SCHOOL OF ECONOMICS AND FINANCE

Discussion Paper 2012-02

Testing for Partial Exogeneity with Weak Identification

Firmin Doko Tchatoka

ISSN 1443-8593
ISBN 978-1-86295-667-4

Testing for Partial Exogeneity with Weak Identification *

Firmin Doko Tchatoka†

School of Economics and Finance

University of Tasmania

May 31, 2012

* The author thanks Peter C.B. Phillips, Mardi Dungey, Alistair Hall, Jean-Marie Dufour, Russell Davidson, Denise Osborn and Jan Jacobs for several useful comments.

† 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/

ABSTRACT

We consider the following problem. A structural equation of interest contains two sets of explanatory variables which economic theory predicts may be endogenous. The researcher is interesting in testing the exogeneity of only one of them. Standard exogeneity tests are in general unreliable from the view point of size control to assess such a problem. We develop four alternative tests to address this issue in a convenient way. We provide a characterization of their distributions under both the null hypothesis (level) and the alternative hypothesis (power), with or without identification. We show that the usual χ^2 critical values are still applicable even when identification is weak. So, all proposed tests can be described as robust to weak instruments. We also show that test consistency may still hold even if the overall identification fails, provided partial identification is satisfied. We present a Monte Carlo experiment which confirms our theory. We illustrate our theory with the widely considered returns to education example. The results underscore: $(\mathbf{1})$ how the use of standard tests to assess partial exogeneity hypotheses may be misleading, and $(\mathbf{2})$ the relevance of using our procedures when checking for partial exogeneity.

Key words: Subset of endogenous regressors; Generated structural equation; Robustness to weak identification; Consistency.

JEL classification: C3; C12; C15; C52.

1. Introduction

Inference methods using instrumental variables (IV) methods are mainly motivated by the fact that explanatory variables may be correlated with the error term, so ordinary least squares (OLS) yields biased and inconsistent estimators. It is well known that when explanatory variables are endogenous, OLS estimators measure only the magnitude of association, rather than the magnitude and direction of causation which is needed for policy analysis. IV estimation provides a way to nonetheless obtain consistent parameter estimates, once the effect of common driving variables has been eliminated. Usually, researchers need to pretest the exogeneity of the regressors to decide whether OLS or IV method is appropriate. In the linear IV regression, exogeneity tests of the type proposed by Durbin (1954); Wu (1973, 1974), Revankar and Hartley (1973), and Hausman (1978), henceforth DWHRH tests, are often used as pretests for exogeneity. Recent studies[1] have established that they never over reject the null hypothesis of exogeneity even when model parameters are weakly identified.

A drawback of DWHRH tests however is that the null hypothesis of interest is specified on the whole set of supposedly endogenous regressors. When more than one regressor is involved, these tests cannot pinpoint which regressor is endogenous and which is not, once joint exogeneity has been rejected. This is particularly problematic from the viewpoint of estimation, since efficiency requires to use available instruments only for the regressors which are endogenous. The use of instruments for exogenous regressors often yields inefficient estimates of model parameters. To avoid such situations, it is important to know which variables are endogenous and which are not before inference. In models involving more than one supposedly endogenous variable, as it is often the case in most empirical applications, it is important to find ways to assess the exogeneity of the regressors separately.

However, the literature has focused on testing hypotheses specified on the structural parameters and inference procedures that are robust to identification problems[2]. Although these robust procedures extend to hypotheses specified on subsets of structural parameters [Dufour and Taamouti (2005, 2007), Kleibergen (2004, 2005), and Guggenberger and Smith (2005)], not much is known about testing for partial exogeneity, especially when identification is weak.

In this paper we propose alternative tests for assessing partial exogeneity hypotheses in linear

[1] See for example, Staiger and Stock (1997), Guggenberger (2010), and Hahn, Ham and Moon (2010).
[2] Anderson and Rubin (1949, AR-test), Kleibergen (2002, KLM-test),Moreira (2003, MQLR-test).

simultaneous equations models. The proposed tests do not require the exogeneity of the regressors not being tested or strong instruments, so they can be described as identification-robust. To be more specific, we consider a model of the form

$$y \quad = \quad Y\beta + W\theta + u$$

where y is an observed dependent variable, Y and W are matrices of observed (possibly) endogenous regressors. We wish to test the exogeneity of $Y,$ i.e. the hypothesis $\mathbf{cov}(Y, u) = 0.$

First, we stress the fact that the regressors W whose exogeneity is not being tested can be orthogonalized through a methodology built on four steps. We refer to the transformed equation where W has been replaced by the orthogonalized regressors, $\tilde{W},$ as the *generated structural equation*. An interesting feature of this *generated structural equation* is the structural parameters of interest β and θ have the same interpretation as in the original model.

Second, we show that the exogeneity hypothesis of Y can be assessed by testing whether Y is uncorrelated with the error of this *generated structural equation*, though the latter error typically differs to the original structural one. We then follow Durbin (1954), Wu (1973), and Hausman (1978) in proposing four statistics based on the vector of contrasts between ordinary least squares (OLS) and instrumental variables (IV) estimators of β in the transformed model, upon scaling by appropriate factors to guarantee the usual asymptotic χ^2 distributions.

Finally, after formulating generic assumptions on model variables which allow one to characterize the behaviour of the tests under both the null hypothesis (level) and the alternative hypothesis (power), we consider two main setups. In the first setup, model parameters are strongly identified, i.e., the reduced form parameter matrix that characterizes the strength of the instruments has full rank. The second setup is Staiger and Stock's (1997) *local-to-zero weak instrument asymptotics*. In this setup, the parameter matrix that controls the strength of the instruments approaches zero at rate $[n^{-\frac{1}{2}}]$ as the sample size n increases. The later case is often interpreted as a situation where some linear combinations of the structural parameters are ill-determined by the data [see the review of Andrews and Stock (2006), Dufour (2003), and Stock, Wright and Yogo (2002)].

In all setups, we show that under the null hypothesis of interest, the usual χ^2 critical values are applicable whether the instruments are strong or weak. Furthermore, our analysis indicates that test consistency may still hold over a wide range of cases where overall identification fails, pro-

vided partial identification is satisfied. However, the tests exhibit lower power when all instruments are weak. We present a Monte Carlo experiment and an empirical application which confirm our theoretical results.

The paper is organized as follows. Section 2 formulates the model studied. Section 3 describes the test statistics. Sections 3.1-3.2 study the asymptotic properties (level and power) of the tests in both strong and weak identification setups. Section 3.3 presents the Monte Carlo experiment while Section 4 deals with the empirical application. Conclusions are drawn in Section 5 and proofs are presented in the Appendix.

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

2. Framework

We consider the following linear IV regression model

$$y = Y\beta + W\theta + u\,, \tag{2.1}$$

$$Y = Z\mathbf{\Pi} + \upsilon, \quad W = Z\mathbf{\Gamma} + \xi\,, \tag{2.2}$$

where $y \in \mathbb{R}^n$ is a vector of observations on a dependent variable, $Y \in \mathbb{R}^{n \times m_y}$ and $W \in \mathbb{R}^{n \times m_w}$ ($m_y + m_w = m \geq 1$) are two matrices of (possibly) endogenous explanatory variables, $Z \in \mathbb{R}^{n \times l}$ is a matrix of exogenous instruments, $u = (u_1, \ldots, u_n)' \in \mathbb{R}^n$ is the vector of structural disturbances, $\upsilon \in \mathbb{R}^{n \times m_y}$ and $\xi \in \mathbb{R}^{n \times m_w}$ are matrices of reduced form disturbances, $\beta \in \mathbb{R}^{m_y}$ and $\theta \in \mathbb{R}^{m_w}$ are unknown structural parameter vectors, while $\mathbf{\Pi} \in \mathbb{R}^{l \times m_y}$ and $\mathbf{\Gamma} \in \mathbb{R}^{l \times m_w}$ are unknown reduced form coefficient matrices. An extension of model (2.1)-(2.2) that is more relevant for practical purposes arises when we add included exogenous variables Z_1. However, the results of this paper do not alter qualitatively if we replace the variables that are currently in (2.1)-(2.2) by the

residuals that result from their projection onto Z_1. We shall assume that the instrument matrix Z has full-column rank l with probability one and $l \geq m$. The full rank assumption requires excluding redundant columns from Z. It is particularly satisfied when Z_i is generated by power series or splines through an underlying scalar instrument x_i, i.e. if $Z_i = p(x_i) = (1, x_i, \ldots, x_i^{l-1})'$ [see Hansen, Hausman and Newey (2008, Assumption 1) for further details].

The usual necessary and sufficient condition for identification of model (2.1)-(2.2) is $\mathbf{rank}(\mathbf{\Pi}_{YW}) = m$, where $\mathbf{\Pi}_{YW} = [\mathbf{\Pi}, \mathbf{\Gamma}]$. If $\mathbf{\Pi}_{YW} = 0$, the instruments Z are irrelevant, and $(\theta', \beta')'$ is completely unidentified. If $1 \leq \mathbf{rank}(\mathbf{\Pi}_{YW}) < m$, $(\beta', \theta')'$ is not identifiable, but some linear combinations of its elements are identifiable [see Choi and Phillips (1992), Dufour and Hsiao (2008)]. If $\mathbf{\Pi}_{YW}$ is close not to have full rank [*e.g.*, if some eigenvalues of $\mathbf{\Pi}'_{YW}\mathbf{\Pi}_{YW}$ are close to zero], some linear combinations of $(\beta', \theta')'$ are ill-determined by the data, a situation often called "weak identification" in this type of setup [See for example, Staiger and Stock (1997); Stock et al. (2002); Dufour (2003); Andrews and Stock (2006)]. We shall now introduce the statistical problem of interest.

2.1. Statistical problem

We consider the problem of testing the partial exogeneity of Y, *i.e.* the hypothesis

$$\mathrm{H}_0^p : \mathbf{cov}(Y, u) = \sigma_{vu} = 0 \tag{2.3}$$

where the regressors W not being tested may be endogenous [$\mathbf{cov}(W, u) = \sigma_{\xi u} \neq 0$]. By convention, we consider that a matrix is not present if its number of columns is equal to zero. We assume $m_y \geq 1$ but $m_w = 0$ is allowed. In particular, if the null hypothesis (2.3) is specified in the whole set of (possibly) endogenous regressors, we have $m_w = 0$ and W drops out of model (2.1)-(2.2) and H_0^p is the standard exogeneity problem considered by Durbin (1954); Wu (1973); Revankar and Hartley (1973); and Hausman (1978). In this case, Staiger and Stock (1997) and more recently Guggenberger (2010) showed that DWH tests apply even when model parameters are weakly identified.

Our concerned in this paper is how to test H_0^p if $m_w \neq 0$, as DWH-RH tests are no longer valid except when W is exogenous. In this perspective, we aim to provide valid procedures for assessing H_0^p whether W is exogenous or not, with or without weak instruments.

To illustrate the problem, consider the following workhorse example from Card (1995) that analyzes the return on education to earnings.

Example 2.1 The structural equation of interest is given by

$$y_i \quad = \quad Y_i\beta + W_i'\theta + Z_{1i}'\gamma + u_i \tag{2.4}$$

where Y_i is the length of education of individual i; $W_i = (exper_i, exper_i^2)'$ contains the experience ($exper$) and experience squared of individual i where $exper_i = age_i - 6 - Y_i$; $Z_{1i} = (1, race_i, south_i, IQ_i)'$ consists of a constant and indicator variables for race, residence in the south of the United States and IQ score; and y_i is the logarithm of the wage of individual i. All variables in Z_{1i} are assumed exogenous. It is well documented that both Y_i and W_i are potentially endogenous, hence instrumental variables are needed to consistently estimate β and θ in (2.4). The matrix instruments Z contains $age,\ age^2$ of individual i and two proximity-to-college indicators for educational attainment; these are *proximity to 2- and 4-year college.*

To access the joint exogeneity of $(educ,\ exper,\ exper^2)$ in (2.4), we use Wu (1973) T_2-statistic and three alternative Hausman (1978) type-statistics, namely, $\mathcal{H}_j,\ j = 1,\ 2,\ 3$. All these tests are robust to weak instruments, i.e., there are still valid even when model parameters are not identified. We use data from the National Longitudinal Survey of Young Men, which run from 1966 until 1981. We exploit the cross-sectional 1976 subsample that contains originally 3,010 observations. When accounting for missing data, the final sample has 2061 observations.

Our calculations give $\mathcal{T}_2 = 7.01,\ \mathcal{H}_1 = 8.33,\ \mathcal{H}_2 = 8.53$ and $\mathcal{H}_3 = 20.92$ as sample values of the statistics, which correspond to p-values 0.000, 0.040, 0.036 and $0.000,$ respectively. This indicates clearly the evidence against $educ,\ exper$ and $exper^2$ joint exogeneity for all tests. Since joint exogeneity is rejected, one important question is: should we apply IV method to all the regressors $educ,\ exper,\ exper^2$? Note that because the joint exogeneity has been rejected does not imply that all three regressors are endogenous. It could be that only one is endogenous and the two others are not. If so, applying IV to all of them may result in inefficient estimates of model parameters. This underscores the necessity of having ways to check for the exogeneity of each regressor separately.

2.2. Approach and model assumptions

In this paper, we aim to provide valid procedure for assessing H_0^p even when W is endogenous and model identification is weak. The main challenge we are facing is how to deal with the possible simultaneity driving W and u. The strategy that we propose is to replace W by a $\tilde{W}$ that is asymptotically independent with u under H_0^p. Suppose we have regressors $\tilde{W}$ satisfying this condition. We can then express (2.1) as

$$y \quad = \quad Y\beta + \tilde{W}\theta + \tilde{u} \tag{2.5}$$

where $\tilde{u} = u + (W - \tilde{W})\theta$ is asymptotically uncorrelated with $\tilde{W}$. We call equation (2.5) the "*generated structural equation*" to underscore the fact that $\tilde{W}$ are generated regressors. Along with being uncorrelated with $\tilde{u}$, a suitable candidate $\tilde{W}$ in (2.5) should further leave invariant the null hypothesis of interest in (2.3), i.e. $\mathbf{cov}(Y,\, \tilde{u}) = 0$ if $\mathbf{cov}(Y,\, u) = 0$.

We now wish to discuss the choice of $\tilde{W}$. Note first that if ξ has zero mean, the choice of the conditional mean of W given Z is plausible, i.e., $\tilde{W} = \mathbf{E}(W|Z) = Z\mathbf{\Gamma}$. This choice then entails that $\tilde{u} = u + (W - \tilde{W})\theta = u + \xi\theta$. Because Z is exogenous and $\mathbf{\Gamma}$ is fixed, $\tilde{W}$ are also exogenous, hence uncorrelated with $\tilde{u}$. A difficulty however is that $\mathbf{\Gamma}$ is unknown. This suggests we replace $\mathbf{\Gamma}$ by an estimator, say $\tilde{\mathbf{\Gamma}}$, which meets the above requirements. At first, one is tempted to use the least squares estimator $\hat{\mathbf{\Gamma}} = (Z'Z)^{-1}Z'W$ obtained from the first-step regression. Even though $\hat{\mathbf{\Gamma}}$ is a consistent estimator of $\mathbf{\Gamma}$ when the model is correctly specified, it is well known that $\sqrt{n}(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}) = (Z'Z/n)^{-1}Z'\xi/\sqrt{n}$ and $Z'\tilde{u}/\sqrt{n}$ are not independent, even asymptotically. Hence, we will still face a simultaneity problem choosing $\tilde{W} = Z\hat{\mathbf{\Gamma}}$.

Now, assume that $\sigma_{u\xi} = \mathbf{E}(u'\xi) < \infty$ and $0 < \sigma_u^2 = \mathbf{E}(u'u) < +\infty$. Suppose further that $(u,\, v,\, \xi)$ have zero mean and $\frac{1}{\sqrt{n}}Z'[u,\, v,\, \xi]$ is asymptotically Gaussian. Then, we can show that $Z'u/\sqrt{n}$ and $\frac{1}{\sqrt{n}}Z'[(W - Z\mathbf{\Gamma}) - \frac{1}{\sigma_u^2}u\sigma_{u\xi}] = \frac{1}{\sqrt{n}}Z'[\xi - \frac{1}{\sigma_u^2}u\sigma_{u\xi}]$ are asymptotically independent [see Kleibergen (2002)]. Let

$$\tilde{W} \quad = \quad Z\tilde{\mathbf{\Gamma}}, \quad \tilde{\mathbf{\Gamma}} = (Z'Z)^{-1}(Z'W - \frac{1}{\sigma_u^2}Z'u\sigma_{u\xi}) = \hat{\mathbf{\Gamma}} - \frac{1}{\sigma_u^2}(Z'Z)^{-1}Z'u\sigma_{u\xi}. \tag{2.6}$$

The choice of $\tilde{W}$ in (2.6) then implies $\tilde{u} = u + (W - \tilde{W})\theta = u + M_Z\xi\theta + \sigma_\theta P_Z u$ so that $Z'\tilde{u}/\sqrt{n} = (1 + \sigma_\theta)Z'u/\sqrt{n}$ is proportional to $Z'u/\sqrt{n}$, where $\sigma_\theta = \sigma_{u\xi}\theta/\sigma_u^2$ is a scalar. Since $Z'u/\sqrt{n}$

is asymptotically independent of $\frac{1}{\sqrt{n}}Z'[\xi - \frac{1}{\sigma_u^2}u\sigma_{u\xi}]$, hence $Z'\tilde{u}/\sqrt{n}$ and $\frac{1}{\sqrt{n}}Z'[\xi - \frac{1}{\sigma_u^2}u\sigma_{u\xi}]$ are also asymptotically independent. Hence, $Z'\tilde{u}/\sqrt{n}$ and $\sqrt{n}(\tilde{\mathbf{\Gamma}} - \mathbf{\Gamma})$ are asymptotically independent; which means that the choice of $\tilde{W}$ in (2.6) weighs out the simultaneity problem. $\tilde{\mathbf{\Gamma}}$ can be viewed here as the part of $\hat{\mathbf{\Gamma}}$ that is asymptotically orthogonal to u. Furthermore, when the above regularity conditions hold, we have $Y'\tilde{u}/n \xrightarrow{p} \sigma_{\upsilon u} + \Sigma_{\upsilon\xi}\theta$, where $\Sigma_{\upsilon\xi} = \mathbf{E}(\upsilon_i\xi_i')$ for all i. In particular, if υ and ξ are uncorrelated (i.e. if $\Sigma_{\upsilon\xi} = 0$) under H_0^p, we have $p\lim_{n\to\infty}(Y'\tilde{u}/n) = 0$ and H_0^p can in principle be assessed by testing whether Y is exogenous in model (2.5).

However, it is practically impossible to exploit (2.6) as u, $\sigma_{u\xi}$ and σ_u^2 are unknown. To alleviate this difficulty, we suggest a strategy built on the following four steps:

1. project W on Z to obtain $\bar{W} = P_Z W$;
2. regress y on Y and $\bar{W}$ by OLS and recover the residuals, say $\hat{u}_*$;
3. estimate $\sigma_{u\xi}$ by $\hat{\sigma}_{uW} = \hat{u}_*'M_Z W/(n-m)$ and σ_u^2 by $\hat{\sigma}_u^2 = \hat{u}_*'M_Z\hat{u}_*/(n-m)$;
4. and generate $\tilde{W}$ as

$$\tilde{W} = Z\tilde{\mathbf{\Gamma}}, \quad \tilde{\mathbf{\Gamma}} = \hat{\mathbf{\Gamma}} - (Z'Z)^{-1}Z'\hat{u}_*(\hat{u}_*'M_Z\hat{u}_*)^{-1}\hat{u}_*'M_Z W. \tag{2.7}$$

Note that $\tilde{\mathbf{\Gamma}}$ in (2.7) can be expressed as $\tilde{\mathbf{\Gamma}} = (Z'Z)^{-1}Z'\mathbf{A}(\hat{u}_*)W$, where $\mathbf{A}(\hat{u}_*) = I - \hat{u}_*(\hat{u}_*'M_Z\hat{u}_*)^{-1}\hat{u}_*'M_Z$. If $Z'Z/n = O_p(1)$ and $Z'W/n = O_p(1)$ along with the exogeneity of Z, then we have $\hat{u}_*'M_Z\hat{u}_*/(n-m) = \hat{u}_*'\hat{u}_*/(n-m) + o_p(1)$ and $\hat{u}_*'M_Z W/(n-m) = \hat{u}_*'W/(n-m) + o_p(1)$, so that $\tilde{\mathbf{\Gamma}} = (Z'Z)^{-1}Z'M_{\hat{u}_*}W + o_p(1)$, where $M_{\hat{u}_*}$ is the projection matrix onto the orthogonal of the space spanned by the residuals $\hat{u}_*$. Hence, $\tilde{\mathbf{\Gamma}}$ is asymptotically orthogonal to the residual $\hat{u}_*$. When identification is strong, $\tilde{\mathbf{\Gamma}} \xrightarrow{p} \mathbf{\Gamma}$ under standard regularity conditions, which is always independent with the asymptotic distribution of $Z'\tilde{u}/\sqrt{n}$. However, when identification is weak, $\tilde{\mathbf{\Gamma}}$ converges to a random variable which is correlated with the asymptotic distribution of $Z'u/\sqrt{n}$. The aim of the orthogonalization by $\tilde{W}$ is guarantee asymptotically, the independence between $Z'\tilde{u}/\sqrt{n}$ and $\mathbf{\Gamma}_\psi$. It is worthwhile noting that the choice of $\tilde{W}$ in (2.7) implies the following form of the errors $\tilde{u}$ in (2.5):

$$\tilde{u} = u + (W - \tilde{W})\theta = u + M_Z\xi\theta + \hat{\sigma}_\theta P_Z\hat{u}_* \quad \text{where} \quad \hat{\sigma}_\theta = \hat{\sigma}_{uW}\theta/\hat{\sigma}_u^2. \tag{2.8}$$

We now make the following generic assumptions on the behaviour of model variables.

Assumption 2.2 *The errors* $\left\{U_i = \left(u_i,\, \upsilon_i',\, \xi_i'\right)' :\, 1 \leq i \leq n\right\}$ *are i.i.d. across* i *and* n *with zero mean and the same nonsingular covariance matrix* $\mathbf{\Sigma}$ *given by*

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_u^2 & \sigma_{Vu}' \\ \sigma_{Vu} & \Sigma_V \end{pmatrix} :\, (m+1) \times (m+1), \quad \textit{where } \Sigma_V = \begin{pmatrix} \Sigma_\upsilon & \Sigma_{\xi\upsilon}' \\ \Sigma_{\xi\upsilon} & \Sigma_\xi \end{pmatrix},$$

$\sigma_{Vu} = (\sigma_{\upsilon u}', \sigma_{\xi u}')',\ \sigma_u^2 : 1 \times 1,\ \sigma_{\upsilon u} : m_y \times 1,\ \sigma_{\xi u} : m_w \times 1,\ \Sigma_\upsilon : m_y \times m_y,\ \Sigma_{\xi\upsilon} : m_w \times m_y,$ $\Sigma_\xi : m_w \times m_w,$ *and* $\sigma_u^2 - \theta'\Sigma_\xi\theta > 0$. *Furthermore, we have* $\mathbf{E}(Z_iU_i') = 0$ *for all* $i = 1, \ldots, n$.

Assumption **2.2** requires model errors to be homoskedastic. However, it can be adapted to account for serially correlated errors.

Assumption 2.3 *When the sample size* n *converges to infinity, the following convergence results hold jointly*: $(\mathbf{a})$ $\frac{1}{n}\sum_{i=1}^n U_iU_i' \xrightarrow{p} \mathbf{\Sigma}$, $\frac{1}{n}\sum_{i=1}^n Z_iU_i' \xrightarrow{p} 0$, $\frac{1}{n}\sum_{i=1}^n Z_iZ_i' \xrightarrow{p} Q_Z$; *and* $(\mathbf{b})$ $\frac{1}{\sqrt{n}}\sum_{i=1}^n \left(Z_iU_i',\, \upsilon_iu_i - \sigma_{\upsilon u}\right) \xrightarrow{d} \mathbf{\Psi} = (\Psi_Z,\, \psi_{\upsilon u})$, *where* $\Psi_Z = (\psi_{Zu},\, \psi_{Z\upsilon},\, \psi_{Z\xi})$, $\mathbf{vec}(\mathbf{\Psi}) \sim \mathbf{N}\left(0\,,\, \mathbf{\Omega}\right),\ \mathbf{vec}(\Psi_Z) \sim \mathbf{N}\left(0\,,\, \mathbf{\Sigma} \otimes Q_Z\right)$ *and* $\psi_{\upsilon u} \sim \mathbf{N}\left(0\,,\, \sigma_u^2\Sigma_\upsilon\right)$.

Assumption **2.3**-$(\mathbf{b})$ entails that Z is weakly exogenous for $(\beta',\, \theta')',\ \mathbf{\Pi},$ and $\mathbf{\Gamma}$ [see Engle, Hendry and Richard (1982)]. The normality assumption on the limiting distributions is implied by Assumption **2.2** and the central limit theorem (CLT).

Assumption 2.4 *Under* H_0^p, *the following two conditions hold*: $(\mathbf{a})$ $\frac{1}{n}\sum_{i=1}^n \upsilon_i\xi_i' = O_p(n^{-\nu})$ *for some* $\nu > 1/2$; *and* $(\mathbf{b})$ $\frac{1}{n}\sum_{i=1}^n W_i\hat{u}_{*i} = O_p(n^{-\frac{1}{2}})$, *where* $\{\hat{u}_{*i} :\, 1 \leq i \leq n\}$ *are the residuals from the OLS regression in* (2.7).

It is worth noting that Assumption **2.4** needs not to be satisfied under the alternative. Assumption **2.4**-$(\mathbf{a})$ along with Assumptions **2.2-2.3** entail that $\frac{1}{n}\sum_{i=1}^n \upsilon_i\xi_i' \xrightarrow{p} \mathbf{E}(\upsilon_i\xi_i') \to 0$ and $n^\nu\mathbf{E}(\upsilon_i\xi_i') = O_p(1),$ as $n \to \infty$ for some $\nu > 1/2$. This means that the covariance matrix, Σ_V, of the reduced-form errors $(\upsilon,\, \xi)$ is asymptotically diagonal under H_0^p. This assumption is particularly satisfied under H_0^p if υ and ξ are uncorrelated ($\Sigma_{\upsilon\xi} = 0$) or more generally if $\Sigma_{\upsilon\xi} = \bar{\Sigma}_{\upsilon\xi}/n^\nu$ for some $\nu > 1/2,$ where $\bar{\Sigma}_{\upsilon\xi}$ is a $m_y \times m_w$ constant matrix. Furthermore, note that we also have $\frac{1}{\sqrt{n}}\sum_{i=1}^n \upsilon_i\xi_i' = n^{\frac{1}{2}-\nu}n^\nu\frac{1}{n}\sum_{i=1}^n \upsilon_i\xi_i' = o_p(1).O_p(1) = o_p(1),$ since $\nu > 1/2$. The condition that

$\frac{1}{n}\sum_{i=1}^{n} W_i\hat{u}_{*i} = O_p(n^{-\frac{1}{2}})$ in Assumption **2.4**-(**b**) implies that the correlation between the residuals from the OLS regression in (2.7) and W converge to zero in probability, as the sample size n increases. It follows that $\hat{u}_*'W/\sqrt{n} = O_p(1)$. Remark that $\hat{u}_*'W/n \xrightarrow{p} 0$ does not implies that the covariance between the structural error u and W (here $\sigma_{\xi u}$) converges to zero. However, it implies a restriction of the form $\sigma_{u\xi} = -\theta'\Sigma_\xi$ involving $\sigma_{\xi u}$, Σ_ξ and θ. Clearly, u and W may still be asymptotically correlated even if $\hat{u}_*'W/n \xrightarrow{p} 0$ [3].

In this paper, we consider two main setups related to the identification of model parameters: (**i**) $\mathbf{\Pi}_{YW} = [\mathbf{\Pi},\ \mathbf{\Gamma}]$ is fixed with $\text{rank}(\mathbf{\Pi}_{YW}) = m$; and (**ii**) $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0,\ \mathbf{\Gamma}_0]$, where $\mathbf{\Pi}_0$ and $\mathbf{\Gamma}_0$ are constant $l \times m_y$ and $l \times m_w$ matrices (possibly zero). The setup for (**i**) implies that $(\beta',\ \theta')'$ is identified, hence the instruments Z are strong. However, our results can be extended to cases where $(\beta',\ \theta')'$ is partially identified [i.e., $\mathbf{\Pi}_{YW}$ is fixed with $0 \leq \text{rank}(\mathbf{\Pi}_{YW}) < m$], upon rotating model variables in an appropriate way [See for example, Choi and Phillips (1992), Doko Tchatoka and Dufour (2011), and Doko Tchatoka (2011)]. (**ii**) is Staiger and Stock (1997) *local-to-zero* weak instruments asymptotic. The parameter that controls the strength of the instruments approaches zero at rate $1/\sqrt{n}$ as the sample size n increases.

We can now prove the following lemma on the asymptotic behaviour of $Z'\hat{u}_*/n,\ Z'\tilde{u}/n,\ \tilde{W}'\tilde{u}/n,$ and $Y'\tilde{u}/n$.

Lemma 2.5 *Suppose Assumptions* **2.2-2.4** *hold and let* $\sigma_{\upsilon u} = 0$. *Then we have*: $Z'\tilde{u}/n,\ \tilde{W}'\tilde{u}/n,\ Z'\hat{u}_*/n,\ Y'\tilde{u}/n \xrightarrow{p} 0,$ *irrespective of whether the instrument are strong or weak.*

Lemma **2.5** shows clearly that $\tilde{W}$ is asymptotically uncorrelated with $\tilde{u}$ in (2.5) and further, that H_0^p is asymptotically invariant by the transformation (2.7).

We now consider the following transformed model:

$$y^\perp = Y^\perp\beta + \tilde{u}^\perp, \quad Y^\perp = Z^\perp\mathbf{\Pi} + \upsilon^\perp \tag{2.9}$$

where the superscript "$\perp$" means residual from projection onto the space spanned by the columns of $\tilde{W}$. As $\tilde{W}$ is asymptotically uncorrelated with $\tilde{u}$ under H_0^p by Lemma **2.5**, $Z^\perp$ is asymptotically a valid instrument for $Y^\perp$. Furthermore, by exploiting (2.8), we can easily show that $Y^{\perp\prime}\tilde{u}^\perp/n \xrightarrow{p} \sigma_{\upsilon u} + \Sigma_{\upsilon\xi}\theta$. If Assumption **2.4** and H_0^p are satisfied, we have $\Sigma_{\upsilon\xi} = 0$ and $\sigma_{\upsilon u} = 0$ so that

[3] Under Assumptions **2.2-2.4**, we have $p\lim_{\to\infty}\left(\frac{W'\hat{u}_*}{n}\right) = \sigma^*_{\xi u} = \sigma_{\xi u} + \Sigma_\xi\theta$. Hence, $\sigma^*_{\xi u} = 0 \Leftrightarrow \sigma_{u\xi} = -\theta'\Sigma_\xi$ so that the remark follows.

$Y^{\perp\prime}\tilde{u}^{\perp}/n \xrightarrow{p} 0,$ which means that H_0^p can be assessed by testing whether $Y^{\perp}$ is uncorrelated with $\tilde{u}^{\perp}$ in (2.9).

If β is identified[4] in (2.9), both the OLS estimator (namely $\hat{\beta}_{LS}$) and IV estimator ($\hat{\beta}_{IV}$) of β are consistent under H_0^p, and $\hat{\beta}_{LS}$ is efficient. Hence, the magnitude of the vector of contrasts is small in that case [$\hat{\beta}_{LS} - \hat{\beta}_{IV} = o_p(1)$]. However, when H_0^p is not satisfied ($\sigma_{uv} \neq 0$), $\hat{\beta}_{IV}$ is still consistent but $\hat{\beta}_{LS}$ is not, so that $\hat{\beta}_{LS} - \hat{\beta}_{IV} = O_p(1)$. Therefore, in the same spirit as Durbin (1954), Wu (1973), and Hausman (1978), we can build the test statistics for assessing H_0^p on $\hat{\beta}_{LS} - \hat{\beta}_{IV},$ upon scaling by appropriate factors to guarantee the usual asymptotic χ^2-distributions.

More interestingly, Lemma **2.6** shows that $(Z^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n},\, v^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n})$ is asymptotically independent of $\sqrt{n}(\tilde{\mathbf{\Gamma}} - \mathbf{\Gamma}),$ whether identification is strong or weak. So, the (possible) simultaneity driving W and u has been eliminated by the transformation (2.7), as required.

Lemma 2.6 *Suppose Assumptions* **2.2-2.4** *hold and let* $\sigma_{vu} = 0$. *Then we have* $(Z^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n},\, v^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n}) \xrightarrow{d} (\psi_{Z^{\perp}\tilde{u}},\, \psi_{v^{\perp}\tilde{u}})$ *where*: (**i**) $(\psi_{Z^{\perp}\tilde{u}},\, \psi_{v^{\perp}\tilde{u}}) \sim \mathbf{N}\left[0,\, \sigma_u^2 \mathbf{diag}(Q_{Z^{\perp}},\, \Sigma_v)\right]$, *with* $Q_{Z^{\perp}} = Q_Z^{1/2} M_{Q_Z^{1/2}\mathbf{\Gamma}} Q_Z^{1/2}$, *rank*$(\mathbf{\Pi}_{YW}) = m$; *and* (**ii**) $(\psi_{Z^{\perp}\tilde{u}}, \psi_{v^{\perp}\tilde{u}}) \sim \int_{\mathbb{R}^{l\times m_w}} \mathbf{N}\left[0, \sigma_u^2 \mathbf{diag}(Q_Z^{1/2} M_{Q_Z^{1/2}\mathbf{\Gamma}(x_2)} Q_Z^{1/2}, \Sigma_v)\right] pdf(x_2)dx_2$ *when* $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0,\, \mathbf{\Gamma}_0],\ \mathbf{\Gamma}(x_2) = \mathbf{\Gamma}_0 + Q_Z^{-1}x_2$ *and* $pdf(x_2)$ *is the probability density function of* $\psi_{Z\xi}$ *evaluated at* x_2.

Three remarks are in order.

1. The results indicate that $Z^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n}$ is asymptotically uncorrelated with $v^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n}$ and $v^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n} \xrightarrow{d} \psi_{vu} \sim \mathbf{N}\left[0,\, \sigma_u^2\Sigma_v\right]$, whether identification is strong or not. Consequently, weak identification does not affect the asymptotic behaviour of $v^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n}$ but the asymptotic behaviour of $Z^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n}$ relies strongly on instrument quality.

2. When identification is strong [rank$(\mathbf{\Pi}_{YW}) = m$], $\tilde{\mathbf{\Gamma}} \xrightarrow{p} \mathbf{\Gamma}$ which is a constant $l \times m_w$ full rank matrix. Hence, $(Z^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n},\, v^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n})$ is asymptotically Gaussian, as expected [see Lemma **2.6**-(**i**)]. However, when identification is weak (weak instruments), $\tilde{\mathbf{\Gamma}} \xrightarrow{d} \mathbf{\Gamma}(\psi_{Z\xi}) = \mathbf{\Gamma}_0 + Q_Z^{-1}\psi_{Z\xi}$ which is a non-degenerated random process with probability one. As a result, the asymptotic distribution of $(Z^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n},\, v^{\perp\prime}\tilde{u}^{\perp}/\sqrt{n})$ is a mixture of Gaussian processes

[4] It is well known that IV methods produce inconsistent estimates when identification is weak, see for example, Dufour (2003), Stock et al. (2002), Stock and Wright (2000), Bekker (1994), Choi and Phillips (1992), Nelson and Startz (1990a, 1990b), Phillips (1989).

with zero mean, as showed Lemma **2.6**-(**ii**). Note that mixture is in the marginal distribution of $\psi_{Z^{\perp}\tilde{u}}$, because $\psi_{\upsilon^{\perp}\tilde{u}}$ is independent of both $\mathbf{\Gamma}(\psi_{Z\xi})$ and $\psi_{Z^{\perp}\tilde{u}}$ when Assumptions **2.2-2.4** and H_0^p hold.

3. When identification is weak, the independence between $(\psi_{Z^{\perp}\tilde{u}},\ \psi_{\upsilon^{\perp}\tilde{u}})$ and $\mathbf{\Gamma}(\psi_{Z\xi})$ is crucial to establish the validity of the tests that are proposed in the next section for assessing H_0^p.

3. Test statistics and their asymptotic behaviour

We propose four alternative statistics to assess H_0^p, namely

$$\mathscr{D}_j^p \quad = \quad \kappa_j(\hat{\beta}_{LS}-\hat{\beta}_{IV})'\hat{\Sigma}_j^{-1}(\hat{\beta}_{LS}-\hat{\beta}_{IV}),\ j=1,\ 2,\ 3,\ 4 \tag{3.1}$$

where $\kappa_1=(n-2m_y)/m_y,\ \kappa_i=n,$ for $j=2,\ 3,\ 4,$ and

$$\begin{aligned}
\hat{\beta}_{LS} &= (Y^{\perp\prime}Y^{\perp})^{-1}Y^{\perp\prime}y,\ \hat{\beta}_{IV}=(Y'P_{Z^{\perp}}Y)^{-1}Y'P_{Z^{\perp}}y,\\
\hat{\Sigma}_1 &= \tilde{\sigma}_2^2\hat{\Delta},\ \hat{\Delta}=\hat{\Omega}_{IV}^{-1}-\hat{\Omega}_{LS}^{-1},\ \hat{\Sigma}_2=\tilde{\sigma}^2\hat{\Omega}_{IV}^{-1}-\hat{\sigma}^2\hat{\Omega}_{LS}^{-1},\ \hat{\Sigma}_3=\tilde{\sigma}^2\hat{\Delta},\ \hat{\Sigma}_4=\hat{\sigma}^2\hat{\Delta},\\
\hat{\Omega}_{IV} &= Y'P_{Z^{\perp}}Y/n,\ \hat{\Omega}_{LS}=Y^{\perp\prime}Y^{\perp}/n,\ \tilde{\sigma}^2=(y-Y\hat{\beta}_{IV})'M_{\tilde{W}}(y-Y\hat{\beta}_{IV})/n,\\
\hat{\sigma}^2 &= (y-Y\hat{\beta}_{LS})'M_{\tilde{W}}(y-Y\hat{\beta}_{LS})/n,\ \tilde{\sigma}_2^2=\hat{\sigma}^2-(\hat{\beta}_{LS}-\hat{\beta}_{IV})'\hat{\Delta}^{-1}(\hat{\beta}_{LS}-\hat{\beta}_{IV}).
\end{aligned}$$

The above expressions of $\hat{\beta}_{LS},\ \hat{\beta}_{IV}$ and $\hat{\Omega}_{IV}$ are derived from the identities $Y^{\perp\prime}y^{\perp}=Y^{\perp\prime}y$, $P_{Z^{\perp}}Y^{\perp}=P_{Z^{\perp}}Y$ and $P_{Z^{\perp}}y^{\perp}=P_{Z^{\perp}}y$. The statistics in (3.1) differ only through the variance estimators of the errors $\tilde{u}^{\perp}$ in (2.9) and the scaling factors $\kappa_j,\ j=1,\ 2,\ 3,\ 4.$ $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ are the usual OLS-and IV-based estimators of the errors (without correction for degrees of freedom), while $\tilde{\sigma}_2^2$ can be interpreted as an alternative IV-based scaling factor. The use of different estimators of the variance of the errors that leads to four versions of the test is important to discriminate between the OLS-and IV-based residuals, especially when identification is weak. When identification is weak, the OLS estimator often outperforms [in terms of minimum mean squared errors (MSE)] the IV estimator [see Kiviet and Niemczyk (2007) and Doko Tchatoka and Dufour (2011)]. The statistic $\mathscr{D}_1^p$ is an analogue to Wu (1973) T_2-statistic and can be interpreted as a usual F-test[5] of $\gamma=0$ in

[5]Further details on the regression interpretation of Durbin-Wu-Hausman tests can be found in Doko Tchatoka and Dufour (2011) and Davidson and Mackinnon (1993, sec. 8.7).

the extended regression

$$y^{\perp} \quad = \quad Y^{\perp}\beta + \hat{v}^{\perp}\gamma + e \tag{3.2}$$

where $\hat{v}^{\perp} = M_{Z^{\perp}}Y^{\perp}$, $e = P_{Z^{\perp}}v^{\perp}\gamma + \varepsilon$, and ε is independent of $v^{\perp}$. The statistics $\mathscr{D}_j^p$ ($j = 2, 3, 4$) are analogues to alternative Hausman (1978) type-statistics considered in Staiger and Stock (1997)[6]. The subscript "p" in the notation of the statistics, as well as the null hypothesis, refers to partial exogeneity. The corresponding tests reject H_0^p when the test statistic is "large". Section 3.1 investigates the size and power properties of the tests when identification is strong (strong instruments).

3.1. Test behaviour with strong instruments

Before investigating the properties (size and power) of the tests, we shall first examine the behaviour of the vector of contrasts $\hat{\beta}_{LS} - \tilde{\beta}_{IV}$. Lemma **3.1** present the results under both the null hypothesis ($\sigma_{vu} = 0$) and the alternative hypothesis ($\sigma_{vu} \neq 0$ is fixed).

Lemma 3.1 *Suppose Assumptions* $\mathbf{2.2}$-$\mathbf{2.4}$ *hold and rank*$(\mathbf{\Pi}_{YW}) = m$. *Then we have*:

$$\begin{array}{lllll} (\mathbf{i}) & \hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{p} 0, & \sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV}) \xrightarrow{d} \mathbf{N}\left[0,\, \sigma_u^2(\tilde{\mathbf{\Sigma}}_{\pi}^{-1} - \mathbf{\Sigma}_{\pi}^{-1})\right] & \textit{when } \sigma_{vu} = 0; \\ (\mathbf{ii}) & \hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{p} & \mathbf{\Sigma}_{\pi}^{-1}\sigma_{vu}, \quad \sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV}) \xrightarrow{d} \infty & \textit{when } \sigma_{vu} \neq 0; \end{array}$$

where $\mathbf{\Sigma}_{\pi} = \tilde{\mathbf{\Sigma}}_{\pi} + \Sigma_v$, $\tilde{\mathbf{\Sigma}}_{\pi} = \mathbf{\Pi}' Q_{Z^{\perp}} \mathbf{\Pi}$, $Q_{Z^{\perp}}$ *is defined in Lemma* $\mathbf{2.6}$-$(\mathbf{i})$.

Lemma **3.1**-$(\mathbf{i})$ states the *consistency* to zero and the $\sqrt{n}$*-consistency* of the vector of contrasts $\hat{\beta}_{LS} - \tilde{\beta}_{IV}$ when H_0^p holds and identification is strong. As expected, the limiting distribution of $\sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV})$ is Gaussian with zero mean and constant positive definite covariance matrix $\sigma_u^2(\tilde{\mathbf{\Sigma}}_{\pi}^{-1} - \mathbf{\Sigma}_{\pi}^{-1})$. Under the alternative hypothesis ($\sigma_{vu} \neq 0$ is fixed, i.e., does not depend on the sample size[7]), $\hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{p} \mathbf{\Sigma}_{\pi}^{-1}\sigma_{vu} \neq 0$ so that $\sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV})$ explodes, as showed Lemma **3.1**-$(\mathbf{ii})$. We can now characterize the asymptotic distributions of the statistics under both the null hypothesis (level) and the alternative hypothesis (power). Theorem **3.2** presents the results.

[6] See also Guggenberger (2010) and Hahn et al. (2010).

[7] Throughout this paper, our analysis is based on alternative hypotheses of the form $\mathrm{H}_1^p : \sigma_{vu} \neq 0$ where σ_{vu} is a $m_y \times 1$ constant vector. However, it is easy to show that under *local-to-zero* alternative hypotheses of the form $\mathrm{H}_{1c}^p : \sigma_{vu} = \mathbf{c}/\sqrt{n}$ where $\mathbf{c} \neq 0$ is constant, $\sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV})$ converges to a Gaussian process with nonzero mean when identification is strong. As a result, all tests in (3.1) exhibit power against *local-to-zero* alternatives, though they are no longer consistent.

Theorem 3.2 *Suppose Assumptions* **2.2-2.4** *are satisfied and rank*$(\mathbf{\Pi}_{YW}) = m$. *Then we have*: $(\mathbf{a})$ $\mathscr{D}_1^p \xrightarrow{d} \frac{1}{m_y}\chi^2(m_y)$, $\mathscr{D}_j^p \xrightarrow{d} \chi^2(m_y)$ $\forall\, j = 2, 3, 4$, *when* $\sigma_{\upsilon u} = 0$; *and* $(\mathbf{b})$ $\mathscr{D}_j^p \xrightarrow{d} +\infty$ $\forall\, j = 1, 2, 3, 4$, *when* $\sigma_{\upsilon u} \neq 0$.

Theorem **3.2**-$(\mathbf{a})$ shows that all $\mathscr{D}^p$ statistics are asymptotically pivotal when identification is strong. Hence, the corresponding tests are asymptotically valid (level is controlled). Theorem **3.2**-$(\mathbf{b})$ indicates that test consistency holds, thus confirming the previous results in Lemma **3.1**-$(\mathbf{ii})$. The Monte Carlo experiment shows that: $(\mathbf{1})$ level is still controlled for moderate samples [see Figure 1 for $n = 100$], and $(\mathbf{2})$ test consistency may still hold in a wide range of cases where the overall identification breaks down, provided partial identification is satisfied [i.e., $\mathbf{\Pi}_{YW}$ is fixed and $0 < \text{rank}(\mathbf{\Pi}_{YW}) < m$]. So, the above results extend to partial identification of model parameters. More generally, it can be shown that the necessary and sufficient condition for consistency is that $\mathbf{\Pi}\Sigma_{\upsilon}^{-1}\sigma_{u\upsilon} \neq 0$. We now study the behaviour of the tests under Staiger and Stock's (1997) *local-to-zero weak instrument asymptotic*.

3.2. Test behaviour with weak instruments

In this section, we assume that model parameters are weakly identified, i.e., $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0,\, \mathbf{\Gamma}_0]$, where $\mathbf{\Pi}_0$ and $\mathbf{\Gamma}_0$ are constant matrices (possibly zero). As in the previous section, we first examine the behaviour of the vector of contrast $\hat{\beta}_{LS} - \tilde{\beta}_{IV}$. Lemma **3.3** presents the results under both the null hypothesis and the alternative hypothesis.

Lemma 3.3 *Suppose Assumptions* **2.2-2.4** *hold and* $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0,\, \mathbf{\Gamma}_0]$. *Then, we have*:

$$(\mathbf{i}) \quad \hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{d} \int_{\mathbb{R}^{l\times m_w}} \int_{\mathbb{R}^{l\times m_y}} \mathbf{N}\left[0, \sigma_u^2 \mathbf{\Psi}_{Z\upsilon}^{-1}\right] pdf(x_1, x_2)dx_1dx_2 \quad \textit{when } \sigma_{\upsilon u} = 0;$$

$$(\mathbf{ii}) \quad \hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{d} \int_{\mathbb{R}^{l\times m_w}} \int_{\mathbb{R}^{l\times m_y}} \mathbf{N}\left[\mu, \sigma_u^2 \mathbf{\Psi}_{Z\upsilon}^{-1}\right] pdf(x_1, x_2)dx_1dx_2 \quad \textit{when } \sigma_{\upsilon u} \neq 0$$

where $\mu \equiv \mu(x_1, x_2) = \mathbf{\Psi}_{Z\upsilon}^{-1}(x_1, x_2)(\mathbf{\Pi}_0 + Q_Z^{-1}x_1)' Q_Z^{1/2} M_{Q_Z^{1/2}\mathbf{\Gamma}(x_2)} Q_Z^{1/2} \mathbf{\Pi}_0 \rho_{\upsilon u}$, $\mathbf{\Psi}_{Z\upsilon} \equiv \mathbf{\Psi}_{Z\upsilon}(x_1, x_2) = (\mathbf{\Pi}_0 + Q_Z^{-1}x_1)' Q_Z^{1/2} M_{Q_Z^{1/2}\mathbf{\Gamma}(x_2)} Q_Z^{1/2} (\mathbf{\Pi}_0 + Q_Z^{-1}x_1)$, $pdf(x_1, x_2)$ *is the joint probability density function of* $(\psi_{Z\upsilon}, \psi_{Z\xi})$, *and* $\mathbf{\Gamma}(x_2) = \mathbf{\Gamma}_0 + Q_Z^{-1}x_2$.

In contrast of Lemma **3.1**, observe now that $\hat{\beta}_{LS} - \tilde{\beta}_{IV}$ converges to a non degenerated random variable, $\tilde{\Psi}_\beta$, under H_0^p. Though $\hat{\beta}_{LS}$ is still consistent under H_0^p despite the lack of identification, $\hat{\beta}_{IV}$ is not. The lack of identification then implies that $\hat{\beta}_{LS} - \hat{\beta}_{IV} = O_p(1)$ under H_0^p. Because

$M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zu},$ is independent of $Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})$ and ψ_{Zv} under $\mathrm{H}_0^p,$ the conditional limiting distribution of $\hat{\beta}_{LS}-\tilde{\beta}_{IV},$ given $(\psi_{Zv},\, \psi_{Z\xi}),$ is Gaussian with zero mean. So, its unconditional null limiting distribution is a mixture of Gaussian processes with zero mean. Under the alternative hypothesis ($\sigma_{vu} \neq 0$), the conditional limiting distribution of $\hat{\beta}_{LS} - \tilde{\beta}_{IV},$ given $(\psi_{Zv},\, \psi_{Z\xi}),$ is Gaussian with nonzero mean so that its unconditional limiting distribution is a mixture of Gaussian processes with nonzero mean.

Let $\phi_0(x_1,x_2) = [1 + \|\sigma_u^{-1}\Sigma_v^{1/2}\mathbf{N}\left(0,\, \sigma_u^2\mathbf{\Psi}_{Zv}^{-1}(x_1,x_2)\right)\|^2]^{-1} \leq 1$ and $\phi(x_1,x_2) = [1 + \|\sigma_u^{-1}\Sigma_v^{1/2}\mathbf{N}\left(\mu(x_1,x_2) - \rho_{vu},\, \sigma_u^2\mathbf{\Psi}_{Zv}^{-1}(x_1,x_2)\right)\|^2]^{-1} \leq 1$. Theorem **3.4** characterizes the asymptotic distributions of $\mathscr{D}^p$ statistics when instruments are *local-to-zero*.

Theorem 3.4 *Suppose Assumptions* $\mathbf{2.2}$-$\mathbf{2.4}$ *are satisfied and* $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0,\, \mathbf{\Gamma}_0]$. $(\mathbf{a})$ *If* $\sigma_{vu} = 0$, *then we have*:

$$\mathscr{D}_1^p \xrightarrow{d} \frac{1}{m_y}\chi^2(m_y),\ \mathscr{D}_4^p \xrightarrow{d} \chi^2(m_y),$$
$$\mathscr{D}_j^p \xrightarrow{d} \chi^2(m_y)\int_{\mathbb{R}^{l\times m_w}}\int_{\mathbb{R}^{l\times m_y}} \phi_0(x_1,x_2)pdf(x_1,x_2)dx_1dx_2 \leq \chi^2(m_y)$$

for $j = 2,\, 3.$ $(\mathbf{b})$ *If* $\sigma_{vu} \neq 0,$ *then we have*:

$$\mathscr{D}_1^p \xrightarrow{d} \frac{1}{m_y}\int_{\mathbb{R}^{l\times m_w}}\int_{\mathbb{R}^{l\times m_y}} \chi^2(m_y;\, \|\sigma_u^{-1}\mathbf{\Psi}_{Zv}^{1/2}\mu\|^2)pdf(x_1,x_2)dx_1dx_2,$$
$$\mathscr{D}_4^p \xrightarrow{d} \int_{\mathbb{R}^{l\times m_w}}\int_{\mathbb{R}^{l\times m_y}} \chi^2(m_y;\, \|\sigma_u^{-1}\mathbf{\Psi}_{Zv}^{1/2}\mu\|^2)pdf(x_1,x_2)dx_1dx_2,$$
$$\mathscr{D}_j^p \xrightarrow{d} \int_{\mathbb{R}^{l\times m_w}}\int_{\mathbb{R}^{l\times m_y}} \phi(x_1,x_2)\chi^2(m_y;\, \|\sigma_u^{-1}\mathbf{\Psi}_{Zv}^{1/2}\mu\|^2)pdf(x_1,x_2)dx_1dx_2$$
$$\leq \int_{\mathbb{R}^{l\times m_w}}\int_{\mathbb{R}^{l\times m_y}} \chi^2(m_y;\, \|\sigma_u^{-1}\mathbf{\Psi}_{Zv}^{1/2}\mu\|^2)pdf(x_1,x_2)dx_1dx_2$$

for $j = 2,\, 3,$ *where* $\mathbf{\Psi}_{Zv} \equiv \mathbf{\Psi}_{Zv}(x_1,x_2)$ *and* $\mu \equiv \mu(x_1,x_2)$ *are defined in Lemma* $\mathbf{3.3}$.

Firstly, we note that under H_0^p ($\sigma_{vu} = 0$), $\mathscr{D}_1^p$ and $\mathscr{D}_4^p$ are still asymptotically pivotal despite identification issues. Hence, these tests have correct size with weak instruments. However, $\mathscr{D}_2^p$ and $\mathscr{D}_3^p$ are boundedly asymptotically pivotal. The upper bound of their limiting distribution correspond to their asymptotic distribution when identification is strong. So, the usually χ^2 critical values are still applicable to these tests, even though doing so leads to conservative procedures. Clearly, all proposed $\mathscr{D}^p$-tests can be described as identification-robust. Secondly, when $\sigma_{vu} \neq 0$, $\mathscr{D}_1^p$ and $\mathscr{D}_4^p$

converge to a mixtures of noncentral χ^2 distributions, while $\mathscr{D}_2^p$ and $\mathscr{D}_3^p$ are asymptotically bounded by a mixture of noncentral χ^2 distributions. Hence the tests $\mathscr{D}_1^p$ and $\mathscr{D}_4^p$ are more powerful than $\mathscr{D}_2^p$ and $\mathscr{D}_3^p$. Moreover, as $\mathbf{\Psi}_{Z\upsilon}(x_1, x_2) > 0$ with probability one and $\mu(x_1, x_2) \neq 0$ with probability one when $\mathbf{\Pi}_0\rho_{\upsilon u} \neq 0,$ hence the non centrality parameter in the asymptotic distribution of the statistics is positive with probability one when $\mathbf{\Pi}_0\rho_{\upsilon u} \neq 0$. This suggests that all tests may still exhibit when identification is weak. This is conform with the necessary and sufficient condition for test consistency which was that $\mathbf{\Pi}\rho_{\upsilon u} \neq 0$ when $\mathit{\Pi}$ is fixed (does not depend on the sample size as it the case here). However, if $\mathbf{\Pi}_0\rho_{\upsilon u} = 0,$ the limiting distribution of all statistics is the same under the null hypothesis and the alternative hypothesis. As a result, the power of the tests cannot exceed their nominal level in that case. This is particularly the case when $\mathit{\Pi}_0 = 0$ (complete non identification of β). An interesting observation also is that even if the parameter of the regressor which exogeneity is not being tested in the structural is completely unidentified ($\mathbf{\Gamma}_0 = 0$), the tests may still have power as long as $\mathbf{\Pi}_0\rho_{\upsilon u} \neq 0$ [see Panels $(\mathbf{B})\&(\mathbf{C})$ in Figure 1]. In the other side, if $\mathbf{\Pi}_0\rho_{\upsilon u} \neq 0,$ the power of all tests is low even when θ is identified or close so [as an illustration of this, see Panel $(\mathbf{D})$ in Figure 1]. We now study in Section 3.3, the behaviour of the tests in a Monte Carlo experiment.

3.3. Size and power comparison

We consider the following data generating process (DGP):

$$\begin{aligned} y &= Y_1\beta_1 + Y_2\beta_2 + W\theta + u\,, \\ (Y_1, Y_2, W) &= Z\,(\mathbf{\Pi}_1, \mathbf{\Pi}_2, \mathbf{\Gamma}) + (\upsilon_1, \upsilon_2, \xi)\,, \end{aligned} \tag{3.3}$$

where $Y = [Y_1\,,\, Y_2]$ is a $n \times 2$ matrix of regressors of interest. W (here a $n \times 1$ vector)[8] is the endogenous variable which exogeneity is not being tested. Z contains l instruments each generated i.i.d $\mathbf{N}(0,\, 1)$ and is kept fix within experiment. So, $\mathbf{\Pi}_1,\ \mathbf{\Pi}_2$ and $\mathbf{\Gamma}$ are l-dimensional vectors.

The errors $(u,\, \upsilon_1,\, \upsilon_2,\, \xi)$ are generated such that:

$$\begin{aligned} u_i &= (1 + \rho_{\upsilon_1}^2 + \rho_{\upsilon_2}^2 + \rho_{\xi}^2)^{-1/2}(\varepsilon_{1i} + \rho_{\upsilon_1}\varepsilon_{2i} + \rho_{\upsilon_2}\varepsilon_{3i} + \rho_{\xi}\varepsilon_{4i}), \\ \upsilon_{1i} &= (1 + \rho_{\upsilon_1}^2)^{-1/2}(\rho_{\upsilon_1}\varepsilon_{1i} + \varepsilon_{2i}),\ \upsilon_{2i} = (1 + \rho_{\upsilon_2}^2)^{-1/2}(\rho_{\upsilon_2}\varepsilon_{1i} + \varepsilon_{3i}), \end{aligned}$$

[8] Note that the results are qualitatively the same when W contains more than one regressor.

$$\xi_i \;=\; (1+\rho_\xi^2)^{-1/2}(\rho_\xi \varepsilon_{1i}+\varepsilon_{4i}),\,(\varepsilon_{1i},\,\varepsilon_{2i},\,\varepsilon_{3i},\,\varepsilon_{4i})' \overset{i.i.d}{\sim} \mathbf{N}\,(0,\,I_4) \tag{3.4}$$

for all $i = 1, \ldots, n,\ -1 \leq \rho_{\upsilon_1} \leq 1,\ \rho_{\upsilon_2} = \rho_{\upsilon_1}/\sqrt{3},$ and ρ_ξ is kept at $\rho_\xi = 0.8$. From this parametrization, the partial null exogeneity of Y is then expressed as $\mathrm{H}_0^p :\ \rho_{\upsilon_1} = 0$. As seen from (3.4), ξ is not correlated with $(\upsilon_1,\,\upsilon_2)$ under $\mathrm{H}_0^p,$ but is under the alternative hypothesis. To extend the model to cases where ξ is *locally correlated* with $(\upsilon_1,\,\upsilon_2),$ as required Assumption **2.4**, we weakened the non correlation assumption between ξ and $(\upsilon_1,\,\upsilon_2)$. The results for this setup are presented in Figure 5 of Appendix B. They indicate that the tests are still valid even for moderate correlation between ξ and $(\upsilon_1,\,\upsilon_2)$.

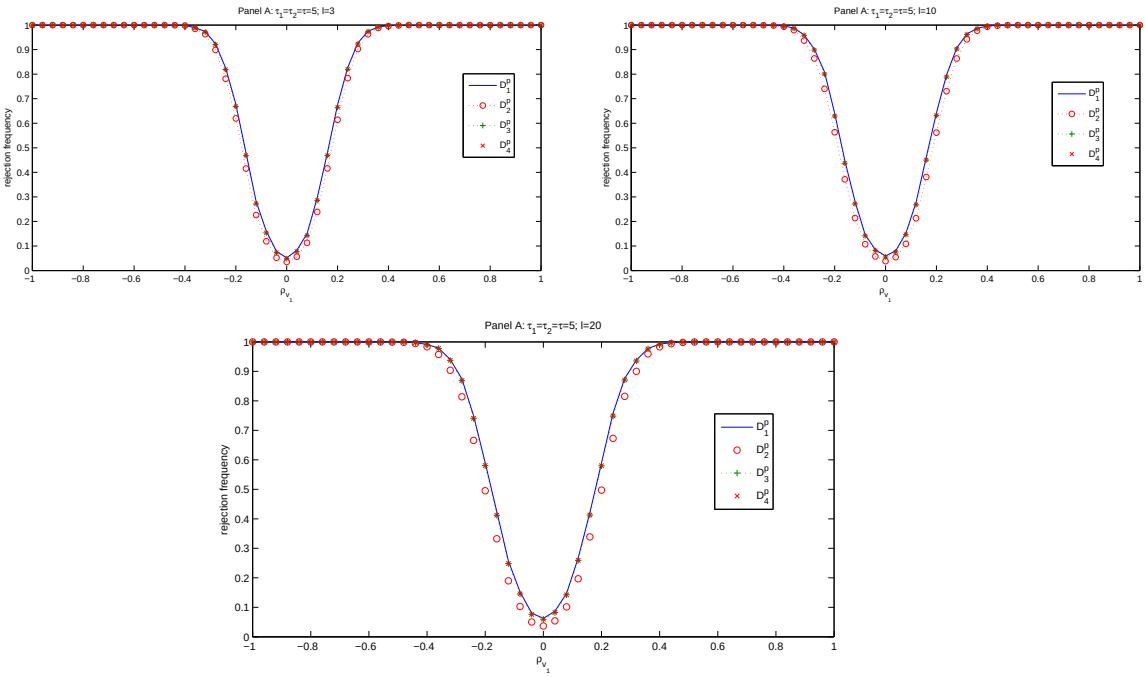
The values of $\beta_1,\ \beta_2$ and θ are set at $2,\ -3$ and $1/2,$ respectively. $\mathbf{\Pi}_1,\ \mathbf{\Pi}_2$ and $\mathbf{\Gamma}$ are chosen as: $\mathbf{\Pi}_1 = \tau_1\mathbf{\Pi}_{01},\ \mathbf{\Pi}_2 = \tau_2\mathbf{\Pi}_{02},\ \mathbf{\Gamma} = \tau\mathbf{\Gamma}_0,$ where $[\mathbf{\Pi}_{01},\,\mathbf{\Pi}_{02},\,\mathbf{\Gamma}_0]$ is obtained by taking the first three columns of the identity matrix of dimension l. To account for strong, partial and weak identification of model parameters, we consider six panels for the values of τ_1, τ_2 and τ as follows: $(\mathbf{A})\ \tau_1 = \tau_2 = \tau = 5,$ i.e. $\beta_1,\ \beta_2$ and θ are identified; $(\mathbf{B})\ \tau_1 = \tau_2 = 5,\ \tau = 0,$ so, β_1 and β_2 are identified but θ is not (partial identification); $(\mathbf{C})\ \tau_1 = 5,\ \tau_2 = \tau = \frac{0.1}{\sqrt{n}},$ i.e. β_1 is identified but β_2 and θ are weakly identified; $(\mathbf{D})\ \tau_1 = \tau_2 = \frac{0.1}{\sqrt{n}},\ \tau = 5,$ hence θ is identified but β_1 and β_2 are weakly identified; $(\mathbf{E})\ \tau_1 = \tau_2 = \frac{0.5}{\sqrt{n}},\ \tau = \frac{1}{\sqrt{n}},$ i.e., all model parameters are weakly identified; and finally $(\mathbf{F})\ \tau_1 = \tau_2 = 0,\ \tau = \frac{1}{\sqrt{n}}\ :\ \beta_1$ and β_2 are completely non identified (irrelevant instruments), and θ is weakly identified. The number of instruments l belong to $\{3,\,10,\,20\}$. Since we have $m = 3$ endogenous regressors in (3.3), $l = 3$ corresponds to the usual "just-identified" setup, while $l > 3$ corresponds to the "overidentification". The simulations are run with sample sizes 100 and $300,$ while the number of replications is $N = 10,000.$ In all cases, the nominal level is set at $5\,\%$.

Figures 1- 2 presents the power curves of the tests for $n = 100,$ while Figures 3-4 in Appendix B is for $n = 300.$ The results are qualitatively the same in terms of level control in both cases. However, the power improves substantially when $n = 300,$ as expected. First, we observe that all tests have correct level whether identification is strong, partial or weak. Furthermore, $\mathscr{D}_1^p$ and $\mathscr{D}_4^p$ have approximately a good level even when IVs are weak [for example, see Figure 2 below and Figure 4 in Appendix B where identification is weak]. However, the same figures show clearly that $\mathscr{D}_2^p$ and $\mathscr{D}_3^p$ are overly conservative. In the same vain, all tests have similar power when identification is strong strong (see Panel $(\mathbf{A})$ in Figure 1& 3), but $\mathscr{D}_1^p$ and $\mathscr{D}_4^p$ exhibit more power than $\mathscr{D}_2^p$ and

$\mathscr{D}_3^p$ when identification is partial or weak. In addition, the results confirm that the tests have power when the parameter of the regressors which exogeneity is tested (here β) is identified (for example, see Panel $(\mathbf{B})$& $(\mathbf{C})$ in Figure 1). But power is low when β is weakly identified, even when θ is strongly identified (see Panel $(\mathbf{D})$ Figure 1). Overall, the recommendation is to use the tests $\mathscr{D}_1^p$ and $\mathscr{D}_4^p$ which outperform the others in all possible configurations of model identification.

Figure 1. Size and power at nominal level 5% when identification is strong or partial, $n = 100$

Strong identification: Panel A



Partial identification of β : Panels B and C

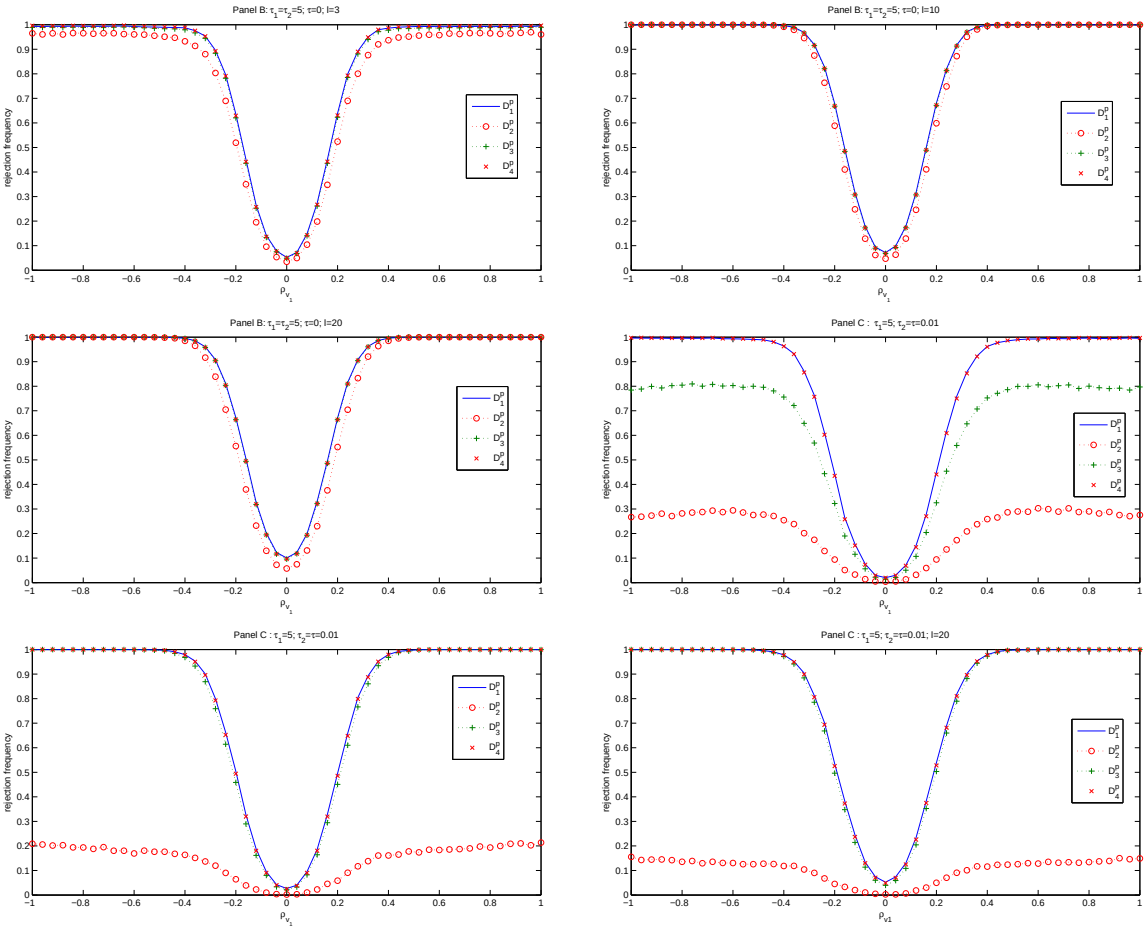
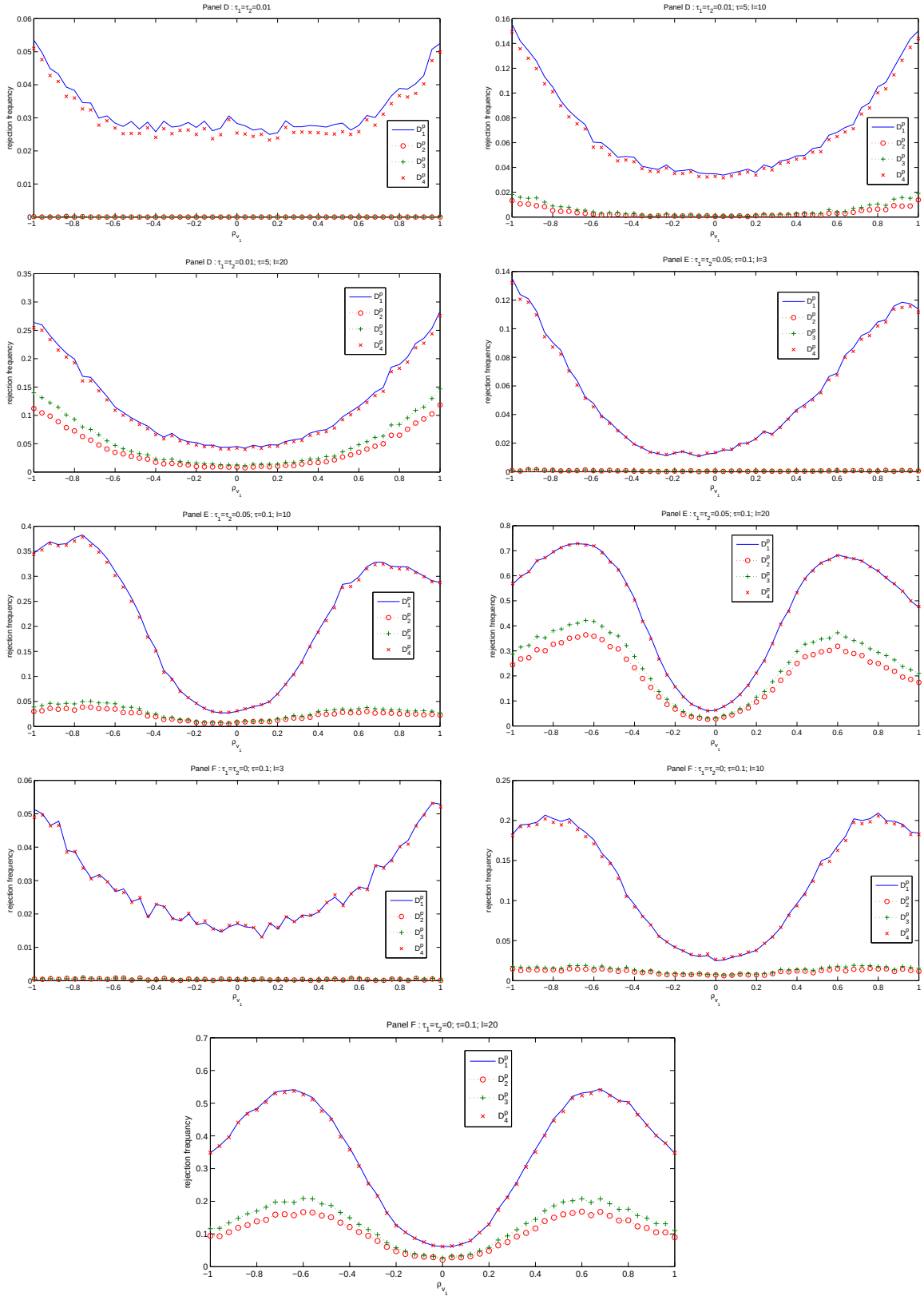


Figure 2. Size and power at nominal level 5% when identification is weak, $n = 100$

Partial identification of θ and complete weak identification of all parameters: Panels D, E and F



4. Empirical illustration

We consider the return to education model from Card (1995) in Example **2.1**. The first-stage specifications for $educ$ and $(exper,\ exper^2)$ are given by

$$educ_i \quad = \quad Z'\mathbf{\Pi} + Z_{1i}'\delta_1 + \upsilon_i,\ (exper_i,\ exper_i^2) = Z_i'\mathbf{\Gamma} + Z_{1i}'\delta_1 + \xi_i,\ i = 1,\ \ldots,\ n \qquad (4.1)$$

where Z_1 and Z are the same as in (2.4). In Example **2.1**, we found that DWH-tests rejected the joint exogeneity of $(educ,\ exper,\ exper^2),$ but we do not know if some regressors are exogenous. In this application, we want to test the exogeneity of $educ$ and $(exper,\ exper^2)$ separately. So, two null hypotheses are considered: $(\mathbf{i})$ H_0^p : $\mathbf{cov}(\upsilon_i,\ u_i) = 0$ for all i (partial exogeneity of $educ$) and $(\mathbf{ii})$ H_0^p : $\mathbf{cov}(\xi_i,\ u_i) = 0$ for all i [partial exogeneity of $(exper,\ exper^2)$], where u is the structural error term in (2.4). Note that in the setup for $(\mathbf{i}),\ \xi$ may be correlated with u [i.e. $(exper,\ exper^2)$ may be endogenous], while in those for $(\mathbf{ii}),\ \upsilon$ may be correlated with u (i.e. $educ$ may be endogenous).

Table 1 reports the outcomes of the DWH-tests and the $\mathscr{D}^p$ tests proposed in this paper. The DWH-tests are run under the assumption that the regressors not being tested are exogenous, while the $\mathscr{D}^p$ tests do not require this questionable restriction. It is important to observe that because $exper$ is generated as $exper = qge - 6 - educ$, we have $\mathbf{cov}(exper_i, u_i) = -\mathbf{cov}(educ_i, u_i),$ as age is exogenous. So, any valid procedure that rejects the partial exogeneity of $educ$ should also reject those of $exper$. This is not however the case for the DWH-tests, as they all fail to rejected the partial exogeneity of $(exper,\ exper^2)$. This result is not surprising because $educ$ is likely endogenous and DWH procedures do not account for that when testing the exogeneity of $(exper,\ exper^2)$. The outcomes of the $\mathscr{D}^p$ tests indicate strong evidence against the exogeneity of both $educ$ and $(exper,\ exper^2)$ as showed Table 1. Overall, these results underscore: $(\mathbf{1})$ how the use of DWH tests to assess partial exogeneity hypotheses may be misleading, and $(\mathbf{2})$ the relevance of using $\mathscr{D}^p$ tests when checking for partial exogeneity.

Table 1. Testing for partial exogeneity of $educ$ and $(exper,\ exper^2)$

	DWH-tests of the exogeneity of $educ$ by keeping $(exper,\ exper^2)$ exogenous		
Statistics	Sample value	p-value	decision
$\mathcal{T}_2$	36.62	0.000	reject
$\mathcal{H}_1$	18.41	0.000	reject
$\mathcal{H}_2$	18.58	0.000	reject
$\mathcal{H}_3$	36.08	0.000	reject
	DWH-tests of the exogeneity of $(exper,\ exper^2)$ by keeping $educ$ exogenous		
Statistics	Sample value	p-value	decision
$\mathcal{T}_2$	1.44	0.236	do not reject
$\mathcal{H}_1$	2.87	0.238	do not reject
$\mathcal{H}_2$	2.89	0.236	do not reject
$\mathcal{H}_3$	2.89	0.235	do not reject
	$\mathscr{D}^p$-tests of the exogeneity of $educ$		
Statistics	Sample value	p-value	decision
$\mathscr{D}_1^p$	27.52	0.000	reject
$\mathscr{D}_2^p$	9.86	0.002	reject
$\mathscr{D}_3^p$	9.91	0.002	reject
$\mathscr{D}_4^p$	27.23	0.000	reject
	$\mathscr{D}^p$-tests of the exogeneity of $(exper,\ exper^2)$		
Statistics	Sample value	p-value	decision
$\mathscr{D}_1^p$	99.05	0.000	reject
$\mathscr{D}_2^p$	151.51	0.000	reject
$\mathscr{D}_3^p$	170.94	0.000	reject
$\mathscr{D}_4^p$	181.38	0.000	reject

5. Conclusion

In this paper, we propose alternative tests for assessing partial exogeneity in a linear IV regression. The tests are easy to implement as they only require OLS and IV regressions. We provide an analysis of their asymptotic behaviour (level and power) which shows that all tests are valid (level is controlled) whether model parameters are identified or not. So, the proposed tests robust to weak instruments. Moreover, our analysis indicates that test consistency may still hold over a wide range of cases where the overall identification fails, provided partial identification is satisfied. However, all tests have low power when model parameters are completely not identified.

A Monte Carlo experiment confirms our theoretical results. We illustrate our theoretical finding through the workhorse example of returns to education from Card (1995). Our results clearly indicate that standard exogeneity tests of the type proposed by Durbin (1954), Wu (1973, 1974), and Hausman (1978) are not appropriate to assess partial exogeneity hypotheses, as they are valid only when the regressors not being tested are exogenous. For example, we find these tests fail to rejected the exogeneity of experience variables in this model if education is assumed exogenous. In contrast, all proposed tests in this paper find strong evidence against the exogeneity of both education and experience variables, separately. Overall, this application underscores the relevance of using $\mathscr{D}^p$-tests when checking for partial exogeneity.

APPENDIX

A. Proofs

PROOF OF LEMMA **2.5** Assume that rank($\mathbf{\Pi}_{YW}$) $= m$. First, write $\tilde{u}$ and $\hat{u}_*$ as:

$$\tilde{u} \;=\; u + (W - \tilde{W})\theta = u + M_Z\xi\theta + \hat{\sigma}_\theta P_Z\hat{u}_*,\; \hat{u}_* = M_{\bar{X}}u_* = M_{\bar{X}}u + M_{\bar{X}}M_Z\xi\theta \qquad \text{(A.1)}$$

where $\bar{X} = [Y, \bar{W}]$ and $\hat{\sigma}_\theta = \hat{\sigma}_{u\xi}\theta/\hat{\sigma}_u^2$. Hence, we have $Z'\tilde{u}/n = Z'u/n + \hat{\sigma}_\theta Z'\hat{u}_*/n$ and $Z'\hat{u}_*/n = Z'M_{\bar{X}}u/n + Z'M_{\bar{X}}M_Z\xi\theta/n$. When Assumptions **2.2-2.4** are satisfied and if further H_0 holds, then $\bar{X}'u/n \xrightarrow{p} (\sigma'_{vu}, 0)' = 0$ and

$$\bar{X}'\bar{X}/n \xrightarrow{p} Q_{\bar{X}} = \begin{pmatrix} \mathbf{\Pi}'Q_Z\mathbf{\Pi} + \Sigma_v & \mathbf{\Pi}'Q_Z\mathbf{\Gamma} \\ \mathbf{\Gamma}'Q_Z\mathbf{\Pi} & \mathbf{\Gamma}'Q_Z\mathbf{\Gamma} \end{pmatrix} > 0,\; Z'\bar{X}/n \xrightarrow{p} Q_{Z\bar{X}} = \begin{pmatrix} Q_Z\mathbf{\Pi} & Q_Z\mathbf{\Gamma} \end{pmatrix}.$$

This then implies that $Z'M_{\bar{X}}u/n = Z'u/n - (Z'\bar{X}/n)(\bar{X}'\bar{X}/n)^{-1}(\bar{X}'u/n) \xrightarrow{p} 0$. Since $v'\xi/n \xrightarrow{p} 0$ from Assumption **2.4**-(**a**), we also get $Z'M_{\bar{X}}M_Z\xi\theta/n \xrightarrow{p} 0$ so that $Z'\hat{u}_*/n = Z'M_{\bar{X}}u/n + Z'M_{\bar{X}}M_Z\xi\theta/n \xrightarrow{p} 0$. So, we have $\hat{\sigma}_{u\xi} = \hat{u}_*'W/(n-m) - (\hat{u}_*'Z/n)(Z'Z/n)^{-1}(ZW/(n-m)) = \hat{u}_*'W/(n-m) + o_p(1) \xrightarrow{p} \sigma^*_{u\xi} = \sigma_{u\xi} + \theta'\Sigma_\xi$ and $\hat{\sigma}_u^2 \xrightarrow{p} \sigma_u^{*^2} = \sigma_u^2 + \sigma_{u\xi}\theta$. From Assumption **2.4**-(**b**), we have $\sigma_{u\xi} = -\theta'\Sigma_\xi$ so that $\sigma^*_{u\xi} = 0$ and $\sigma_u^{*^2} = \sigma_u^2 - \theta'\Sigma_\xi\theta > 0$ (by Assumption **2.2**). Hence, we have $\hat{\sigma}_\theta = \hat{\sigma}_{u\xi}\theta/\hat{\sigma}_u^2 \xrightarrow{p} \sigma_\theta = 0$ and $Z'\tilde{u}/n = Z'u/n + \hat{\sigma}_\theta Z'\hat{u}_*/n \xrightarrow{p} 0$. We shall now show that $\tilde{W}'\tilde{u}/n \xrightarrow{p} 0$ and $Y'\tilde{u}/n \xrightarrow{p} 0$. Observe first that $\tilde{W}'\tilde{u}/n = \tilde{\mathbf{\Gamma}}'Z'\tilde{u}/n$. Since $\tilde{\mathbf{\Gamma}} \xrightarrow{p} \mathbf{\Gamma}$, and from (**??**) $Z'\tilde{u}/n \xrightarrow{p} 0$, we have $\tilde{W}'\tilde{u}/n \xrightarrow{p} 0$. By the same way, we get $Y'\tilde{u}/n = Y'(u + M_Z\xi\theta + \hat{\sigma}_\theta P_Z\hat{u}_*)/n \xrightarrow{p} \sigma_{vu}$. As $\sigma_{vu} = 0$ under Assumption H_0, it is clear that $Y'\tilde{u}/n \xrightarrow{p} 0$. The proof is similar for weak values of $\mathbf{\Pi}_{YW}$, i.e., $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0, \mathbf{\Gamma}_0]$.

□

PROOF OF LEMMA **2.6** Firstly, from Lemma **2.6**, we have $\tilde{u} = u + M_Z\xi\theta + \hat{\sigma}_\theta P_Z\hat{u}_* = u + M_Z\xi\theta + P_Z P_{\hat{u}_*}W\theta + o_p(1)$, where $P_{\hat{u}_*} = \hat{u}_*(\hat{u}_*'\hat{u}_*)^{-1}\hat{u}_*'$ is the projection matrix in the space spanned by the residuals $\hat{u}_*$. So, we can write $Z^{\perp'}\tilde{u}^\perp/\sqrt{n}$ and $v^{\perp'}\tilde{u}^\perp/\sqrt{n}$ as:

$$Z^{\perp'}\tilde{u}^\perp/\sqrt{n} \;=\; Z^{\perp'}u^\perp/\sqrt{n} + Z^{\perp'}M_Z\xi\theta/\sqrt{n} + Z^{\perp'}P_Z P_{\hat{u}_*}W\theta/\sqrt{n} \qquad \text{(A.2)}$$

$$v^{\perp'}\tilde{u}^\perp/\sqrt{n} \;=\; v^{\perp'}u^\perp/\sqrt{n} + v^{\perp'}M_Z\xi\theta/\sqrt{n} + v^{\perp'}P_Z P_{\hat{u}_*}W\theta/\sqrt{n}. \qquad \text{(A.3)}$$

Observe that $Z^{\perp'}M_Z = Z'M_Z = 0$ and $v^{\perp'}M_Z\xi\theta/\sqrt{n} = o_p(1)$ when Assumption **2.4**-(**a**) holds, hence

(A.2)-(A.3) become:

$$Z^{\perp'}\tilde{u}^{\perp}/\sqrt{n} = Z^{\perp'}u^{\perp}/\sqrt{n} + Z^{\perp'}P_Z P_{\hat{u}_*}W\theta/\sqrt{n} \quad (A.4)$$

$$v^{\perp'}\tilde{u}^{\perp}/\sqrt{n} = v^{\perp'}u^{\perp}/\sqrt{n} + v^{\perp'}P_Z P_{\hat{u}_*}W\theta/\sqrt{n} + o_p(1). \quad (A.5)$$

As $Z'\hat{u}_*/n = o_p(1)$, $v^{\perp'}Z/n = o_p(1)$ and $\hat{u}_*'W/\sqrt{n} = O_p(1)$, we have $v^{\perp'}P_Z P_{\hat{u}_*}W\theta/\sqrt{n} = (v^{\perp'}Z/n)(Z'Z/n)^{-1})(Z'\hat{u}_*/n)(\hat{u}_*'\hat{u}_*/n)^{-1}(\hat{u}_*'W/\sqrt{n}) = 0_p(1)$. Moreover, since $\tilde{\mathbf{\Gamma}} \xrightarrow{p} \mathbf{\Gamma}$ (with $\mathbf{\Gamma} = 0$ when $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0, \mathbf{\Gamma}_0]$), we have $v'\tilde{W}/n = (v'Z/n)\tilde{\mathbf{\Gamma}} \xrightarrow{p} 0$ so that $v^{\perp'}u^{\perp}/\sqrt{n} = v'u/\sqrt{n} - (v'\tilde{W}/n)(\tilde{W}'\tilde{W}/n)^{-1}\tilde{\mathbf{\Gamma}}(Z'/\sqrt{n}) = v'u/\sqrt{n} + o_p(1)$. By the same way, we get $Z^{\perp'}P_Z P_{\hat{u}_*}W\theta/\sqrt{n} = (Z^{\perp'}Z/n)(Z'Z/n)^{-1})(Z'\hat{u}_*'/n)(\hat{u}_*'\hat{u}_*/n)^{-1}(\hat{u}_*'W\theta/\sqrt{n}) = o_p(1)$ so that we can express (A.4)-(A.5) as:

$$\begin{pmatrix} Z^{\perp'}\tilde{u}^{\perp}/\sqrt{n} \\ v^{\perp'}\tilde{u}^{\perp}/\sqrt{n} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{1n} & 0 \\ 0 & I_{m_y} \end{pmatrix} \begin{pmatrix} Z'u/\sqrt{n} \\ v'u/\sqrt{n} \end{pmatrix} + o_p(1) \quad (A.6)$$

where $\mathbf{A}_{1n} = I_l - (Z'Z/n)\tilde{\mathbf{\Gamma}}(\tilde{\mathbf{\Gamma}}'(Z'Z/n)\tilde{\mathbf{\Gamma}})^{-1}\tilde{\mathbf{\Gamma}}'$ and $\begin{pmatrix} Z'u/\sqrt{n} \\ v'u/\sqrt{n} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_{Zu} \\ \psi_{vu} \end{pmatrix} \sim \mathbf{N}\left[0, \sigma_u^2 \begin{pmatrix} Q_Z & 0 \\ 0 & \Sigma_v \end{pmatrix}\right]$ by Assumption **2.3**. We shall now distinguish two cases: $(\mathbf{1})$ rank$(\mathbf{\Pi}_{YW}) = m$, and $(\mathbf{2})$ $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0, \mathbf{\Gamma}_0]$.

$(\mathbf{1})$ Suppose first that rank$(\mathbf{\Pi}_{YW}) = m$. Then, $\mathbf{A}_{1n} \xrightarrow{p} \mathbf{A}_1 = I_l - Q_Z\mathbf{\Gamma}(\mathbf{\Gamma}'Q_Z\mathbf{\Gamma})^{-1}\mathbf{\Gamma}' = Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}}Q_Z^{-1/2}$ and from (A.6) we have

$$\begin{pmatrix} Z^{\perp'}\tilde{u}^{\perp}/\sqrt{n} \\ v^{\perp'}\tilde{u}^{\perp}/\sqrt{n} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_{Z^{\perp}\tilde{u}} \\ \psi_{v^{\perp}\tilde{u}} \end{pmatrix} = \begin{pmatrix} Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}}Q_Z^{-1/2} & 0 \\ 0 & I_{m_y} \end{pmatrix} \begin{pmatrix} \psi_{Zu} \\ \psi_{vu} \end{pmatrix}$$

$$\sim \mathbf{N}\left[0, \sigma_u^2 \begin{pmatrix} Q_{Z^{\perp}} & 0 \\ 0 & \Sigma_v \end{pmatrix}\right], Q_{Z^{\perp}} = Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}}Q_Z^{1/2}.$$

$(\mathbf{2})$ Suppose now that $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0, \mathbf{\Gamma}_0]$ and write $\sqrt{n}\tilde{\mathbf{\Gamma}} = \mathbf{\Gamma}_0 + (Z'Z/n)^{-1}(Z'\xi/\sqrt{n}) - \hat{\sigma}_\theta(Z'Z/n)^{-1}(Z'\hat{u}_*/\sqrt{n})$. From the proof in Lemma **2.5**, we have $\hat{\sigma}_\theta = \hat{\sigma}_{u\xi}\theta/\hat{\sigma}_u^2 \xrightarrow{p} \sigma_\theta = 0$. From Assumption **2.3**, we also have $(Z'Z/n)^{-1}(Z'\xi/\sqrt{n}) \xrightarrow{d} Q_Z^{-1}\psi_{Z\xi}$. We now focus on $Z'\hat{u}_*/\sqrt{n}$. Let us decompose $M_{\bar{X}}$ as $M_{\bar{X}} = M_{\bar{W}} - P_{M_{\bar{W}}Y}$ and write $Z'\hat{u}_*/\sqrt{n}$ as:

$$\begin{aligned} Z'\hat{u}_*/\sqrt{n} = Z'M_{\bar{X}}u_*/\sqrt{n} &= Z'M_{\bar{W}}u_*/\sqrt{n} - (Z'M_{\bar{W}}Y/n)(Y'M_{\bar{W}}Y/n)^{-1}(Y'M_{\bar{W}}u_*/\sqrt{n}) \\ &= [I_l - (Z'M_{\bar{W}}Y/n)(Y'M_{\bar{W}}Y/n)^{-1}\mathbf{\Pi}']Z'M_{\bar{W}}u_*/\sqrt{n} + \\ &\quad (Z'M_{\bar{W}}Y/n)(Y'M_{\bar{W}}Y/n)^{-1}v'u_*/\sqrt{n}. \end{aligned} \quad (A.7)$$

Since $Z'M_{\bar{W}}u_*/\sqrt{n} = [I_l - (Z'Z/n)\hat{\mathbf{\Gamma}}(\hat{\mathbf{\Gamma}}'(Z'Z/n)\hat{\mathbf{\Gamma}})^{-1}\hat{\mathbf{\Gamma}}']Z'u_*/\sqrt{n} = [I_l - (Z'Z/n)\hat{\mathbf{\Gamma}}(\hat{\mathbf{\Gamma}}'(Z'Z/n)\hat{\mathbf{\Gamma}})^{-1}\hat{\mathbf{\Gamma}}']Z'u/\sqrt{n}$ and $\upsilon'u_*/\sqrt{n} = \upsilon'u/\sqrt{n} + \upsilon'M_Z\xi\theta/\sqrt{n} = \upsilon'u/\sqrt{n} + o_p(1)$ [because $\upsilon'M_Z\xi\theta/\sqrt{n} = o_p(1)$ under H_0], we can express (A.7) as:

$$Z'\hat{u}_*/\sqrt{n} = \mathbf{A}_{2n}Z'u/\sqrt{n} + \mathbf{A}_{3n}\upsilon'u/\sqrt{n} + o_p(1) \tag{A.8}$$

where $\mathbf{A}_{2n} = [I_l - (Z'M_{\bar{W}}Y/n)(Y'M_{\bar{W}}Y/n)^{-1}\mathbf{\Pi}'][I_l - (Z'Z/n)\hat{\mathbf{\Gamma}}(\hat{\mathbf{\Gamma}}'(Z'Z/n)\hat{\mathbf{\Gamma}})^{-1}\hat{\mathbf{\Gamma}}']$ and $\mathbf{A}_{3n} = (Z'M_{\bar{W}}Y/n)(Y'M_{\bar{W}}Y/n)^{-1}$. As $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0, \mathbf{\Gamma}_0]$, we find: $Z'M_{\bar{W}}Y/n \xrightarrow{p} 0$ $Y'M_{\bar{W}}Y/n \xrightarrow{p} \Sigma_\upsilon$, $\mathbf{A}_{2n} \xrightarrow{p} \mathbf{A}_2 = Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}_\xi}Q_Z^{-1/2}$ where $\mathbf{\Gamma}_\xi = \mathbf{\Gamma}_0 + Q_Z^{-1}\psi_{Z\xi}$, and $\mathbf{A}_{3n} \xrightarrow{p} 0$. Hence, we get and $Z'\hat{u}_*/\sqrt{n} \xrightarrow{d} Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}_\xi}Q_Z^{-1/2}\psi_{Zu}$ and $\sqrt{n}\tilde{\mathbf{\Gamma}} \xrightarrow{d} \mathbf{\Gamma}(\psi_{Z\xi}) = Q_Z^{-1/2}(Q_Z^{1/2}\mathbf{\Gamma}_\xi - \sigma_\theta M_{Q_Z^{1/2}\mathbf{\Gamma}_\xi}Q_Z^{-1/2}\psi_{Zu}) \equiv \mathbf{\Gamma}_\xi$ (since $\sigma_\theta = 0$). Moreover, we have $\mathbf{A}_{1n} \xrightarrow{d} \mathbf{A}_1 = I_l - Q_Z\mathbf{\Gamma}(\psi_{Z\xi})(\mathbf{\Gamma}(\psi_{Z\xi})'Q_Z\mathbf{\Gamma}(\psi_{Z\xi}))^{-1}\mathbf{\Gamma}(\psi_{Z\xi})' = Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}$ and (A.6) then implies that

$$\begin{pmatrix} Z^{\perp'}\tilde{u}^\perp/\sqrt{n} \\ \upsilon^{\perp'}\tilde{u}^\perp/\sqrt{n} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_{Z^\perp\tilde{u}} \\ \psi_{\upsilon^\perp\tilde{u}} \end{pmatrix} = \begin{pmatrix} Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2} & 0 \\ 0 & I_{m_y} \end{pmatrix} \begin{pmatrix} \psi_{Zu} \\ \psi_{\upsilon u} \end{pmatrix}. \tag{A.9}$$

Because $\psi_{Z^\perp\tilde{u}} = Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zu}$, it is clear that $Q_Z^{-1/2}\psi_{Z^\perp\tilde{u}}$ is independent of $Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})$. Since Q_Z is fixed, $\psi_{Z^\perp\tilde{u}}$ is also independent of $Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})$. So, conditionally on $Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi}) = Q_Z^{1/2}\mathbf{\Gamma}(x_2)$, (A.9) implies that

$$\begin{pmatrix} \psi_{Z^\perp\tilde{u}} \\ \psi_{\upsilon^\perp\tilde{u}} \end{pmatrix} |_{Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \sim \mathbf{N}\left[0, \sigma_u^2\mathbf{diag}(Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(x_2)}Q_Z^{1/2}, \Sigma_\upsilon)\right]. \tag{A.10}$$

By integrating (A.10) with respect to all possible realization of $\psi_{Z\xi}$, the result follows.

□

PROOF OF LEMMA **3.1** (**i**) Assume first that $\sigma_{\upsilon u} = 0$. We have

$$\begin{aligned}
\hat{\beta}_{LS} - \tilde{\beta}_{IV} &= (Y^{\perp'}Y^\perp/n)^{-1}Y^{\perp'}\tilde{u}^\perp/n - (Y^{\perp'}P_{Z^\perp}Y^\perp/n)^{-1}Y^{\perp'}P_{Z^\perp}\tilde{u}^\perp/n, \\
\sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV}) &= (Y^{\perp'}Y^\perp/n)^{-1}Y^{\perp'}\tilde{u}^\perp/\sqrt{n} - (Y^{\perp'}P_{Z^\perp}Y^\perp/n)^{-1}Y^{\perp'}P_{Z^\perp}\tilde{u}^\perp/\sqrt{n}, \\
Y^{\perp'}Y^\perp/n &= Y'Y/n - (Y'Z/n)\sqrt{n}\tilde{\mathbf{\Gamma}}[\sqrt{n}\tilde{\mathbf{\Gamma}}'(Z'Z/n)\sqrt{n}\tilde{\mathbf{\Gamma}}]^{-1}\sqrt{n}\tilde{\mathbf{\Gamma}}'(Z'Y/n), \\
Y^{\perp'}P_{Z^\perp}Y^\perp/n &= (Y'M_{\tilde{W}}Z/n)(Z'M_{\tilde{W}}Z/n)^{-1}(Z'M_{\tilde{W}}Y/n), \\
Y^{\perp'}\tilde{u}^\perp/n &= Y'\tilde{u}/n - (Y'\tilde{W}/n)(\tilde{W}'\tilde{W}/n)^{-1}(\tilde{W}'\tilde{u}/n), \\
Y^{\perp'}P_{Z^\perp}\tilde{u}^\perp/n &= (Y'M_{\tilde{W}}Z/n)(Z'M_{\tilde{W}}Z/n)^{-1}(Z'M_{\tilde{W}}\tilde{u}/n).
\end{aligned}$$

From Lemmas **2.5-2.6**, it easy to see that $Y^{\perp'}Y^{\perp}/n \xrightarrow{p} \mathbf{\Sigma}_{\pi} = \mathbf{\Pi}'Q_{Z^{\perp}}\mathbf{\Pi} + \Sigma_{v}$, $Y^{\perp'}P_{Z^{\perp}}Y^{\perp}/n \xrightarrow{p} \tilde{\mathbf{\Sigma}}_{\pi} = \mathbf{\Pi}'Q_{Z^{\perp}}\mathbf{\Pi}$, $Y^{\perp'}\tilde{u}^{\perp}/n = \mathbf{\Pi}'Z^{\perp'}\tilde{u}^{\perp}/n + v^{\perp'}\tilde{u}^{\perp}/n \xrightarrow{p} 0$, and $Y^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp}/n = \mathbf{\Pi}'Z^{\perp'}\tilde{u}^{\perp}/n + v^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp}/n \xrightarrow{p} 0$. So, we find $\hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{p} 0$. Moreover, from results in by Lemma **2.6**, we have $Y^{\perp'}\tilde{u}^{\perp}/\sqrt{n} = \mathbf{\Pi}'Z^{\perp'}\tilde{u}^{\perp}/\sqrt{n} + v^{\perp'}\tilde{u}^{\perp}/\sqrt{n} \xrightarrow{d} \mathbf{\Pi}'\psi_{Z^{\perp}\tilde{u}} + \psi_{v^{\perp}\tilde{u}} = \mathbf{\Pi}'Q_{Z^{\perp}}Q_{Z}^{-1}\psi_{Zu} + \psi_{vu}$, and $Y^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp}/\sqrt{n} = \mathbf{\Pi}'Z^{\perp'}\tilde{u}^{\perp}/\sqrt{n} + v^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp}/\sqrt{n} \xrightarrow{d} \mathbf{\Pi}'\psi_{Z^{\perp}\tilde{u}} = \mathbf{\Pi}'Q_{Z^{\perp}}Q_{Z}^{-1}\psi_{Zu}$. So, from Lemma **2.6**-(**i**), we get $\sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV}) \xrightarrow{d} \Psi_{\beta}$ where

$$\Psi_{\beta} = \mathbf{\Sigma}_{\pi}^{-1}(\mathbf{\Pi}'Q_{Z^{\perp}}Q_{Z}^{-1}\psi_{Zu} + \psi_{vu}) - \tilde{\mathbf{\Sigma}}_{\pi}^{-1}\mathbf{\Pi}'Q_{Z^{\perp}}Q_{Z}^{-1}\psi_{Zu} \sim \mathbf{N}\left[0,\, \sigma_u^2(\tilde{\mathbf{\Sigma}}_{\pi}^{-1} - \mathbf{\Sigma}_{\pi}^{-1})\right].$$

(**ii**) Suppose now that $\sigma_{vu} \neq 0$. It is easy to see from the above proof that $Y^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp}/n \xrightarrow{p} 0$, $Y^{\perp'}\tilde{u}^{\perp}/n \xrightarrow{p} \sigma_{vu} \neq 0$ so that we obtain $\hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{p} \mathbf{\Sigma}_{\pi}^{-1}\sigma_{vu} \neq 0$. This then entails that $\sqrt{n}(\hat{\beta}_{LS} - \tilde{\beta}_{IV}) \xrightarrow{p} \infty$.

□

PROOF OF THEOREM **3.2** Let $\sigma_{vu} = 0$ and recall that

$$\mathscr{D}_j^p = \kappa_j(\hat{\beta}_{LS} - \hat{\beta}_{IV})'\hat{\Sigma}_j^{-1}(\hat{\beta}_{LS} - \hat{\beta}_{IV}) \tag{A.11}$$

where $\hat{\Sigma}_j$ and κ_j, $j = 1, 2, 3, 4$ are defined in (3.1). By noting that all $\hat{\sigma}_2, \tilde{\sigma}^2, \tilde{\sigma}_2^2 \xrightarrow{p} \sigma_u^2$ and by Lemma **3.1**-(**i**), we have $\kappa_j(\hat{\beta}_{LS} - \hat{\beta}_{IV}) \xrightarrow{d} \Psi_{\beta} \sim \mathbf{N}\left[0,\, \sigma_u^2(\tilde{\mathbf{\Sigma}}_{\pi}^{-1} - \mathbf{\Sigma}_{\pi}^{-1})\right]$ for $j = 2, 3, 4$; $\sqrt{\kappa_1}(\hat{\beta}_{LS} - \hat{\beta}_{IV}) \xrightarrow{d} \frac{1}{\sqrt{m_y}}\Psi_{\beta} \sim \frac{1}{\sqrt{m_y}}\mathbf{N}\left[0,\, \sigma_u^2(\tilde{\mathbf{\Sigma}}_{\pi}^{-1} - \mathbf{\Sigma}_{\pi}^{-1})\right]$, we have $\mathscr{D}_1^p \xrightarrow{d} \frac{1}{m_y\sigma_u^2}\Psi_{\beta}'(\tilde{\mathbf{\Sigma}}_{\pi}^{-1} - \mathbf{\Sigma}_{\pi}^{-1})^{-1}\Psi_{\beta} \sim \frac{1}{m_y}\chi^2(m_y)$, $\mathscr{D}_j^p \xrightarrow{d} \frac{1}{\sigma_u^2}\Psi_{\beta}'(\tilde{\mathbf{\Sigma}}_{\pi}^{-1} - \mathbf{\Sigma}_{\pi}^{-1})^{-1}\Psi_{\beta} \sim \chi^2(m_y)$ for all $j = 2, 3, 4$.

If $\sigma_{vu} \neq 0$, we have $\mathscr{D}_j^p \xrightarrow{d} +\infty$, for all $j = 1, 2, 3, 4$, by applying directly results in Lemma **3.1**-(**ii**).

□

PROOF OF LEMMA **3.3** (**i**) Assume first that $\sigma_{vu} = 0$ and express $\hat{\beta}_{LS} - \tilde{\beta}_{IV}$ as:

$$\hat{\beta}_{LS} - \tilde{\beta}_{IV} = (Y^{\perp'}Y^{\perp}/n)^{-1}(Y^{\perp'}\tilde{u}^{\perp}/n) - (Y^{\perp'}P_{Z^{\perp}}Y^{\perp})^{-1}Y^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp}. \tag{A.12}$$

Since $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0,\, \mathbf{\Gamma}_0]$, it is easy to see that $Y^{\perp'}Y^{\perp}/n \xrightarrow{p} \Sigma_v$ and $Y^{\perp'}\tilde{u}^{\perp}/n \xrightarrow{p} 0$, so that the first term in the right hand sight of (A.12) is $o_p(1)$. Hence, $\hat{\beta}_{LS} - \tilde{\beta}_{IV} = -(Y^{\perp'}P_{Z^{\perp}}Y^{\perp})^{-1}Y^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp} + o_p(1)$. Now, write $Y'P_{Z^{\perp}}Y = (Y^{\perp'}Z^{\perp}/\sqrt{n})(Z^{\perp'}Z^{\perp}/n)^{-1}(Z^{\perp'}Y^{\perp}/\sqrt{n})$ and $Y^{\perp'}P_{Z^{\perp}}\tilde{u}^{\perp} = (Y^{\perp'}Z^{\perp}/\sqrt{n})(Z^{\perp'}Z^{\perp}/n)^{-1}(Z^{\perp'}\tilde{u}^{\perp}/\sqrt{n})$. By observing that we now have: $Z^{\perp'}Z^{\perp}/n \xrightarrow{p} Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{1/2}$, $Z^{\perp'}Y^{\perp}/\sqrt{n} \xrightarrow{d} Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{1/2}\mathbf{\Pi}_0 + Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zv} = (Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{1/2})(\mathbf{\Pi}_0 + Q_Z^{-1}\psi_{Zv})$, and $Z^{\perp'}\tilde{u}^{\perp}/\sqrt{n} \xrightarrow{d} Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zu}$ (by Lemma

2.6), it follows that $Y'P_{Z^\perp}Y \xrightarrow{d} \mathbf{\Psi}_{Zv} = (\mathbf{\Pi}_0 + Q_Z^{-1}\psi_{Zv})'Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{1/2}(\mathbf{\Pi}_0 + Q_Z^{-1}\psi_{Zv})$ and $Y^{\perp'}P_{Z^\perp}\tilde{u}^\perp \xrightarrow{d} (\mathbf{\Pi}_0 + Q_Z^{-1}\psi_{Zv})'Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zu}$. Thus we get

$$\hat{\beta}_{LS} - \tilde{\beta}_{IV} \quad \xrightarrow{d} \quad \tilde{\mathbf{\Psi}}_\beta = -\mathbf{\Psi}_{Zv}^{-1}(\mathbf{\Pi}_0 + Q_Z^{-1}\psi_{Zv})'Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zu}. \tag{A.13}$$

Because $M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zu}$ is independent of $Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})$ and ψ_{Zu} is also independent of ψ_{Zv} under H_0^p, with a little manipulation (and using results in Lemma **2.6**-(**ii**)), we find that conditionally on $(\psi_{Zv},\, \psi_{Z\xi}) = (x_1, x_2)$,

$$\tilde{\mathbf{\Psi}}_\beta \mid_{\psi_{Zv}=x_1, \psi_{Z\xi}=Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \sim \mathbf{N}\left[0,\, \sigma_u^2\mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)\right] \tag{A.14}$$

where $\mathbf{\Psi}_{Zv}(x_1, x_2) = (\mathbf{\Pi}_0 + Q_Z^{-1}x_1)'Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(x_2)}Q_Z^{1/2}(\mathbf{\Pi}_0 + Q_Z^{-1}x_1)$. By taking the integral with respect to all possible realizations $(\psi_{Zv},\, \psi_{Z\xi}) = (x_1, x_2),$ the result follows.

(**ii**) Suppose now that $\sigma_{vu} \neq 0$. The proof is similar to those in (**i**). Firstly, note that we now have $Y^{\perp'}\tilde{u}^\perp/n \xrightarrow{p} \sigma_{vu}$ and the other limits in (**i**) do not change. So, we have

$$\hat{\beta}_{LS} - \tilde{\beta}_{IV} \quad \xrightarrow{d} \quad \Sigma_v^{-1}\sigma_{vu} + \tilde{\mathbf{\Psi}}_\beta. \tag{A.15}$$

Now, observe that $M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zu}$ is still independent of $Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})$. However, ψ_{Zu} and ψ_{Zv} are correlated. Since $(\psi_{Zu},\, \psi_{Zv},\, \psi_{Z\xi})$ is Gaussian by Assumption **2.3**-(**b**), we have $\mathbf{E}(\psi_{Zu} \mid_{\psi_{Zv}, \psi_{Z\xi}}) = \mathbf{E}(\psi_{Zu} \mid_{\psi_{Zv}}) = \psi_{Zv}\rho_{vu}$ where $\rho_{vu} = \Sigma_v^{-1}\sigma_{vu}$. As a result, we have $\mathbf{E}(\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta \mid_{\psi_{Zv}, \psi_{Z\xi}}) = \rho_{vu} - \mathbf{\Psi}_{Zv}^{-1}(\mathbf{\Pi}_0 + Q_Z^{-1}\psi_{Zv})'Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{-1/2}\psi_{Zv}\rho_{vu} = \mathbf{\Psi}_{Zv}^{-1}(\mathbf{\Pi}_0 + Q_Z^{-1}\psi_{Zv})'Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})}Q_Z^{1/2}\mathbf{\Pi}_0\rho_{vu} \equiv \mu$. So, we get

$$\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta \mid_{\psi_{Zv}=x_1, \psi_{Z\xi}=Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \quad \sim \quad \mathbf{N}\left[\mu(x_1, x_2),\, \sigma_u^2\mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)\right] \tag{A.16}$$

where $\mu(x_1, x_2) = \mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)(\mathbf{\Pi}_0 + Q_Z^{-1}x_1)'Q_Z^{1/2}M_{Q_Z^{1/2}\mathbf{\Gamma}(x_2)}Q_Z^{1/2}\mathbf{\Pi}_0\rho_{vu}$. By integrating (A.16) with respect to all possible realizations of $(\psi_{Zv},\, \psi_{Z\xi}),$ the result follows.

□

PROOF OF THEOREM **3.4** Note first that we still have $\hat{\sigma}^2, \hat{\sigma}_2^2 \xrightarrow{p} \sigma_u^2$ when $\mathbf{\Pi}_{YW} = \frac{1}{\sqrt{n}}[\mathbf{\Pi}_0,\, \mathbf{\Gamma}_0],$ whether $\sigma_{vu} = 0$ or not. Moreover, we can write $\tilde{\sigma}^2$ as:

$$\begin{aligned}
\tilde{\sigma}^2 &= (y - Y\hat{\beta}_{2SLS})'M_{\tilde{W}}(y - Y\hat{\beta}_{2SLS})/n \\
&= [y - Y\hat{\beta}_{LS} - Y(\hat{\beta}_{2SLS} - \hat{\beta}_{LS})]'M_{\tilde{W}}[y - Y\hat{\beta}_{LS} - Y(\hat{\beta}_{2SLS} - \hat{\beta}_{LS})]/n \\
&= \hat{\sigma}^2 - 2(y - Y\hat{\beta}_{LS})'Y^\perp(\hat{\beta}_{2SLS} - \hat{\beta}_{LS})/n + (\hat{\beta}_{2SLS} - \hat{\beta}_{LS})'(Y^{\perp'}Y^\perp/n)(\hat{\beta}_{2SLS} - \hat{\beta}_{LS})
\end{aligned}$$

$$= \quad \hat{\sigma}^2 + (\hat{\beta}_{2SLS} - \hat{\beta}_{LS})'(Y^{\perp}{}'Y^{\perp}/n)(\hat{\beta}_{2SLS} - \hat{\beta}_{LS}) \tag{A.17}$$

where the last equality holds because $(y - Y\hat{\beta}_{LS})'Y^{\perp} = 0$. So, it follows from Lemma **3.3** that $\tilde{\sigma}^2 \xrightarrow{d} \bar{\sigma}_u^2 = \sigma_u^2 + \tilde{\mathbf{\Psi}}_\beta' \Sigma_v \tilde{\mathbf{\Psi}}_\beta = \sigma_u^2(1 + \sigma_u^{-2}\tilde{\mathbf{\Psi}}_\beta' \Sigma_v \tilde{\mathbf{\Psi}}_\beta) \geq \sigma_u^2,$ where $\tilde{\mathbf{\Psi}}_\beta$ is defined by (A.13).

Suppose that $\sigma_{vu} = 0$. From Lemma **3.3**-(**i**), we have $\hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{d} \tilde{\mathbf{\Psi}}_\beta$ so that

$$\mathscr{D}_1^p \xrightarrow{d} \frac{1}{m_y \sigma_u^2} \tilde{\mathbf{\Psi}}_\beta' \mathbf{\Psi}_{Zv} \tilde{\mathbf{\Psi}}_\beta,\ \mathscr{D}_4^p \xrightarrow{d} \frac{1}{\sigma_u^2} \tilde{\mathbf{\Psi}}_\beta' \mathbf{\Psi}_{Zv} \tilde{\mathbf{\Psi}}_\beta, \tag{A.18}$$

$$\mathscr{D}_j^p \xrightarrow{d} \frac{\sigma_{zv}^2}{\sigma_u^2} \tilde{\mathbf{\Psi}}_\beta' \mathbf{\Psi}_{Zv} \tilde{\mathbf{\Psi}}_\beta,\ j = 2,\ 3;\ \sigma_{zv}^2 = (1 + \sigma_u^{-2}\tilde{\mathbf{\Psi}}_\beta' \Sigma_v \tilde{\mathbf{\Psi}}_\beta)^{-1} \leq 1. \tag{A.19}$$

Because $M_{Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})} Q_Z^{-1/2} \psi_{Zu}$ is independent of $Q_Z^{1/2}\mathbf{\Gamma}(\psi_{Z\xi})$ and ψ_{Zu} is also independent of ψ_{Zv} under H_0^p, From (A.14) we have $\tilde{\mathbf{\Psi}}_\beta \ |_{\psi_{Zv}=x_1, \psi_{Z\xi}=Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \sim \mathbf{N}\left[0,\ \sigma_u^2 \mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)\right]$ so that $\frac{1}{\sigma_u^2}\tilde{\mathbf{\Psi}}_\beta' \mathbf{\Psi}_{Zv} \tilde{\mathbf{\Psi}}_\beta \ |_{\psi_{Zv}=x_1, \psi_{Z\xi}=Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \sim \chi^2(m_y)$. Hence, we have $\mathscr{D}_1^p \ |_{\psi_{Zv}=x_1, \psi_{Z\xi}=Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \xrightarrow{d} \frac{1}{m_y}\chi^2(m_y)$, $\mathscr{D}_4^p \ |_{\psi_{Zv}=x_1, \psi_{Z\xi}=Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \xrightarrow{d} \chi^2(m_y)$, and $\mathscr{D}_2^p,\ \mathscr{D}_3^p \ |_{\psi_{Zv}=x_1, \psi_{Z\xi}=Q_Z^{1/2}\mathbf{\Gamma}(x_2)} \xrightarrow{d} \sigma_{zv}^2(x_1, x_2)\chi^2(m_y)$. As the conditional asymptotic distribution of $\mathscr{D}_1^p$ and $\mathscr{D}_4^p$ does not depend on (x_1, x_2), we have $\mathscr{D}_1^p \xrightarrow{d} \frac{1}{m_y}\chi^2(m_y)$, $\mathscr{D}_4^p \xrightarrow{d} \chi^2(m_y)$, unconditionally. However, the conditional asymptotic distribution of $\mathscr{D}_2^p$ and $\mathscr{D}_3^p$ depends on (x_1, x_2) through σ_{zv}^2, the unconditional distribution is obtained by integrating with respect to possible values of (x_1, x_2); where $\sigma_{zv}^2|_{x_1,\, x_2} \sim \phi_0(x_1, x_2) = [1 + \sigma_u^{-2}\mathbf{N}\left(0,\ \sigma_u^2 \mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)\right)' \Sigma_v \mathbf{N}\left(0,\ \sigma_u^2 \mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)\right)]^{-1} = [1 + \|\sigma_u^{-1}\Sigma_v^{1/2}\mathbf{N}\left(0,\ \sigma_u^2 \mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)\right)\|^2]^{-1}$. So, part (**a**) and (**b**) of Theorem **3.4** follow.

Suppose that $\sigma_{vu} \neq 0$. From Lemma **3.3**-(**ii**), we have $\hat{\beta}_{LS} - \tilde{\beta}_{IV} \xrightarrow{d} \rho_{vu} + \tilde{\mathbf{\Psi}}_\beta$ so that

$$\begin{aligned} \mathscr{D}_1^p &\xrightarrow{d} \frac{1}{m_y \sigma_u^2}(\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta)' \mathbf{\Psi}_{Zv}(\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta),\ \mathscr{D}_4^p \xrightarrow{d} \frac{1}{\sigma_u^2}(\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta)' \mathbf{\Psi}_{Zv}(\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta), \\ \mathscr{D}_j^p &\xrightarrow{d} \frac{\sigma_{zv}^2}{\sigma_u^2}(\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta)' \mathbf{\Psi}_{Zv}(\rho_{vu} + \tilde{\mathbf{\Psi}}_\beta),\ j = 2,\ 3. \end{aligned} \tag{A.20}$$

Furthermore, we can see from (A.16) that $\sigma_{zv}^2|_{x_1,\, x_2} \sim \phi(x_1, x_2) = [1 + \|\sigma_u^{-1}\Sigma_v^{1/2}\mathbf{N}\left(\mu(x_1, x_2) - \rho_{vu},\ \sigma_u^2 \mathbf{\Psi}_{Zv}^{-1}(x_1, x_2)\right)\|^2]^{-1}$ and $\mathscr{D}_1^p|_{x_1,x_2} \xrightarrow{d} \frac{1}{m_y}\chi^2(m_y\,;\|\sigma_u^{-1}\mathbf{\Psi}_{Zv}^{1/2}\mu(x_1, x_2)\|^2)$, $\mathscr{D}_4^p|_{x_1,x_2} \xrightarrow{d} \chi^2(m_y\,;\|\sigma_u^{-1}\mathbf{\Psi}_{Zv}^{1/2}\mu(x_1, x_2)\|^2)$ and $\mathscr{D}_2^p,\ \mathscr{D}_3^p|_{x_1,x_2} \xrightarrow{d} \phi(x_1, x_2)\chi^2(m_y\,;\|\sigma_u^{-1}\mathbf{\Psi}_{Zv}^{1/2}\mu(x_1, x_2)\|^2)$. Part (**c**) and (**d**) of Theorem **3.4** follow by integrating with respect $(\psi_{Zv}, \psi_{Z\xi})$.

□

B. Additional simulation results

Figure 3. Size and power at nominal level 5% when identification is strong or partial, $n = 300$

Strong and partial identification: Panels A and C

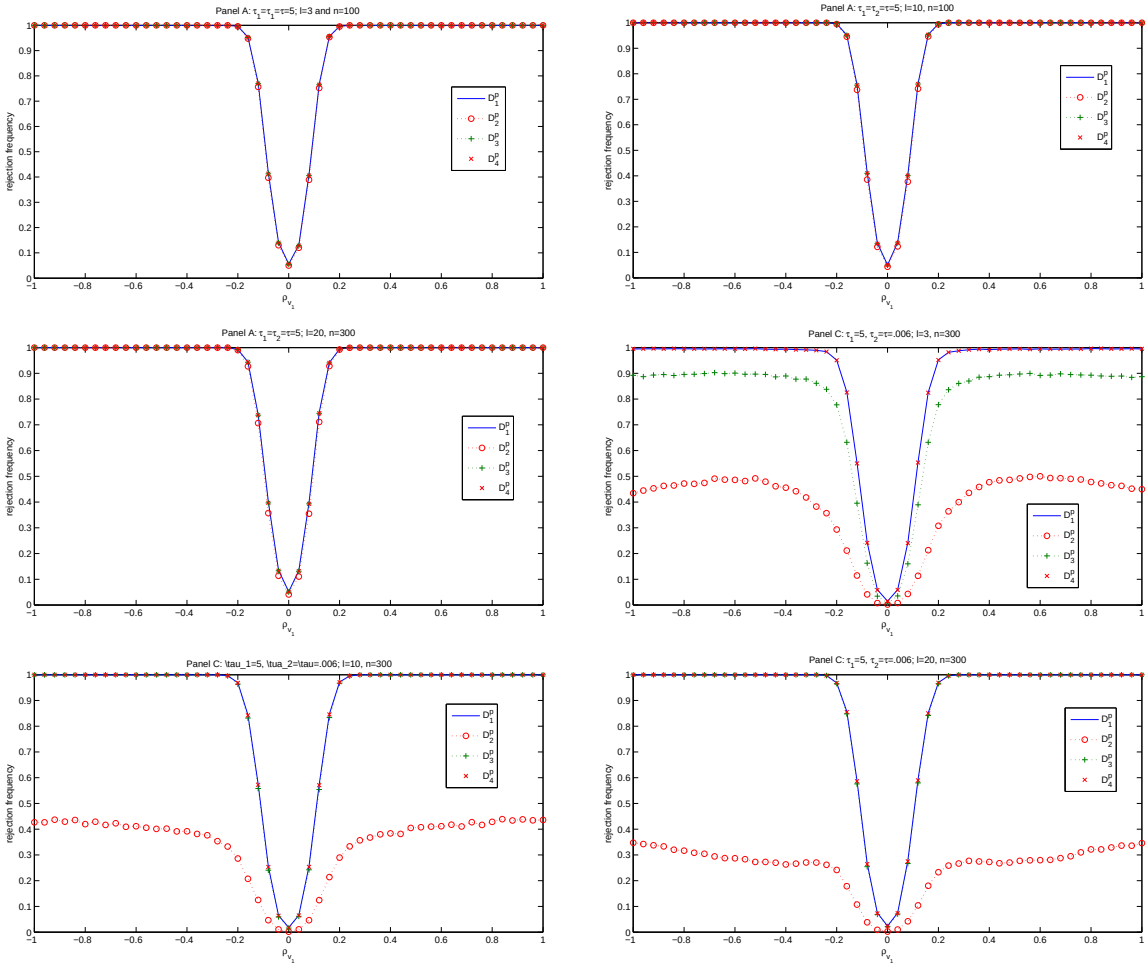


Figure 4. Size and power at nominal level 5% when identification is weak, $n = 300$

Partial identification of θ : Panels E and F

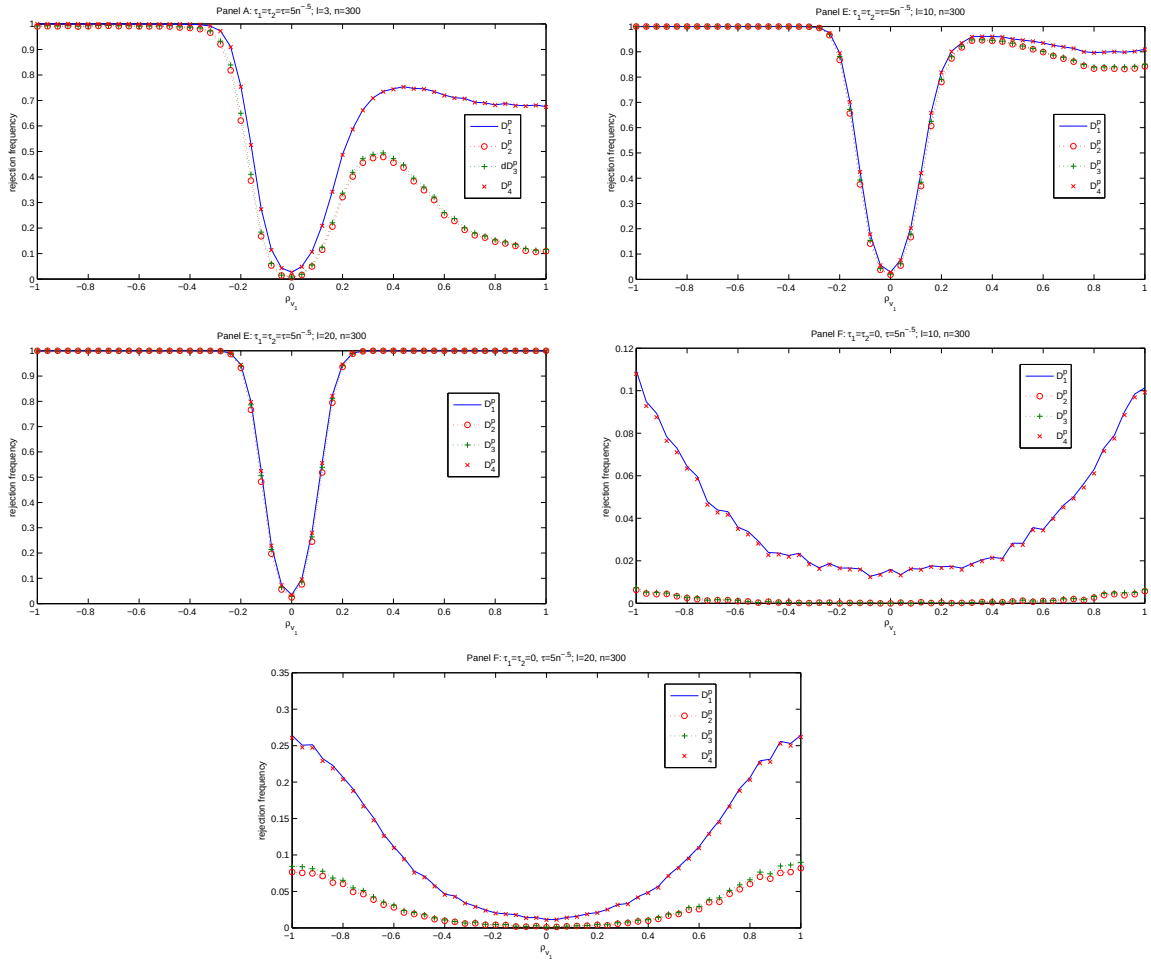
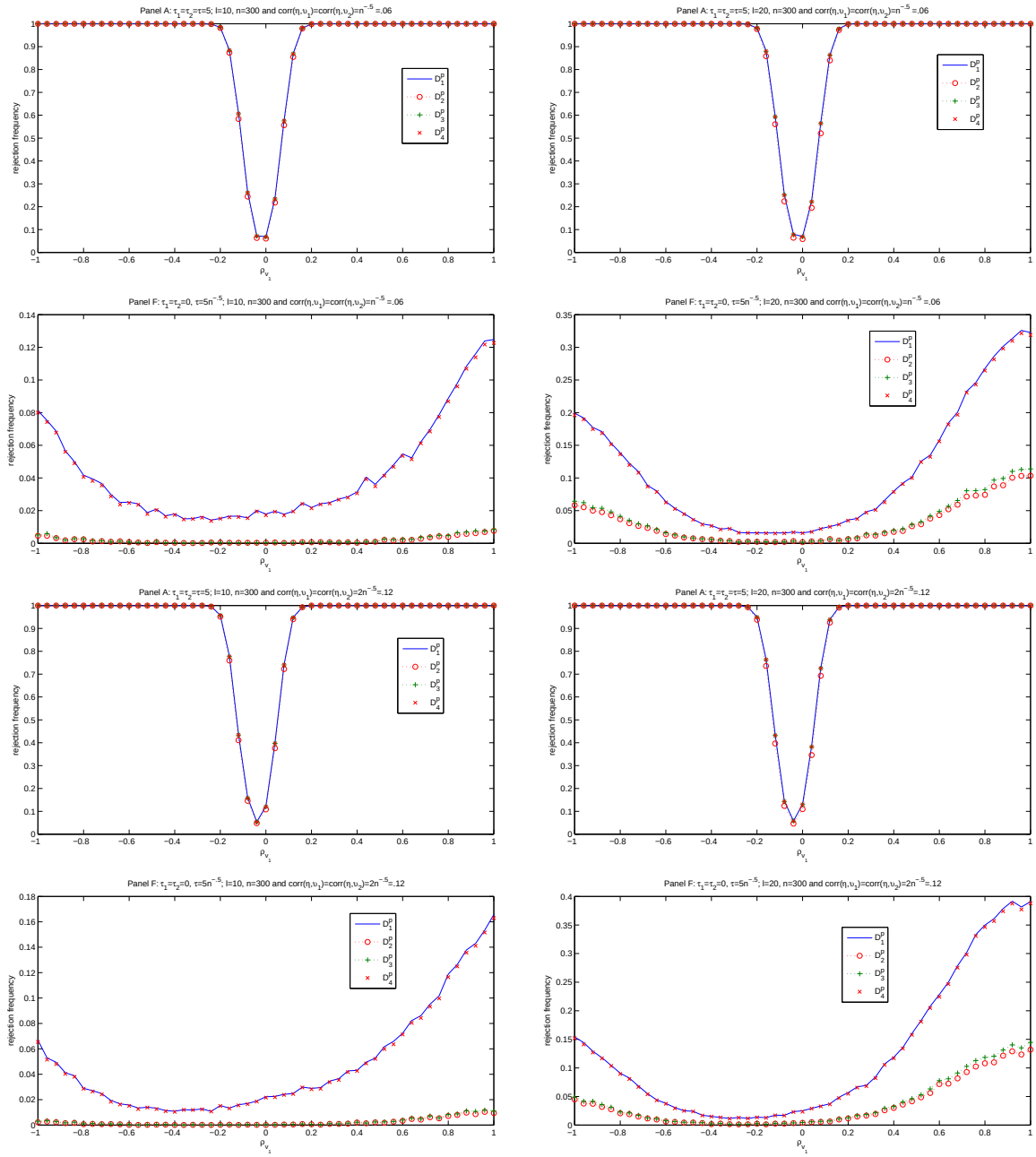


Figure 5. Size and power at nominal level 5% when the reduced form errors are correlated under the null hypothesis



References

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.

Andrews, D. W. K., Stock, 2006. Inference with weak instruments. In: R. Blundell, W. Newey, T. Pearson, eds, Advances in Economics and Econometrics, Theory and Applications, 9th Congress of the Econometric Society Vol. 3. Cambridge University Press, Cambridge, U.K., chapter 6.

Bekker, P., 1994. Alternative approximations to the distributions of instrumental variable estimators. Econometrica 62, 657–681.

Card, D., 1995. Using geographic variation in college proximity to estimate the return to schooling.. University of Toronto Press: in L. N. Christo. des, E. K. Grant, and R. Swidinsky (Eds.), Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp, .

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.

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

Doko Tchatoka, F., 2011. Subset hypotheses testing and instrument exclusion in the linear IV regression. Technical report, School of Economics and Finance, University of Tasmania Hobart, Australia.

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

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

Dufour, J.-M., Hsiao, C., 2008. Identification. In: L. E. Blume, S. N. Durlauf, eds, The New Palgrave Dictionary of Economics 2nd edn. Palgrave Macmillan, Basingstoke, Hampshire, England. forthcoming.

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., 2007. 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.

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

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

Guggenberger, P., Smith, R., 2005. Generalized empirical likelihood estimators and tests under partial, weak and strong identification. Econometric Theory 21, 667–709.

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

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

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

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.

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

Kleibergen, F., 2005. Testing parameters in GMM without assuming that they are identified. Econometrica 73, 1103–1124.

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

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

Nelson, C., Startz, R., 1990b. Some further results on the exact small properties of the instrumental variable estimator. Econometrica 58, 967–976.

Phillips, P. C. B., 1989. Partially identified econometric models. Econometric Theory 5, 181–240.

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.

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

Stock, J. H., Wright, J. H., 2000. GMM with weak identification. Econometrica 68, 1055–1096.

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.

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

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

List of Tables

List of Figures

School of Economics and Finance Discussion Papers

2012-02 Testing for Partial Exogeneity with Weak Identification, **Firmin Doko Tchatoka**

2012-01 On the Correspondence Between Data Revision and Trend-Cycle Decomposition, **Mardi Dungey, Jan PAM Jacobs** and **Jian Tian**

2011-06 Systematic and Liquidity Risk in Subprime-Mortgage Backed Securities, **Mardi Dungey, Gerald P. Dwyer** and **Thomas Flavin**

2011-05 A SVECM Model of the UK Economy and The Term Premium**, Mardi Dungey** and **M. Tugrul Vehbi**

2011-04 Do Contact Matter in the Process of Getting a Job in Cameroon? **Firmin Doko Tchatoka** and Urbain Thierry Yogo

2011-03 Subset Hypotheses Testing and Instrument Exclusion in the Linear IV Regression, **Firmin Doko Tchatoka**

2011-02 First home Buyers' Support Schemes in Australia – Results Spreadsheet, **Mardi Dungey**, **Graeme Wells** and **Sam Thompson**

2011-01 First home Buyers' Support Schemes in Australia, **Mardi Dungey**, **Graeme Wells** and **Sam Thompson**

2010-12 Financial Crises in Asia: Concordance by Asset Market or Country?, **Mardi Dungey**, Jan P.A.M. Jacobs and Lestano

2010-11 Innovation Contracts with Leakage Through Licensing, **Shane B. Evans**

2010-10 Franchise Contracts with Ex Post Limited Liability, **Shane B. Evans**

2010-09 Menus of Linear Contracts in Procurement with Type-Dependent Reservation Utility, **Shane B. Evans**

2010-08 Decomposing the Price Effects on the Cost of Living for Australian Households, **Paul Blacklow**

2010-07 Modelling the Time Between Trades in the After-Hours Electronic Equity Futures Market, **Mardi Dungey, Nagaratnam Jeyasreedharan** and **Tuo Li**

2010-06 Cojumping: Evidence from the US Treasury Bond and Futures Markets, **Mardi Dungey** and **Lyudmyla Hvozdyk**

2010-05 Assessing the Impact of Worker Compensation Premiums on Employment in Tasmania, **Paul Blacklow**

2010-04 Non-Linear Pricing with Homogeneous Customers and Limited Unbundling, **Hugh Sibly**

2010-03 Detecting Contagion with Correlation: Volatility and Timing Matter, **Mardi Dungey** and **Abdullah Yalama**

2010-02 From Trade-to-Trade in US Treasuries, **Mardi Dungey**, **Olan Henry** and **Michael McKenzie**

2010-01 Economic Assessment of the Gunns Pulp Mill 2004-2008, **Graeme Wells**

Copies of the above mentioned papers and a list of previous years' papers are available from our home site at http://www.utas.edu.au/ecofin