sigma_{\bar{Z}_2}\Pi_2^0(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}^{-1}\Pi_2^0+\Sigma_V)^{-1}(\Pi_2^{0\prime}\bar{S}_{2u}+S_{V\varepsilon}+\delta_0) \\
&\sim N[-\Sigma_{\bar{Z}_2}\Pi_2^0(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}^{-1}\Pi_2^0+\Sigma_V)^{-1}\delta_0,\, \sigma_u^2\Sigma_{\bar{Z}_2}^{*}].
\end{aligned}
$$

Thus,

$$
\mathscr{RH} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2,\nu_{\delta_0})\,, \qquad (A.85)
$$

where $\nu_{\delta_0} = \frac{1}{\sigma_u^2}\delta_0'(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Sigma_{\bar{Z}_2}^{*-1}\Sigma_{\bar{Z}_2}\Pi_2^0(\Pi_2^{0\prime}\Sigma_{\bar{Z}_2}\Pi_2^0+\Sigma_V)^{-1}$.

(B) Suppose now that $\text{rank}(\Pi_2^0) \le G$. Since $\delta \to 0$ as $T \to +\infty$, we can observe that equations (A.22)-(A.42) still hold so that we have

$$
\mathscr{H}_3 \xrightarrow{L} \chi^2(G). \qquad (A.86)
$$

By proceeding as in Theorem **3.1**, we get the results for the other statistics.

□

PROOF OF THEOREM **4.1** Assume that $\delta = 0$. Under the assumptions of the model and if $\Pi_2 = \Pi_2^0/\sqrt{T}$ where $\Pi_2^0$ is a $k \times G$ constant matrix ($\Pi_2^0 = 0$ is allowed), then we have

$$T\hat{\Omega}_{IV} \xrightarrow{L} \Psi_V = (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V}),\, \hat{\Omega}_{LS} \xrightarrow{p} \Sigma_V\,, \tag{A.87}$$

$$\frac{1}{T}Y'M_1 u \xrightarrow{p} \delta = 0,\, Y'(M_1 - M)u \xrightarrow{L} (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}\,, \tag{A.88}$$

$$\hat{\sigma}^2 = u'u/T - (u'M_1Y/T)\hat{\Omega}_{LS}^{-1}(Y'M_1u/T) \xrightarrow{p} \sigma_u^2 - \delta'\Sigma_V^{-1}\delta = \sigma_u^2\,, \tag{A.89}$$

$$\tilde{\beta} - \beta = (T\hat{\Omega}_{IV})^{-1}Y'(M_1 - M)u \xrightarrow{p} \Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}\,. \tag{A.90}$$

Thus, we get

$$\tilde{\sigma}^2 = \frac{u'u}{T} - 2\frac{u'M_1Y}{T}(\tilde{\beta} - \beta) + (\tilde{\beta} - \beta)'\hat{\Omega}_{LS}(\tilde{\beta} - \beta) \xrightarrow{L} \bar{\sigma}_u^2,\, \hat{\Sigma}_i \xrightarrow{L} \frac{1}{\bar{\sigma}_u^2}\Psi_V^{-1},\, i = 1,\, 2,$$

where $\bar{\sigma}_u^2 = \sigma_u^2 + \bar{S}_{2u}'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})\Psi_V^{-1}\Sigma_V\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u},$ so that

$$\mathscr{H}_i \xrightarrow{L} \frac{1}{\bar{\sigma}_u^2}\bar{S}_{2u}'\Sigma_A\bar{S}_{2u},\, i = 1,\, 2\,,$$

where $\Sigma_A = \Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}$. Since $\bar{\sigma}_u^2 \geq \sigma_u^2,$ we have

$$\mathscr{H}_i \leq \frac{1}{\sigma_u^2}\bar{S}_{2u}'\Sigma_A\bar{S}_{2u},\, i = 1,\, 2.$$

Because $\bar{S}_{2u}$ and $\bar{S}_{2V}$ are independent when $\delta = 0$, it follows that

$$\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}|_{\bar{S}_{2V}} \sim N(0,\, \sigma_u^2\Psi_V)\,, \tag{A.91}$$

where $\Psi_V$ is defined in (A.87). Hence,

$$\frac{1}{\sigma_u^2}\bar{S}_{2u}'\Sigma_A\bar{S}_{2u}|_{\bar{S}_{2V}} \sim \chi^2(G)\,,$$

and $H_i \leq \chi^2(G),\; i = 1, 2.$ Furthermore, $\hat{\Sigma}_3 \xrightarrow{p} \frac{1}{\sigma_u^2}\Psi_V^{-1},$ which entails that $H_3|_{\bar{S}_{2V}} \xrightarrow{L} \chi^2(G),$ *i.e.* $H_3 \xrightarrow{L} \chi^2(G)$. By the same way, we can also show that

$$\mathscr{T}_1 \xrightarrow{L} F(G, k_2 - G), \quad \mathscr{T}_2 \xrightarrow{L} \frac{1}{G}\chi^2(G), \quad \mathscr{T}_4 \xrightarrow{L} \chi^2(G)\,, \tag{A.92}$$

$$\mathscr{T}_3 \xrightarrow{L} \frac{1}{\bar{\sigma}_u^2}\bar{S}_{2u}'\Sigma_A\bar{S}_{2u} \leq \chi^2(G), \quad \text{and} \quad \mathscr{R}\mathscr{H} \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2)\,. \tag{A.93}$$

□

PROOF OF THEOREM **4.2** Suppose that $\Pi_2 = \Pi_2^0/\sqrt{T}$ where $\Pi_2^0 = 0$ is allowed. Then we have

$$T\hat{\Omega}_{IV} \xrightarrow{L} \Psi_V = (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V}),\, \hat{\Omega}_{LS} \xrightarrow{p} \Sigma_V\,, \tag{A.94}$$

$$\frac{1}{T}Y'M_1u \xrightarrow{p} \delta \neq 0,\, Y'(M_1 - M)u \xrightarrow{L} (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}\,, \tag{A.95}$$

$$\hat{\sigma}^2 = u'u/T - (u'M_1Y/T)\hat{\Omega}_{LS}^{-1}(Y'M_1u/T) \xrightarrow{p} \sigma_u^2 - \delta'\Sigma_V^{-1}\delta = \sigma_\varepsilon^2\,, \tag{A.96}$$

$$\tilde{\beta} - \beta \;=\; (T\hat{\Omega}_{IV})^{-1}Y'(M_1 - M)u \xrightarrow{p} \Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}\,. \tag{A.97}$$

$$\begin{aligned}
\tilde{\sigma}^2 &= \frac{u'u}{T} - 2\frac{u'M_1Y}{T}(\tilde{\beta} - \beta) + (\tilde{\beta} - \beta)'\hat{\Omega}_{LS}(\tilde{\beta} - \beta) \\
&\xrightarrow{L} \tilde{\sigma}_*^2,\, \frac{1}{T}\hat{\blacksquare}_i \xrightarrow{L} \sigma_{i*}^2\Psi_V^{-1},\, i = 1,\,2,\,3
\end{aligned} \tag{A.98}$$

where $\sigma_{1*}^2 = \sigma_{2*}^2 = \tilde{\sigma}_*^2,\ \sigma_{3*}^2 = \sigma_\varepsilon^2$ and

$$\begin{aligned}
\tilde{\sigma}_*^2 &= \sigma_u^2 - 2\delta'\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u} \\
&\quad + \bar{S}_{2u}'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})\Psi_V^{-1}\Sigma_V\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\tilde{\beta} - \hat{\beta} &= \hat{\Omega}_{LS}^{-1}(u'M_1Y/T) - (T\hat{\Omega}_{IV})^{-1}Y'(M_1 - M)u \\
&\xrightarrow{L} \Sigma_V^{-1}\delta - \Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u}\,.
\end{aligned} \tag{A.99}$$

By noting that $\bar{S}_{2u} = \bar{S}_{2V}a + \bar{S}_{2\varepsilon} = \bar{S}_{2V}\Sigma_V^{-1}\delta + \bar{S}_{2\varepsilon}$, we easily get

$$\begin{aligned}
\tilde{\beta} - \hat{\beta} &\xrightarrow{L} \Sigma_V^{-1}\delta - \Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}\Sigma_V^{-1}\delta - \Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon} \\
&= \Psi_V^{-1}[\Lambda_V a - (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon}].
\end{aligned}$$

where $\Lambda_V = \Psi_V - (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V} = (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Pi_2^0$ and $a = \Sigma_V^{-1}\delta$. So,

$$\mathscr{H}_i \xrightarrow{L} \frac{1}{\sigma_{i*}^2}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon}),\, i = 1,\,2,\,3\,,$$

where $\sigma_{1*}^2 = \sigma_{2*}^2 = \tilde{\sigma}_*^2,\ \sigma_{3*}^2 = \sigma_\varepsilon^2,$ and $\Delta_V = (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'$. Moreover, $\bar{S}_{2\varepsilon} \sim N(0, \sigma_\varepsilon^2\Sigma_{\bar{Z}_2})$ and $\bar{S}_{2\varepsilon}$ is independent with $\bar{S}_{2V}$ when $\delta = 0$. Thus

$$\mathscr{H}_3|\bar{S}_{2V} \xrightarrow{L} \chi^2(G, \mu_V),\ \mu_V = \frac{1}{\sigma_\varepsilon^2}a'\Pi_2^{0'}\Delta_V\Pi_2^0 a\,. \tag{A.100}$$

Since $T_3 = (\kappa_3/T)H_2,\ T_3 = (\kappa_4/T)H_3$ and $\kappa_3/T = \kappa_4/T = (T - G)/T \to 1$ as $T \to +\infty,$ it follows that

$$\mathscr{T}_3 \xrightarrow{L} \frac{1}{\tilde{\sigma}_*^2}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon}),$$

$$\mathscr{T}_4 \xrightarrow{L} \frac{1}{\sigma_\varepsilon^2}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon}). \tag{A.101}$$

By conditioning on $\bar{S}_{2V}$, we get

$$\mathscr{T}_4|\bar{S}_{2V} \xrightarrow{L} \chi^2(G, \mu_V). \tag{A.102}$$

Moreover, by noting that $\operatorname*{plim}_{T\to\infty}(\tilde{\sigma}_2^2) = \operatorname*{plim}_{T\to\infty}(\hat{\sigma}^2) = \sigma_\varepsilon^2$, we also find

$$\mathscr{T}_2 \xrightarrow{L} \frac{1}{\sigma_\varepsilon^2 G}(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon}) \text{ and } \mathscr{T}_2|\bar{S}_{2V} \xrightarrow{L} \frac{1}{G}\chi^2(G, \mu_V).$$

Furthermore, we can see that

$$T\tilde{\sigma}_1^2 = u'(M_1 - M)M_{\hat{Y}}(M_1 - M)u \xrightarrow{L} \bar{S}_{2u}'(\Sigma_{\bar{Z}_2}^{-1} - \Sigma_{\bar{Z}_2}^{-1}\Delta_V\Sigma_{\bar{Z}_2}^{-1})\bar{S}_{2u}, \tag{A.103}$$

where the limit term in (A.103) can be written as

$$\bar{S}_{2u}'(\Sigma_{\bar{Z}_2}^{-1} - \Sigma_{\bar{Z}_2}^{-1}\Delta_V\Sigma_{\bar{Z}_2}^{-1})\bar{S}_{2u} = (\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a + \Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2\varepsilon})'\Delta_V^*(\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a + \Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2\varepsilon}),$$

where $\Delta_V^* = I_{k_2} - \Sigma_{\bar{Z}_2}^{-1/2}\Delta_V\Sigma_{\bar{Z}_2}^{-1/2}$ is symmetric idempotent with rank $k_2 - G$. So, we have $T\tilde{\sigma}_1^2|\bar{S}_{2V} \xrightarrow{L} \sigma_\varepsilon^2\chi^2(k_2 - G, \lambda_V)$, where

$$\lambda_V = \frac{1}{\sigma_\varepsilon^2}a'\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1/2}\Delta_V^*\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a = \frac{1}{\sigma_\varepsilon^2}a'\bar{S}_{2V}'(\Sigma_{\bar{Z}_2}^{-1} - \Sigma_{\bar{Z}_2}^{-1}\Delta_V\Sigma_{\bar{Z}_2}^{-1})\bar{S}_{2V}a.$$

Further, we have $\Delta_V(\Sigma_{\bar{Z}_2}^{-1/2}\Delta_V^*\Sigma_{\bar{Z}_2}^{-1/2}) = \Delta_V\Sigma_{\bar{Z}_2}^{-1} - \Delta_V\Sigma_{\bar{Z}_2}^{-1}\Delta_V\Sigma_{\bar{Z}_2}^{-1}$ and since $\Delta_V\Sigma_{\bar{Z}_2}^{-1}\Delta_V = \Delta_V$, it follows that $\Delta_V(\Sigma_{\bar{Z}_2}^{-1/2}\Delta_V^*\Sigma_{\bar{Z}_2}^{-1/2}) = 0$. So, conditionally on $\bar{S}_{2V}$, the quadratic forms

$$(\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a + \Sigma_{\bar{Z}_2}^{-1/2}S_{2\varepsilon})'\Delta_V^*(\Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2V}a + \Sigma_{\bar{Z}_2}^{-1/2}\bar{S}_{2\varepsilon}) \text{ and } (\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})'\Delta_V(\Pi_2^0 a - \Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon})$$

are independent and distributed as noncentral chi-squares. Thus

$$\mathscr{T}_1|\bar{S}_{2V} \xrightarrow{L} F(G, k_2 - G; \mu_V, \lambda_V). \tag{A.104}$$

For the statistic $RH$, the denominator is

$$\frac{1}{T}u'M_{\bar{X}}u = \frac{1}{T}(u'M_{X_1}u - u'M_{X_1}\bar{Z}_2(\bar{Z}_2'M_{X_1}\bar{Z}_2)^{-1}\bar{Z}_2'M_{X_1}u), \tag{A.105}$$

where

$$\begin{aligned}
\frac{1}{T}u'M_{X_1}u &= \frac{1}{T}u'M_1u - \frac{1}{T}u'M_1Y(Y'M_1Y)^{-1}Y'M_1u \xrightarrow{p} \sigma_u^2 - \delta'\Sigma_V^{-1}\delta = \sigma_\varepsilon^2, \\
\text{and} &\quad \frac{1}{T}(u'M_{X_1}\bar{Z}_2(\bar{Z}_2'M_{X_1}\bar{Z}_2)^{-1}\bar{Z}_2'M_{X_1}u) \xrightarrow{p} 0
\end{aligned}$$

under $\delta = 0$. So, we find $\frac{1}{T}u'M_{\bar{X}}u \xrightarrow{p} \sigma_\varepsilon^2$. For the numerator, we have

$$\frac{1}{k_2}u'(M_{X_1} - M_{\bar{X}})u = \frac{1}{k_2}\frac{u'M_{X_1}\bar{Z}_2}{\sqrt{T}}\left(\frac{\bar{Z}_2'M_{X_1}\bar{Z}_2}{T}\right)^{-1}\frac{\bar{Z}_2'M_{X_1}u}{\sqrt{T}}. \tag{A.106}$$

Moreover, $\frac{\bar{Z}_2'M_{X_1}\bar{Z}_2}{T} = \frac{\bar{Z}_2'M_1\bar{Z}_2}{T} - \frac{\bar{Z}_2'M_1Y}{T}\left(\frac{Y'M_1Y}{T}\right)^{-1}\frac{Y'M_1\bar{Z}_2}{T} \xrightarrow{p} \Sigma_{\bar{Z}_2}$ because $\frac{Y'M_1\bar{Z}_2}{T} \xrightarrow{p} 0$. Now, we have

$$\frac{\bar{Z}_2'M_{X_1}u}{\sqrt{T}} = \frac{\bar{Z}_2'u}{\sqrt{T}} - \frac{\bar{Z}_2'M_1Y}{\sqrt{T}}\left(\frac{Y'M_1Y}{T}\right)^{-1}\frac{Y'M_1u}{T},$$

where $\frac{\bar{Z}_2'u}{\sqrt{T}} = \frac{\bar{Z}_2'V}{\sqrt{T}}\Sigma_V^{-1}\delta + \frac{\bar{Z}_2'\varepsilon}{\sqrt{T}} \xrightarrow{L} \bar{S}_{2\varepsilon} + \bar{S}_{2V}\Sigma_V^{-1}\delta$, $\left(\frac{Y'M_1Y}{T}\right)^{-1}\frac{Y'M_1u}{T} \xrightarrow{p} \Sigma_V^{-1}\delta$ and $\frac{\bar{Z}_2'M_1Y}{\sqrt{T}} \xrightarrow{L} \Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V}$. Hence we have

$$\frac{1}{k_2}u'(M_{X_1} - M_{\bar{X}})u \xrightarrow{L} \frac{1}{k_2}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0a)'\Sigma_{\bar{Z}_2}^{-1}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0a),$$

thus

$$\begin{aligned}
\mathscr{R}\mathscr{H} &\xrightarrow{L} \frac{1}{k_2\sigma_\varepsilon^2}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0a)'\Sigma_{\bar{Z}_2}^{-1}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0a) \sim \frac{1}{k_2}\chi^2(k_2, \mu_R), \\
\mu_R &= a'\Pi_2^{0'}\Sigma_{\bar{Z}_2}\Pi_2^0a.
\end{aligned}$$

□

Proof of Theorem **4.3** Let $\Pi_2^0a = 0$ in the proof of Theorem **4.2** above. Then, we have $\mu_V = \lambda_V = \mu_R = 0$. Further, we can observe that

$$\begin{aligned}
\tilde{\sigma}_*^2 = \sigma_{0*}^2 &= \sigma_\varepsilon^2 + \bar{S}_{2\varepsilon}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V}(\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V})^{-1}\Sigma_V(\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2V})^{-1}\bar{S}_{2V}'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2\varepsilon} && \text{(A.107)} \\
&\geq \sigma_\varepsilon^2 && \text{(A.108)}
\end{aligned}$$

and the matrix $\Sigma_{\bar{Z}_2}^{-1} - \Sigma_{\bar{Z}_2}^{-1}\Delta_V\Sigma_{\bar{Z}_2}^{-1}$ is positive semi-definite, *i.e.*

$$\Sigma_{\bar{Z}_2}^{-1} - \Sigma_{\bar{Z}_2}^{-1}\Delta_V\Sigma_{\bar{Z}_2}^{-1} = \Sigma_{\bar{Z}_2}^{-\frac{1}{2}}(I_{k_2} - \Sigma_{\bar{Z}_2}^{-\frac{1}{2}}\Delta_V\Sigma_{\bar{Z}_2}^{-\frac{1}{2}})\Sigma_{\bar{Z}_2}^{-\frac{1}{2}} \geq 0, \tag{A.109}$$

where $I_{k_2} - \Sigma_{\bar{Z}_2}^{-\frac{1}{2}}\Delta_V\Sigma_{\bar{Z}_2}^{-\frac{1}{2}}$ is idempotent of rank $k_2 - G$. Then, the results of Thoerem **4.3** follow. □

PROOF OF LEMMA **5.1** Assume that $\Pi_2$ is fixed. We have

$$\begin{aligned}\hat{\beta} &= (Y'M_1Y/T)^{-1}(Y'M_1y/T) = \beta + (Y'M_1Y/T)^{-1}(Y'M_1u/T) \qquad (A.110)\\ \tilde{\beta} &= \mathscr{S}_1\tilde{\beta}_1 + \mathscr{S}_2\tilde{\beta}_2\\ &= \beta + \mathscr{S}_1(Y_1'EY_1/T)^{-1}(Y_1'Eu/T) + \mathscr{S}_2(Y_2'JY_2)^{-1}(Y_2'Ju), \qquad (A.111)\end{aligned}$$

where $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are defined in (3.11). Since $Y'M_1Y/T \xrightarrow{p} \Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 + \Sigma_V$ and $Y'M_1u/T \xrightarrow{p} \delta$, we have

$$\hat{\beta} - \beta \xrightarrow{p} (\Pi_2'\Sigma_{\bar{Z}_2}\Pi_2 + \Sigma_V)^{-1}\delta \qquad (A.112)$$

irrespective of whether $\mathrm{rank}(\Pi_2) = G$ or not. We now focus on $\tilde{\beta}$. From (A.29)-(A.37), We have

$$Y_1'EY_1/T \xrightarrow{p} \Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21},\, Y_1'Eu/T \xrightarrow{p} 0, \qquad (A.113)$$

and we have

$$\tilde{\beta}_1 - \beta_1 \xrightarrow{p} 0, \qquad (A.114)$$

irrespective of whether $\mathrm{rank}(\Pi_2) = G$ or not. For $\tilde{\beta}_2$, we distinguish 3 cases:
(i) if $\mathrm{rank}(\Pi_2) = G,$ $\beta$ reduces to $\beta_1$ and we have $\tilde{\beta} \xrightarrow{p} \beta$.
(ii) if $\mathrm{rank}(\Pi_2) < G,$ (A.113) still holds and we have $\mathscr{S}_1(Y_1'EY_1/T)^{-1}(Y_1'Eu/T) \xrightarrow{p} 0$. Furthermore, from (A.29)-(A.37), we have

$$\mathscr{S}_2(Y_2'JY_2)^{-1}(Y_2'Ju) \xrightarrow{L} \mathscr{S}_2\mathscr{N}_B,\ \tilde{\beta}_2 \xrightarrow{L} \beta_2 + \mathscr{N}_B, \qquad (A.115)$$

where $\mathscr{N}_B$ is given by (A.37). From (A.111), we get

$$\tilde{\beta} \xrightarrow{p} \beta + \mathscr{S}_2\mathscr{N}_B. \qquad (A.116)$$

$\tilde{\beta}_1$ is always consistent even if identification is deficient while $\tilde{\beta}_2$ is consistent only when identification is strong.

By putting (i)-(ii) together, we have

$$\tilde{\beta} \xrightarrow{p} \beta + \mathscr{S}_2\mathscr{N}_B, \qquad (A.117)$$

where

$$\mathscr{S}_2\mathscr{N}_B = \begin{cases} 0 \text{ if } \mathrm{rank}(\Pi_2) = G, \\ \\ \mathscr{S}_2\mathscr{B}^{-1}\mathscr{S}_2'\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2u} \\ \text{if } \mathrm{rank}(\Pi_2) < G\ , \end{cases} \qquad (A.118)$$

where from (A.70), we have

$$\mathscr{S}_2\mathscr{N}_B|_{\bar{S}_{2V}\mathscr{S}_2} \sim \mathrm{N}[\Sigma_V^{-1}\delta,\, \sigma_\varepsilon^2\left\{\bar{S}_{2V}'[\Sigma_{\bar{Z}_2}^{-1} - \Pi_{21}(\Pi_{21}'\Sigma_{\bar{Z}_2}\Pi_{21})^{-1}\Pi_{21}']\bar{S}_{2V}\right\}^{-1}], \quad \text{(A.119)}$$

or equivalently

$$\mathscr{N}_B|_{\bar{S}_{2V}\mathscr{S}_2} \sim \mathrm{N}\left[\mathscr{S}_2'\Sigma_V^{-1}\delta,\, \sigma_\varepsilon^2\mathscr{B}^{-1}\right]. \quad \text{(A.120)}$$

□

PROOF OF LEMMA **5.2** Suppose that $\Pi_2 = \Pi_2^0/\sqrt{T}$ (asymptotically weak instruments). We have $Y'M_1Y/T \xrightarrow{p} \Sigma_V$ and $Y'M_1u/T \xrightarrow{p} \delta$. Hence, we have

$$\hat{\beta} \xrightarrow{p} \beta + \Sigma_V^{-1}\delta = \beta + a = \beta^*. \quad \text{(A.121)}$$

Now, we have

$$\begin{aligned}\tilde{\beta} &= [Y'(M_1 - M)Y]^{-1}Y'(M_1 - M)y \\ &= \beta + [Y'(M_1 - M)Y]^{-1}Y'(M_1 - M)u. \end{aligned} \quad \text{(A.122)}$$

Moreover, from (A.94)-(A.95), we have

$$\begin{aligned} Y'(M_1 - M)Y &= T\hat{\Omega}_{IV} \xrightarrow{L} \Psi_V, \\ \Psi_V &= (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V}), \end{aligned} \quad \text{(A.123)}$$

$$\begin{aligned} Y'(M_1 - M)u &\xrightarrow{L} (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\bar{S}_{2u} \\ &= \Psi_V a + (\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0 a). \end{aligned} \quad \text{(A.124)}$$

Thus

$$\tilde{\beta} - \beta^* \xrightarrow{p} \mathscr{N}_\Psi^W, \quad \text{(A.125)}$$

where $\mathscr{N}_\Psi^W = \Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}(\bar{S}_{2\varepsilon} - \Sigma_{\bar{Z}_2}\Pi_2^0 a)$, and $\mathscr{N}_\Psi^W|_{\bar{S}_{2V}} \sim \mathrm{N}[-\Psi_V^{-1}(\Sigma_{\bar{Z}_2}\Pi_2^0 + \bar{S}_{2V})'\Sigma_{\bar{Z}_2}^{-1}\Sigma_{\bar{Z}_2}\Pi_2^0 a,\, \sigma_\varepsilon^2\Psi_V^{-1}]$.

□

PROOF OF THEOREM **5.3** Theorem **5.3** follow from the definition of pre-test estimators given by (5.1) - (5.3) and the results of Lemma **5.1** and Lemma **5.2**. □

## References

Anderson, T. W., 2003. An Introduction to Multivariate Statistical Analysis third edn. John Wiley & Sons, New York.

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.

Angrist, J. D., al., 1999. Jackknife instrumental variables estimates. Journal of Applied Econometrics 14, 57–67.

Angrist, J. D. , Krueger, A. B. , 1991. Does compulsory school attendance affect schooling and earning?. Quarterly Journal of Economics CVI, 979–1014.

Angrist, J. D., Krueger, A. B., 1995. Split-sample instrumental variables estimates of the return to schooling. Journal of Business and Economic Statistics 13, 225–235.

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

Choi, I., Phillips, P. C. B., 1992. Asymptotic and finite sample distribution theory for IV estimators and tests in partially identified structural equations. Journal of Econometrics 51, 113–150.

Doko, 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, T. F., Dufour, J.-M., 2010. 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.

Donald, S. G., Newey, W. K., 2001. Choosing the number of instruments. Econometrica 69, 1161–1191.

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

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

Dufour, J.-M., Jasiak, J., 2001. Finite sample limited information inference methods for structural equations and models with generated regressors. International Economic Review 42, 815–843.

Dufour, J.-M., Taamouti, M., 2005. Projection-based statistical inference in linear structural models with possibly weak instruments. Econometrica 73(4), 1351–1365.

Dufour, J.-M., Taamouti, M., 2006. Further results on projection-based inference in IV regressions with weak, collinear or missing instruments. Journal of Econometrics 139(1), 133–153.

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

Frankel, J. A., Romer, D., 1999. Does trade cause growth?. American Economic Review 89(3), 379–399.

Guggenberger, P., 2008. The impact of a Hausman pretest on the size of the hypothesis tests. forthcoming. Econometric Theory .

Hahn, J., Ham, J., Moon, H. R., 2008. The Hausman test and weak instruments. Technical report, Department of Economics, UCLA Los Angeles, California.

Hall, A. R., Peixe, F. P. M., 2003. A consistent method for the selection of relevant instruments. Econometric Reviews 2(3), 269–287.

Hall, A. R., Rudebusch, G. D., Wilcox, D. W., 1996. Judging instrument relevance in instrumental variables estimation. International Economic Review 37, 283–298.

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

Harrison, A., 1996. Oponness and growth: a time-series, cross-country analysis for developing countries. 48, 419–447.

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

Irwin, A.-D., Tervio, M., 2002. Does trade raise income? evidence from twentieth century. Journal of International Economics 58, 1–18.

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.

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

Mankiw, G. N., al., 1992. A contribution to the empirics of economic growth. The Quarterly Journal of Economics 107(2), 407–437.

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

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.

Rodrik, D., 1995. Trade and industrial policy reform. Review of Economic Studies 17, 1–20.

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

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.

Tylavsky, D. , Sohie, G. , 1986. Generalization of the matrix inversion lemma, in Proc. IEEE. 74, 1050–1052.

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

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