Impact of Instrument Endogeneity on some Test-Statistics

Firmin Doko* Jean-Marie Dufour†
Université de Montréal Université de Montréal
preliminary

24 April 2007

Abstract

When in a regression model some of the explanatory variables are correlated with the disturbance term, one needs further variables to use as instruments in order to make reliable inferences. The test-statistics often used for making these inferences are based on so-called orthogonality conditions, i.e. the instrumental variables are not correlated with the disturbances. However, it is hard to assess whether an instrumental variable is valid in practice because instrument validity is based on the questionable identifying assumption that some of them are exogenous. In this paper, we examine the impact of instrument endogeneity on the standard t-test, Anderson and Rubin (AR)-test, Kleibergen K-test and Moreira conditional likelihood ratio (CLR)-test statistics in the context of linear structural model. Our findings are : (i) an AR-type procedure is globally more robust to endogenous instruments than the others procedures and is also robust to missing (endogenous) instrument. This last result Generalizes the finding of Dufour and Taamouti (2005) to the case of endogenous instrument; (ii) a K-test is globally more robust to endogenous instruments than CLR-test but CLR-test is more robust to missing (endogenous) instrument than K-test; (iii) a t-test is less robust to invalid instruments than the others. However, it is more robust to instrument omission than K-test when the excluded instrument is endogenous; (iv) instrument invalidity is much more detrimental than instrument weakness on the inference procedures based on AR, K and CLR statistics, i.e. tests based on strong but invalid instruments have higher size distortions than those based on valid but weak instruments; (v) AR and K statistics are not pivotal even asymptotically when some of the instruments are endogenous.

Key words : structural model; instrumental variables; weak instruments; invalid instruments ; missing instrument; AR test; K test; CLR test.

*Université de Montréal, Département de Sciences Économiques, C.P. 6128, succursale Centre-ville, Montréal, QC H3C 3J7. Téléphone: (514)343-6111 poste 5283; Fax: (514)343-7221. E-mail adress : f.doko.tchatoka@umontreal.ca.

†Canada Research Chair in Econometrics, CIRANO, CIREQ, Université de Montréal; Mailing adress : Département de Sciences Économiques, C.P. 6128, succursale Centre-ville, Montréal, QC H3C 3J7. Téléphone: (514)343-2400; Fax: (514)343-5831. E-mail adress : jean.marie.dufour@umontreal.ca.

1 Introduction

The last decade shows growing interest on so-called weak instruments problems in the econometric literature. Several studies note that weak instruments cause serious statistical difficulties from the viewpoint of estimation, confidence set construction and testing. When the quality of the instruments is poor, the limiting distributions of standard statistics like the t-statistic, Wald, likelihood ratio and Lagrange multiplier, based on the (IV) estimator of a structural model, depend on nuisance parameters; see e.g. Dufour (1997), Bekker (1994), Phillips (1989), Staiger and Stock (1997), and Wang and Zivot (1998). The Anderson and Rubin (1949) statistic, whose distribution is pivotal even in finite sample, is often used to alleviate these problems. Although the AR-test statistic has some defiencies, due to its possible low power when too many instruments are added, several solutions are proposed to palliate these deficiencies : approximately optimal instruments [Dufour and Taamouti (2001b)]; sample-splitting methods [Dufour and Jasiak (2001)], systematic search methods for identifying relevant instruments and excluding unimportant instruments [Hall, Rudebusch and Wilcox (1996); Hall and Peixe (2000); Dufour and Taamouti (2001a); Donald and Newey (2001)]; Pseudo-pivotal LM-type and LR-type statistics [Wang and Zivot (1998); Zivot , Stwartz and Nelson (1998); Kleibergen's K statistic [Kleibergen (2002, 2004, 2005)] and Moreira conditional LR statistic [Moreira (2002, 2003a, 2003b)]. However, all these procedures are built on the prerequisite of having valid instruments, which raises an interesting question : what happens to these procedures when some of the instruments are weak and endogenous? In others words, how robust are these inference procedures to invalid instruments? The problem of invalid instruments on the inference procedures is important because it is hard in practice to assess whether an instrumental variable is valid, i.e is uncorrelated with the disturbance term. Instrument validity or orthogonality tests are built on the questionable identify assumption of having already available a number of undisputed valid instruments, at least as great as the number of coefficients to be estimated, whereas the validity of those initial instruments is untestable. In econometric literature, exogeneity tests do not deal with these problems. Two tendencies Characterized this literature in the past. The first is tests for the hypothesis of independence between a full vector of stochastic explanatory variables and a disturbance term proposed by several authors, see Durbin (1954), Wu (1973, 1974), Revankar and Hartley (1973), Farebrother (1976), Hausman (1978), Revankar (1978), Kariya and Hodoshima (1980), Richard (1980), Holly and Sargan (1982), Dufour (1987). The second is tests for the hypothesis of independence between a subset of stochastic explanatory variables and a disturbance in a structural equation. These tests are proposed by a number of authors : Hwang (1980) and Smith (1984) studied likelihood ratio tests, Hausman and Taylor (1981a), Spencer and Berk (1981) and Wu (1983b) extended the tests previously studied by Wu (1973) and Hausman (1978), Engle (1982) derived the LM tests while Dufour (1987) developed a generalized Wald test. More recently, Hausman and Hahn (2003) deal with both the problem of endogeneity and weakness of instruments, but from the viewpoint of estimation. Analyzing the effect of instrument invalidity on the limiting and empirical distribution of IV estimators, Kiviet and Niemczyk (working paper) conclude that : (a) for the accuracy of asymptotic approximations, instrument weakness

is much more detrimental than instrument invalidity; (b) the realizations of IV estimator obtained from strong but possibly invalid instruments seem usually much closer to the true parameter values than those obtained from valid but weak instruments. Once again, the important question is what happens to the inference procedures (in finite sample) based on AR-type test statistics in this case? R. Blundell and Horowitz (working paper) propose an inference procedure based on a nonparametric function that is identified by a conditional moment restriction involving instrumental variables, but the problem with this procedure is that it is only valid in large sample. Thus, we see that little is known about the inference procedures (in finite sample) when instruments are invalid or both weak and invalid.
In this paper, we focus on structural models and analyze the effects of instrument endogeneity on four test statistics : the standard t-test associated to the 2SLS estimator of structural coefficients (T_{2SLS}) and the tests meant to be robust to weak instruments : Anderson and Rubin AR-test, Kleibergen's K-test and Moreira's conditional likelihood ratio (CLR)-test. First, we show by Monte Carlo studies that the AR-type, K-type and CLR-type procedures are more robust to invalid instruments than the standard t-test procedure. When the number of instruments is large (which is often the case in empirical work), the AR-type procedure is more robust to invalid instruments than K-type and CLR-type procedure whereas K-type procedure is more robust than CLR-type procedure. Our results also show that instrument invalidity is much more detrimental than instrument weakness on the inference procedures based on AR, K and CLR statistics, i.e. tests based on strong but invalid instruments have higher size distortions than those based on valid but weak instruments. Finally, we show that AR-type procedure is robust to missing (endogenous) instrument. This last result generalizes the finding of Dufour and Taamouti (2005). we also show that the t-test statistic is the most affected by instrument omission and is more robust to instrument omission that K-test when the excluded instrument is endogenous. Moreover, the CLR-test is globally more robust to (invalid) instrument omission than the K-test. Second, we show that when some instruments are invalid, then, like the t-statistic and CLR-statistic, the AR-statistic and the K-statistic are not pivotal even in large sample. Irrespective of the quality of the instruments, the AR-statistic has a non central $\frac{1}{k_2}\chi^2$ limiting distribution with k_2[1] degree of freedom and its non centrality depends only on instrument endogeneity parameter τ_0 and the standard error σ_{11} of the structural disturbance . When the instruments are both weak and invalid, the conditional asymptotic null distribution of the K-statistic is a non central χ^2 with G degree of freedom [2] and its non centrality parameter depends not only on τ_0 and σ_{11} , but also on the reduced form equation parameters : Π_0 and the limiting distribution of the linear transformation of the reduced form error. Moreover, when instruments are both strong and invalid, K-statistic has a non central χ^2 limiting distribution with G degree of freedom and its non centrality parameter depends on τ_0, σ_{11} , Π_0, but not on the reduced form error parameters.
The paper is organized as follows. Section 2 formulates the model considered. Section 3 describes briefly the four test statistics. Section 4 presents the Monte Carlo study where we examine the effect of endogeneity on the sizes of the tests and the effect of missing

[1] k_2 is the number of instruments excluded from the structural equation
[2] G is the number of structural parameters

instruments. In section 5, we derive the limiting null distribution of the statistics . Finally the sixth section concludes.

2 Framework

Let us consider the following common simultaneous equation framework, which has been the basis of many works on inference in model with possibly weak instruments [see Dufour(2003), Dufour and Taamouti (2005) and Stock et al.(2002)]:

$$y = Y\beta + X_1\gamma + u \tag{2.1}$$

$$Y = X_1\Pi_1 + X_2\Pi_2 + V \tag{2.2}$$

where y is a $T \times 1$ vector of observations on the dependent variable ; Y is a $T \times G$ matrix of observations on the explanatory endogenous variables ($G \geq 1$); X_1 is a $T \times k_1$ matrix of observations on the included exogenous variables; X_2 is a $T \times k_2$ matrix of observations on the excluded variables(instruments); $u = \begin{bmatrix} u_1 & \dots & u_T \end{bmatrix}'$ and $V = \begin{bmatrix} V_1 & \dots & V_T \end{bmatrix}'$ are respectively a $T \times 1$ vector and a $T \times G$ matrix of disturbances; β and γ are $G \times 1$ and $k_1 \times 1$ vectors of unknown coefficients and Π_1, Π_2 are $k_1 \times G$ and $k_2 \times G$ matrices of unknown coefficients.

We assume that some instruments in X_2 are invalid. Indeed, Let X_2 be partitioned as $X_2 = [\bar{X}_2, \tilde{X}_2]$, the invalidity condition is such that

$$cov\left(\tilde{X}_2, u\right) \neq 0 \tag{2.3}$$

.

To allow for a finite-sample distributional theory, we assume that:

- $$X = [X_1, X_2] \text{ is full-column rank}\, T \times k \,\text{matrix} \tag{2.4}$$
  where $k = k_1 + k_2$;

- $$u \sim N\left(0, \sigma_u^2 I_T\right). \tag{2.5}$$

Condition (2.3) means that the instruments $\tilde{X}_2$ are endogenous. This assumption is relevant because in practice, it is hard to assess whether an instrumental variable is valid. Instrument validity is only viable under just identification or overidentification by truly valid instruments, i.e. it is built on the prerequisite of having already available a number of undisputed valid instruments, at least as great as the number of coefficients to be estimated, whereas the validity of the initial instruments is untestable.

We consider the problem of building tests and confidence sets on β. In particular, we consider the problem $H_0 : \beta = \beta_0$ and analysis the impact of condition (2.3) on four test statistics : the standard t-statistic based on the 2SLS estimator, the AR-test statistic (Anderson-Rubin, 1949), the K-statistic (Kleibergen, 2002) and the conditional Likelihood Ratio test statistic (Moreira, 2003).

3 The Test Statistics

The 2SLS estimator of the parameter of interest β in the system (2.1)-(2.2) is given, according to Frisch-Waugh theorem, by

$$\hat{\beta}_{2SLS} = (Y' M_{X_1} P_X Y)^{-1} Y' M_{X_1} P_X y \tag{3.1}$$

where $P_X = X(X'X)^{-1}X'$ and $M_{X_1} = I - X_1(X_1'X_1)^{-1}X_1'$. The t-statistic is given under H_0 and assumption (2.5) as

$$T_{2SLS_j} = \hat{\Omega}_{jj}^{-1/2}[\hat{\beta}_{2SLS_j} - \beta_{0j}] \tag{3.2}$$

where $\hat{\Omega}_{jj}$, $1 \leq j \leq G$, is the jth element of the diagonal of the matrix

$\frac{1}{T-k}(Y' M_{X_1} P_X Y)^{-1} y' M_{(\hat{Y}, X_1)} y$, $\hat{\beta}_{2SLS} = (\hat{\beta}_{2SLSj})_{1 \leq j \leq G}$, $\beta_0 = (\beta_{0j})_{1 \leq j \leq G}$, $\hat{Y} = P_X Y$,

$M_{(\hat{Y}, X_1)} = I - P_{(\hat{Y}, X_1)}$ and $P_{(\hat{Y}, X_1)}$ is the projection matrix on the space spanned by the columns of $[\hat{Y}, X_1]$.

The Anderson and Rubin (1949) test for H_0 in equation (2.1) involves considering the transformed equation

$$y - Y\beta_0 = X_1 \Psi_1 + X_2 \Psi_2 + \xi \tag{3.3}$$

where $\Psi_1 = \gamma + \Pi_1 (\beta - \beta_0)$, $\Psi_2 = \Pi_2 (\beta - \beta_0)$ and $\xi = u + V(\beta - \beta_0)$. H_0 can then be assessed by testing $H_0' : \Psi_2 = 0$. The AR-statistic for H_0' is given by

$$AR(\beta_0) = \frac{(T-k)}{k_2} \frac{(y - Y\beta_0)'(M_{X_1} - M_X)(y - Y\beta_0)}{(y - Y\beta_0)' M_X (y - Y\beta_0)} \tag{3.4}$$

where $M_B = I - P(B)$, $P_B = B(B'B)^{-1}B'$ is the projection matrix on the space spanned by the columns of B. If condition (2.3) fails to hold, under H_0, equation (3.3) satisfies all the conditions of the classical linear model and $AR(\beta_0) \sim F(k_2, T-k)$. Unfortunately, this will not be the case if condition (2.3) holds, because $\tilde{X}_2$ is correlated with ξ.

Kleibergen (2002) proposed a modification of Anderson-Rubin statistic. His statistic[3] for testing H_0 is

$$K(\beta_0) = \frac{(y - Y\beta_0)' P_{\tilde{Y}(\beta_0)} (y - Y\beta_0)}{\frac{1}{T-k}(y - Y\beta_0)' M_X (y - Y\beta_0)} \tag{3.5}$$

where
$\tilde{Y}(\beta_0) = X\tilde{\Pi}(\beta_0)$,
$\tilde{\Pi}(\beta_0) = (X'X)^{-1}X'[Y - (y - Y\beta_0)\frac{S_{uV}(\beta_0)}{S_{uu}(\beta_0)}]$,
$S_{uu}(\beta_0) = \frac{1}{T-k}(y - Y\beta_0)' M_X (y - Y\beta_0)$,
$S_{uV}(\beta_0) = \frac{1}{T-k}(y - Y\beta_0)' M_X Y$.

[3]Kleibergen derived the K-statistic by supposing without loss of generality that $\gamma = 0$ and $\Pi_1 = 0$ in (2.1)-(2.2). So, there is no exogenous variable included in the model.

Unlike the AR-statistic which projects $(y - Y\beta_0)$ on the k columns of X, the K-statistic projects $(y - Y\beta_0)$ on the G columns of $X\tilde{\Pi}(\beta_0)$. If condition (2.3) fails to hold, i.e if the instruments X are exogenous, $\tilde{\Pi}(\beta_0)$ is both a consistent estimator of Π and asymptotically independent of $(y - Y\beta_0)'X$ under H_0, and $K(\beta_0)$ converges to a $\chi^2(G)$. However, if condition (2.3) holds, $\tilde{\Pi}(\beta_0)$ is not asymptotically independent of $(y - Y\beta_0)'X$ and the asymptotic distribution of the K-statistic will fail to be a $\chi^2(G)$.

Moreira (2003) used a similar test based on conditioning for testing $H_0 = \beta_0$. His statistic is given by :

$$CLR(\beta_0) = \hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\lambda}^{min} \tag{3.6}$$

where $\hat{\lambda}^{min}$ is the smallest eigenvalue of $(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}})$,
$\hat{\mathcal{S}} = (X'X)^{-1/2}X'\underline{Y}b_0(b_0'\hat{\Delta}b_0)^{-1/2}$,
$\hat{\mathcal{J}} = (X'X)^{-1/2}X'\underline{Y}\hat{\Delta}^{-1}A_0(A_0'\hat{\Delta}^{-1}A_0)^{-1/2}$,
$\hat{\Delta} = \underline{Y}'M_X\underline{Y}/(T-k)$ is the estimation of covariance matrix of the reduced-form errors.
$\underline{Y} = [y, Y]$, $M_X = I - P(X)$, $P_X = X(X'X)^{-1}X'$, $b_0 = [1, -\beta_0]'$ and $A_0 = [\beta_0, I_G]'$.
When $G = 1$, the statistic CLR(β_0) is simplified to

$$CLR(\beta_0) = \frac{1}{2}\left[\hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\mathcal{J}}'\hat{\mathcal{J}} + \sqrt{(\hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\mathcal{J}}'\hat{\mathcal{J}})^2 - 4[\hat{\mathcal{S}}'\hat{\mathcal{S}}\hat{\mathcal{J}}'\hat{\mathcal{J}} - (\hat{\mathcal{S}}'\hat{\mathcal{J}})^2]}\right] \tag{3.7}$$

where
$\hat{\mathcal{S}} = \frac{(X'X)^{-1/2}X'\underline{Y}b_0}{\sqrt{b_0'\hat{\Delta}b_0}}$, $\hat{\mathcal{J}} = \frac{(X'X)^{-1/2}X'\underline{Y}\hat{\Delta}^{-1}a_0}{\sqrt{a_0'\hat{\Delta}^{-1}a_0}}$,
$a_0 = [\beta_0, 1]'$.
Even if condition (2.3) fails to hold, the (weak-instrument) asymptotic distribution of CLR(β_0) under the null , conditional on $\hat{\mathcal{J}} = \zeta$, is nonstandard and depends on β_0 and ζ. With condition (2.3), the (weak-instrument) asymptotic distribution of CLR(β_0) under $H_0 = \beta_0$, conditional on $\hat{\mathcal{J}} = \zeta$ must depend not only on β_0 and ζ, but also on τ[4]. So, instrument endogeneity would have some effects on the size and power of the CLR-statistic. In the following section, we explore how robust are these four test statistics [T_{2SLS}, AR, K and CLR] to instrument endogeneity.

4 Impact of instrument endogeneity on the statistics: Monte carlo experiments

In this section, we present a study of the finite sample behavior of the standard t-test and tests meant to be robust to weak instruments[5] when some of the instruments are invalid. First, we study the impact of instrument endogeneity on the size of these tests and second, we analyze their robustness to missing instruments.

[4] τ is the vector of instrument endogeneity parameters, i.e, $\tau = cov(X, u)$.

[5] the tests considered are the Anderson-Rubin AR-test, the Kleibergen's K-test and the Moreira's Conditional Likelihood Ratio CLR-test

4.1 Impact of invalid instruments on the size

This subsection presents the effects of instrument endogeneity on the size of the four tests. We consider the following data generating process :

$$y = Y\beta + u, \quad Y = X_2\Pi_2 + w \tag{4.1}$$

$$(u_t, w_t)' \overset{i.i.d}{\sim} N(0, \Sigma), \Sigma = \begin{pmatrix} 1 & 0.95 \\ 0.95 & 1 \end{pmatrix} \tag{4.2}$$

$$X_2 = [\bar{X}_2, \tilde{X}_2, \tilde{X}_3] \tag{4.3}$$

$$\begin{aligned} \tilde{X}_2 = M(\bar{X}_2)X_{22}, \quad \tilde{X}_3 = M(\bar{X}_2)X_{23} \\ X_{22} = \xi_1 X_0 + \tau_1 u, \, X_{23} = \xi_2 \bar{X}_0 + \tau_2 u \end{aligned} \tag{4.4}$$

where $\xi_1 = \sqrt{1-\tau_1^2}$, $\xi_2 = \sqrt{1-\tau_2^2}$, $M_B = I - P(B)$, $P_B = B(B'B)^{-1}B'$ is the projection matrix on the space spanned by the columns of B and X_0, $\bar{X}_0 \overset{i.i.d}{\sim} N(0, I_T)$. X_2 is the $T \times k_2$ matrix of instruments excluded in the structural equation, $\bar{X}_2$ is $T \times \bar{k}$ matrix i.i.d of $N(0,1)$ variables and $\tilde{X}_2$, $\tilde{X}_3$, X_{22} and X_{23} are $T \times 1$ vectors. Clearly, when $\bar{k} = 0$, they are no exogenous instruments in X_2 and $\bar{X}_2$ vanishes. Without loss of generality, we assume that $\tilde{X}_2$ and $\tilde{X}_3$ are orthogonal to $\bar{X}_2$. From equation (4.4), $\tilde{X}_2$ and $\tilde{X}_3$ are correlated with u unless τ_1 and τ_2 are all equal to zero. Which means that instrument $\tilde{X}_2$ or $\tilde{X}_3$ is endogenous in the model while τ_1 or τ_2 is different from zero. When X_2 contains only one instrument, the Anderson-Rubin test statistic, Kleibergen's K-statistic and Moreira's conditional likelihood ratio test statistic are all identical. In this case, the K-statistic and the CLR-statistic are both pivotal if the instrument is exogenous (because they are reduced to AR-statistic which is pivotal in finite sample). Moreover, when $k_2 > 1$, the three statistics are different and if the instruments are valid, then, unlike the AR-test statistic, the K-test and CLR-test statistics are not pivotal in finite sample. The K-statistic is asymptotically pivotal but the CLR-statistic is not pivotal even in large sample. Moreira (2003) suggested computing the critical values for the conditional LR statistic by Monte Carlo simulation.

When $\tau_1 = \tau_2 = 0$, instruments $\tilde{X}_2$ and $\tilde{X}_3$ are valid, otherwise, at least one instrument is invalid. The matrix Π_2 is such that $\Pi_2 = \eta\Pi$, where η takes the value 0.001 (design of weak identification) or 0.5 (design of strong identification), and Π is a $k_2 \times 1$ vector of one. The correlation coefficient ρ between u and w is set equal to 0.95, i.e, the variable Y is endogenous and the instrumental variables X_2 are necessary. The number of instruments k_2 varies from 1 to 60 in Table 1 and from 5 to 80 in Table 2. We consider three null hypotheses $H_0 = \beta_0$ for β_0 =10, 5 and 0. The sample size is $T = 100$. The number of replications is $N = 1000$ and the conditional LR-critical values are computed using the same number of replications.

The results are presented in Table 1 and Table 2 below. In Table 1, we have at least one invalid instrument whereas in Table 2, we have at most one invalid instrument[6]. In the

[6]In table 1, we set $\tau_1 \in \{0, .05, .1, .15, .2, .3\}$ and $\tau_2 = 0$ whereas in table 2, we kept $\tau_1 \in \{0, .05, .1, .15, .2, .3\}$ but $\tau_2 = .3$

first column of the tables, we report the values of η whereas in the two second columns, we report the values of k_2 and β_0. Finally in the others columns, we report for each value of τ_1 and τ_2, the rejected frequencies at nominal level of 5 % for the four statistics T_{2SLS}, AR, K and CLR.

The main observations from these results are that the sizes of all these tests can be seriously affected when instruments are invalid and the empirical rejection frequencies of the tests (as hight as 100 %) exceeds 5 % by very wide margins. The more invalid the instruments, the larger the distortions.Thus, instrument endogeneity has some important drawbacks on the inference procedures based on these four test statistics. When k_2=1, the three statistics AR, K and CLR are identical and their rejection frequencies are the same. When instruments are valid[7], only the AR-test statistic has good level irrespective of the weakness of instruments [robustness of AR-type test statistic to weak instruments : Anderson and Rubin (1949)]. The K-test and CLR-test statistics show size distortions even if the instruments are valid ($\tau_1 = 0$ in Table 1), with empirical rejection frequencies as high as 23 % for the K-statistic and 22 % for CLR-statistic. When instruments are weak (i.e $\eta = .001$), the larger are the instruments, the larger the distortions for the two statistics K and CLR whereas the AR-statistic gets smaller distortions when the instruments become larger. The same observation remains valid for the AR-statistic when instruments are strong (i.e $\eta = .5$) but vanishes for the K-statistic and CLR-statistic. When instruments are strong and large, the statistic T_{2SLS} shows small distortions (below 5 %) even with invalid instruments (at most 4 % if k_2=80). Thus, the behavior of the statistic T_{2SLS} is similar to the findings of F. Kiviet and J. Niemczyk (working paper) [8]. When instruments are strong, the size distortions of the CLR-statistic due to endogeneity grows with the value of β_0, i.e the greater is β_0, the larger the distortions for the CLR-test. As far as the comparison of performance of the tests is concerned (even if the results do not show an uniform dominance when instruments are strong), our major findings are : (i) the AR-test, K-test and CLR-test statistics are more robust to invalid instruments than the t-test statistic; (ii) when the number of instruments is large (which is often the case in empirical work), the AR-type procedure is more robust to invalid instruments than the K-type and CLR-type procedures whereas the K-type procedure is more robust than the CLR-type procedure; (iii) instrument invalidity is much more detrimental than instrument weakness on the inference procedures based on AR, K and CLR type procedures, i.e. tests based on strong but invalid instruments have higher size distortions than those based on valid but weak instruments.

[7] i.e when $\tau_1 = 0$ in table 1

[8] Jan F. Kiviet and Jerzy Niemczyk find that the realizations of IV estimors obtained from strong but possibly invalid instruments seem usualy much closer to the true parameter values than those obtained from valid but weak instruments.

Table 1: Percent Rejected at nominal level of 5 %

			$\tau_1 = 0$				$\tau_1 = 0.05$				$\tau_1 = 0.1$			
	k_2	β_0	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\eta = 0.001$	1	0	44.4	5.4	5.4	5.4	52.5	8.0	8.0	8.0	61.9	16.5	16.5	16.5
	5	0	97.5	5.5	7.3	7.3	97.1	6.1	7.1	6.7	97.6	7.8	8.2	8.8
	40	0	100.0	4.4	11.1	10.8	100.0	4.6	13.6	13.5	100.0	5.5	15.2	15.0
	60	0	100.0	6.2	23.1	22.5	100.0	4.8	21.7	21.4	100.0	5.6	22.9	21.9
	1	5	46.4	5.6	5.6	5.6	50.3	8.6	8.6	8.6	60.9	16.1	16.1	16.1
	5	5	97.2	5.1	5.9	6.6	97.9	6.0	6.5	7.2	98.1	10.4	8.3	12.0
	40	5	100.0	5.6	14.1	13.2	100.0	5.3	13.1	14.1	100.0	5.4	13.0	11.0
	60	5	100.0	5.1	20.2	18.5	100.0	4.3	20.7	18.7	100.0	5.5	23.6	19.5
	1	10	46.0	4.6	4.6	4.6	50.5	7.3	7.3	7.3	62.6	16.1	16.1	16.1
	5	10	96.3	4.8	6.6	6.0	97.1	7.5	6.5	8.2	97.8	8.2	8.2	8.9
	40	10	100.0	5.3	14.6	14.0	100.0	4.7	14.0	11.4	100.0	5.7	11.2	14.3
	60	10	100.0	4.5	19.9	17.5	100.0	4.7	22.7	17.7	100.0	5.9	22.3	18.8
			$\tau_1 = 0.15$				$\tau_1 = 0.2$				$\tau_1 = 0.3$			
	k_2	β_0	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\eta = 0.001$	1	0	78.4	32.7	32.7	32.7	88.9	51.0	51.0	51.0	98.9	87.2	87.2	87.2
	5	0	98.5	16.0	9.8	10.1	99.3	28.5	14.6	16.8	99.9	64.4	28.6	36.4
	40	0	100.0	6.6	13.8	14.0	100.0	9.8	16.9	16.2	100.0	19.8	18.8	19.0
	60	0	100.0	5.7	25.2	25.1	100.0	9.9	26.3	25.8	100.0	11.6	26.9	26.8
	1	5	79.2	32.8	32.8	32.8	87.5	53.0	53.0	53.0	98.1	88.2	88.2	88.2
	5	5	98.9	16.7	13.7	18.5	98.7	28.5	17.1	31.4	99.9	65.7	32.4	68.1
	40	5	100.0	8.8	16.7	17.0	100.0	9.1	18.3	20.4	100.0	18.3	18.7	32.5
	60	5	100.0	5.8	24.1	21.6	100.0	7.5	25.6	23.9	100.0	12.6	26.0	32.1
	1	10	77.9	33.5	34.4	35.0	89.4	52.5	52.5	52.5	98.1	86.4	86.4	86.4
	5	10	99.0	16.3	11.6	18.4	99.4	30.0	18.0	32.5	99.9	62.3	29.1	64.9
	40	10	100.0	7.4	14.8	18.6	100.0	9.4	16.8	21.1	100.0	16.5	21.2	31.8
	60	10	100.0	7.5	23.0	22.2	100.0	8.9	24.3	25.2	100.0	13.0	29.3	33.8

Table 1 (continued): Percent Rejected at nominal level of 5 %

	k_2	β_0	$\tau_1 = 0$				$\tau_1 = 0.05$				$\tau_1 = 0.1$			
			T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\eta = .5$	1	0	10.5	5.7	5.7	5.7	19.1	7.7	7.7	7.7	38.1	16.1	16.1	16.1
	5	0	12.3	3.5	4.9	5.2	15.8	5.4	4.5	4.4	26.1	10.0	8.4	8.2
	40	0	19.6	4.2	5.4	5.2	21.5	5.0	5.7	6.1	23.0	7.0	6.8	6.9
	60	0	14.9	3.9	5.8	5.6	13.5	5.6	5.5	5.4	16.3	5.6	6.5	6.6
	1	5	8.0	5.1	5.1	5.1	18.1	6.8	6.8	6.8	37.7	16.2	16.2	16.2
	5	5	14.2	5.6	5.7	5.7	18.4	6.8	5.3	6.9	24.4	9.7	8.2	8.8
	40	5	18.9	5.0	4.7	5.6	19.2	5.4	5.0	5.6	22.2	5.9	6.1	6.3
	60	5	12.4	4.0	5.8	6.1	13.8	5.7	6.5	7.6	14.4	7.1	5.4	6.3
	1	10	10.0	5.1	5.1	5.1	20.9	7.0	7.0	7.0	37.5	18.3	18.3	18.3
	5	10	12.9	5.3	5.8	6.6	16.6	6.8	5.1	7.7	24.1	9.4	7.5	11.4
	40	10	19.1	4.7	6.1	11.0	20.1	6.1	4.9	14.6	20.5	7.7	5.8	14.9
	60	10	14.9	5.3	5.7	18.4	14.3	4.6	6.9	17.4	13.9	5.8	6.3	18.3

	k_2	β_0	$\tau_1 = 0.15$				$\tau_1 = 0.2$				$\tau_1 = 0.3$			
			T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\eta = .5$	1	0	59.2	31.3	31.3	31.3	76.4	54.3	54.3	54.3	95.7	87.9	87.9	87.9
	5	0	32.7	14.5	10.3	10.7	45.4	28.2	13.2	13.1	69.9	62.6	26.5	26.9
	40	0	21.6	6.4	5.9	5.7	25.3	11.7	7.2	7.3	33.0	18.4	7.8	8.1
	60	0	14.8	6.4	3.9	4.0	19.3	7.4	7.3	7.6	20.3	12.1	5.5	5.7
	1	5	55.8	31.1	31.1	31.1	75.0	50.7	50.7	50.7	95.8	88.1	88.1	88.1
	5	5	34.3	16.3	11.6	13.7	44.6	27.7	15.5	22.4	71.8	63.1	29.2	49.3
	40	5	22.3	5.7	5.3	6.3	27.4	9.8	6.2	7.6	32.7	20.1	7.8	10.7
	60	5	15.0	6.2	4.9	5.9	16.9	7.4	6.6	8.6	19.7	13.6	7.7	10.7
	1	10	59.1	34.2	34.2	34.2	72.8	48.3	48.3	48.3	95.4	87.2	87.2	87.2
	5	10	35.1	14.4	10.8	15.9	45.9	27.6	14.9	30.3	69.9	64.7	29.8	67.6
	40	10	23.6	6.2	5.9	15.5	26.9	9.7	6.6	20.7	33.5	17.5	9.1	33.2
	60	10	14.7	6.0	6.3	21.6	17.7	9.1	6.2	26.2	18.9	11.8	7.8	32.9

Table 2 : Percent Rejected at nominal level of 5 %

	k_2	β_0	$\tau_1 = 0, \tau_2 = .3$ T_{2SLS}	AR	K	CLR	$\tau_1 = .05, \tau_2 = .3$ T_{2SLS}	AR	K	CLR	$\tau_1 = .1, \tau_2 = .3$ T_{2SLS}	AR	K	CLR
$\eta = .001$	5	0	99.9	55.9	28.1	34.9	100.0	60.7	28.4	34.5	100.0	66.5	33.3	39.8
	40	0	100.0	16.2	21.4	22.2	100.0	19.2	21.8	22.2	100.0	20.5	20.4	21.3
	60	0	100.0	13.1	28.0	28.4	100.0	13.1	26.9	27.2	100.0	12.9	26.1	27.4
	80	0	100.0	7.2	40.8	41.5	100.0	8.9	43.3	43.5	100.0	9.4	42.6	43.0
	5	5	100.0	57.4	28.4	61.0	100.0	58.9	26.9	61.3	100.0	63.7	31.4	66.0
	40	5	100.0	17.1	17.9	29.7	100.0	16.4	19.3	30.7	100.0	19.8	21.0	33.5
	60	5	100.0	10.2	27.1	30.9	100.0	11.3	24.5	37.4	100.0	12.9	28.6	33.3
	80	5	100.0	8.4	42.9	40.0	100.0	9.0	43.2	38.3	100.0	8.2	42.9	38.1
	5	10	99.9	59.7	26.7	61.5	100.0	64.2	30.8	67.4	100.0	62.1	29.5	65.0
	40	10	100.0	16.3	19.9	31.0	100.0	18.0	19.2	32.7	100.0	19.9	19.7	36.4
	60	10	100.0	15.4	26.9	35.1	100.0	10.2	26.9	31.7	100.0	12.8	28.5	35.1
	80	10	100.0	8.0	41.9	35.1	100.0	7.5	42.2	36.1	100.0	9.6	41.3	39.8
	k_2	β_0	$\tau_1 = .15, \tau_2 = .3$ T_{2SLS}	AR	K	CLR	$\tau_1 = .2, \tau_2 = .3$ T_{2SLS}	AR	K	CLR	$\tau_1 = .3, \tau_2 = .3$ T_{2SLS}	AR	K	CLR
$\eta = .001$	5	0	99.9	68.6	31.3	38.8	99.8	79.7	36.0	45.8	99.9	92.7	43.6	56.1
	40	0	100.0	22.5	22.0	22.0	100.0	25.1	21.7	22.9	100.0	40.5	24.8	25.5
	60	0	100.0	14.8	29.9	29.9	100.0	15.6	27.5	27.8	100.0	20.4	34.2	34.8
	80	0	100.0	9.8	49.3	49.3	100.0	10.7	44.5	45.1	100.0	12.4	47.5	47.7
	5	10	100.0	72.2	31.1	75.8	100.0	75.2	36.9	77.9	100.0	92.0	46.4	92.7
	40	5	100.0	20.1	22.3	31.7	100.0	23.9	22.3	40.9	100.0	39.3	23.3	58.3
	60	5	100.0	14.7	28.0	37.2	100.0	16.1	27.6	39.4	100.0	21.7	31.4	48.9
	80	5	100.0	9.5	44.6	39.8	100.0	10.4	44.8	42.9	100.0	15.3	46.9	52.8
	5	10	100.0	70.3	31.6	72.5	100.0	79.0	38.2	80.6	100.0	91.4	44.4	92.5
	40	10	100.0	21.8	19.4	39.6	100.0	23.7	19.4	43.6	100.0	36.2	26.1	56.4
	60	10	100.0	12.7	28.5	35.0	100.0	15.6	29.0	39.4	100.0	21.7	31.0	47.3
	80	10	100.0	9.5	43.8	40.1	100.0	11.7	46.5	42.3	100.0	14.8	49.0	51.2

Table 2 (continued) : Percent Rejected at nominal level of 5 %

	k_2	β_0	$\tau_1 = 0, \tau_2 = .3$				$\tau_1 = .05, \tau_2 = .3$				$\tau_1 = .1, \tau_2 = .3$			
			T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\eta = .5$	5	0	69.9	55.8	26.4	26.3	74.7	57.7	34.5	34.5	79.5	64.2	42.4	42.4
	40	0	29.7	18.4	5.4	5.2	36.0	16.8	8.3	8.6	33.3	22.2	8.6	8.8
	60	0	19.2	12.9	7.9	8.0	22.3	12.0	7.9	7.8	21.3	11.4	7.6	8.0
	80	0	2.3	6.7	6.9	7.2	4.1	9.5	9.5	9.6	3.5	8.5	8.2	8.4
	5	5	69.1	62.1	25.9	46.6	76.7	62.0	32.6	52.6	81.5	65.8	44.9	62.0
	40	5	33.5	17.2	7.0	9.9	33.9	14.9	7.9	11.0	34.3	19.3	9.5	12.5
	60	5	19.6	11.1	6.6	8.5	19.2	12.0	7.2	10.3	20.4	12.7	7.3	9.9
	80	5	3.5	8.7	8.9	10.7	2.5	8.4	7.7	10.1	2.2	8.9	7.2	10.1
	5	10	70.6	63.3	28.8	66.4	77.0	60.1	35.0	63.5	81.3	65.5	42.3	68.1
	40	10	32.7	16.6	7.1	33.2	33.6	17.2	6.4	30.3	38.3	19.5	8.8	34.6
	60	10	19.7	10.3	6.7	30.5	18.3	11.3	6.8	34.3	21.7	12.7	7.5	33.8
	80	10	2.6	8.4	6.3	37.2	3.8	7.6	8.5	38.4	3.2	6.7	7.7	38.2

	k_2	β_0	$\tau_1 = .15, \tau_2 = .3$				$\tau_1 = .2, \tau_2 = .3$				$\tau_1 = .3, \tau_2 = .3$			
			T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\eta = .5$	5	0	87.5	69.8	53.2	53.6	90.8	77.9	58.2	58.5	97.5	90.7	77.2	77.0
	40	0	40.1	23.4	8.2	8.5	40.4	24.7	10.0	10.6	47.9	38.0	13.4	13.8
	60	0	22.1	12.4	7.9	7.8	22.8	18.2	8.7	8.9	28.8	23.1	11.6	12.1
	80	0	2.9	9.3	8.2	8.3	3.7	9.7	9.8	10.1	4.1	12.9	10.4	10.7
	5	5	84.0	71.5	52.4	66.0	91.9	77.6	63.3	77.8	96.7	91.5	75.9	89.5
	40	5	39.0	20.9	8.7	12.2	40.4	23.7	9.8	14.3	48.4	36.5	12.3	19.3
	60	5	24.9	12.7	10.1	13.2	23.0	15.0	8.3	11.9	25.7	25.9	9.1	13.8
	80	5	3.2	10.8	8.8	12.1	2.7	9.2	7.4	11.4	4.3	14.1	9.1	14.2
	5	10	88.4	73.6	52.1	77.3	92.9	76.2	59.0	79.8	97.7	91.2	75.5	93.0
	40	10	38.6	23.0	8.2	38.3	40.8	24.7	10.1	41.0	51.3	38.5	12.3	59.3
	60	10	20.1	13.7	7.7	37.4	23.3	17.0	7.8	41.3	27.0	22.7	9.0	49.9
	80	10	2.6	9.0	8.4	39.7	4.4	11.0	9.1	41.8	4.1	11.4	11.5	47.9

4.2 The effect of instrument exclusion

Dufour and Taamouti(2005) show that Anderson-Rubin-type procedures are robust to instrument exclusion and that Kleibergen's K-test statistic and Moreira's conditional likelihood ratio CLR-test can be seriously affected by the omission of a relevant instrument. In this section, we study the effect of missing endogenous instrument on T_{2SLS}, AR-type, K and CLR test-statistics. The data generating process is :

$$y = Y\beta + u, \quad Y = X_2\Pi_2 + X_3\gamma + w \tag{4.5}$$

$$(u_t, w_t)' \overset{i.i.d}{\sim} N(0, \Sigma), \Sigma = \begin{pmatrix} 1 & 0.95 \\ 0.95 & 1 \end{pmatrix} \tag{4.6}$$

$$\tilde{X}_3 = \xi_1 X_0 + \tau_1 u, \quad X_3 = M(X_2)\tilde{X}_3 \tag{4.7}$$

where $\xi_1 = \sqrt{1-\tau_1^2}$, $X_{0t} \overset{i.i.d}{\sim} N(0,1)$, X_2 is a $T \times k_2$ matrix of excluded exogenous instruments, X_3 is a $T \times 1$ omitted instrument vector which is not taken into account when computing the different statistics. $M(B) = I - P(B)$, $P_B = B(B'B)^{-1}B'$ is the projection matrix on the space spanned by the columns of B. From (4.7), X_3 is orthogonal to X_2 and is endogenous (invalid) unless $\tau_1 = 0$. Once again, the matrix Π_2 is such that $\Pi_2 = \eta\Pi$, where η takes the value 0.001 (design of weak identification) or 1 (design of strong identification), and Π is a $k_2 \times 1$ vector of one. The correlation coefficient ρ between u and w is again set equal to 0.95. So, the variable Y is endogenous and the instrumental variables X_2 are necessary. The number of instruments k_2 varies from 5 to 40. The parameters γ values are set at 3 and $\tau_1 \in \{0, .05, .1, .15, .2, .3\}$. We also consider three null hypotheses $H_0 = \beta_0$ for β_0 =10, 5 and 0. The sample size is $T = 100$. The number of replications is $N = 1000$ and the conditional LR critical values are computed using the same number of replications.
The results are presented in Tables 3-4 below. The main observations from these results are that : (i) AR-test statistic is robust to relevant instrument exclusion even if the excluded instrument is invalid whereas the t-test, K and CLR tests present size distortions [as expected from the theory, Dufour and Taamouti (2005)], with empirical rejection frequencies as high as 94.3 % for the t-test, 99.1 % for the K-statistic and 13.9 % for CLR-statistic; (ii) CLR-type procedure is more robust to instrument omission than K-type procedure; (iii) When the excluded instrument is valid, K-type procedure is more robust to instrument omission than t-test whereas t-test is more robust than K-test when the excluded instrument is endogenous.

Table 3 : Missing instruments (Percent Rejected at Nominal level of 5 %)

	k_2	β_0	$\tau_1 = 0$				$\tau_1 = 0.05$				$\tau_1 = 0.1$			
			T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\gamma = 3$ and $\eta = 0.001$	5	0	0.0	6.8	23.2	7.6	0.0	3.3	23.2	4.8	0.0	5.5	28.1	6.5
	10	0	1.7	3.9	42.1	4.9	1.1	4.6	46.1	6.6	1.0	4.6	52.2	6.1
	20	0	22.3	6.7	65.2	9.8	27.1	5.8	66.1	9.6	22.0	5.3	73.1	9.3
	40	0	93.3	4.0	77.5	11.6	93.4	4.6	78.5	10.7	94.8	5.2	84.5	12.6
	5	5	0.1	5.4	21.1	6.3	0.1	4.5	27.1	5.2	0.1	4.6	28.4	6.2
	10	5	0.9	4.6	42.2	6.4	0.8	4.3	44.0	5.7	0.5	4.6	56.0	5.8
	20	5	22.2	4.4	61.0	7.3	24.0	5.0	68.6	8.4	22.9	5.3	74.3	9.4
	40	5	94.3	4.4	76.5	12.8	93.2	5.2	76.4	12.7	94.1	5.0	83.4	12.0
	5	10	0.1	5.7	22.8	6.4	0.0	4.0	23.4	4.7	0.0	4.2	29.5	4.5
	10	10	0.8	4.1	41.2	5.7	0.8	4.4	43.8	6.4	1.7	4.8	51.9	6.7
	20	10	23.7	5.7	65.4	9.5	23.0	4.6	63.8	7.8	24.6	6.6	73.5	9.7
	40	10	92.6	5.0	79.4	12.5	92.2	5.5	82.5	10.9	93.4	4.8	83.5	12.6
			$\tau_1 = 0.15$				$\tau_1 = 0.2$				$\tau_1 = 0.3$			
	k_2	β_0	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\gamma = 3$ and $\eta = 0.001$	5	0	0.0	5.9	38.0	7.2	0.1	6.2	40.7	7.4	0.1	5.9	42.8	8.6
	10	0	1.8	4.6	63.4	7.0	0.9	5.0	74.9	6.6	1.0	5.5	86.4	6.7
	20	0	21.9	4.5	83.9	7.0	23.0	4.2	90.3	8.4	24.2	5.3	98.0	8.5
	40	0	94.9	5.1	89.0	13.9	93.8	5.7	93.3	13.8	94.1	5.7	98.1	11.2
	5	5	0.0	6.6	37.5	7.6	0.0	5.4	41.0	6.2	0.0	5.2	50.2	5.9
	10	5	1.7	6.1	63.7	8.0	1.1	5.1	75.1	6.6	0.9	4.8	86.2	7.2
	20	5	23.7	4.0	82.1	6.9	21.5	5.6	90.5	10.0	25.1	6.0	99.0	8.9
	40	5	94.1	5.8	88.5	12.6	94.2	5.3	93.6	13.9	93.5	5.1	99.1	10.9
	5	10	0.0	4.7	35.3	5.8	0.1	4.9	38.5	5.8	0.0	5.1	47.8	6.5
	10	10	1.2	5.0	63.6	6.8	0.8	4.8	74.0	6.6	1.9	5.5	86.9	8.0
	20	10	22.7	4.3	82.7	6.8	22.6	4.5	90.4	8.1	25.6	4.5	98.1	8.6
	40	10	93.5	4.4	88.9	11.8	94.1	5.7	92.4	11.8	94.2	4.8	99.0	11.0

Table 4: Missing instruments (Percent Rejected at Nominal level of 5 %)

			$\tau_1 = 0$				$\tau_1 = 0.05$				$\tau_1 = 0.1$			
	k_2	β_0	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\gamma = 3$ and $\eta = 1$	5	0	9.5	6.2	6.3	7.6	9.1	5.1	5.5	5.9	9.1	4.1	6.0	5.1
	10	0	9.5	4.4	4.8	6.9	8.5	5.3	4.6	7.1	11.1	5.3	6.7	6.7
	20	0	7.7	4.8	5.6	8.5	8.3	4.9	4.7	7.9	10.1	5.1	5.3	7.6
	40	0	8.2	4.9	7.2	11.3	6.4	4.9	6.2	10.5	7.4	5.1	6.6	11.5
	5	5	10.6	6.8	6.5	7.6	7.6	3.3	4.2	4.8	10.8	5.5	5.9	6.5
	10	5	8.2	3.9	4.9	4.9	10.4	4.6	5.5	6.6	10.9	4.6	6.1	6.1
	20	5	9.6	4.9	5.2	8.3	7.8	5.6	4.5	9.3	8.2	4.3	5.6	7.7
	40	5	7.8	4.0	6.6	11.6	6.8	4.6	6.0	10.7	5.9	5.2	6.6	12.4
	5	10	10.2	5.4	6.0	6.3	8.6	4.5	4.3	5.2	9.2	4.6	4.8	6.2
	10	10	9.9	4.6	4.6	6.4	10.4	4.3	5.1	5.7	9.9	4.6	5.1	5.8
	20	10	10.4	4.4	6.2	7.3	9.4	5.0	4.5	8.4	10.0	5.3	6.0	9.4
	40	10	7.1	4.4	5.8	12.8	9.3	5.2	7.0	12.7	8.0	5.0	6.0	12.0
			$\tau_1 = 0.15$				$\tau_1 = 0.2$				$\tau_1 = 0.3$			
	k_2	β_0	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR	T_{2SLS}	AR	K	CLR
$\gamma = 3$ and $\eta = 1$	5	0	8.0	4.5	4.6	5.4	10.9	3.7	5.6	4.6	9.7	4.4	6.3	5.8
	10	0	11.1	4.7	6.5	6.8	9.8	4.5	6.1	6.5	10.8	6.5	6.7	8.0
	20	0	9.2	6.0	6.4	10.0	9.9	4.1	6.4	7.9	8.5	5.0	7.7	8.3
	40	0	8.2	4.4	7.0	10.6	9.2	5.5	7.6	11.4	8.4	6.1	9.6	12.1
	5	5	9.7	5.9	5.9	7.2	9.8	6.2	4.8	7.4	9.5	6.1	5.0	6.9
	10	5	10.0	4.6	5.0	7.0	9.5	5.0	4.0	6.6	10.0	5.5	5.3	6.7
	20	5	9.9	3.4	6.2	5.7	10.7	5.6	6.7	9.0	11.5	4.4	6.6	7.6
	40	5	6.5	5.1	7.9	13.8	6.5	5.7	7.2	13.8	7.0	5.7	9.6	11.0
	5	10	11.6	6.6	6.3	7.6	10.9	5.4	5.8	6.2	10.2	5.2	6.4	5.9
	10	10	10.5	6.1	6.6	8.0	9.7	5.1	5.2	6.6	9.3	4.8	4.4	7.2
	20	10	9.3	4.0	7.3	6.9	8.8	5.6	6.8	10.0	10.4	6.0	7.7	8.9
	40	10	7.0	5.8	7.5	12.6	8.0	5.3	7.7	13.9	7.3	5.1	10.6	10.9

5 Asymptotic distributions of the test statistics under instrument endogeneity

Let us reconsider the following form of the model (2.1)-(2.2) :

$$y = Y\beta + u, \quad Y = X_2\Pi_2 + v \tag{5.1}$$

where y is a $T \times 1$ vector of observations on the dependent variable; Y is a $T \times G$ matrix of observations on the explanatory endogenous variables ($G \geq 1$); X_2 is a $T \times k_2$ matrix of observations on the excluded instruments; u and v are respectively a $T \times 1$ vector and a $T \times G$ matrix of disturbances; β is a $G \times 1$ vector of unknown coefficients and Π_2 is a $k_2 \times G$ matrices of unknown coefficients.

$$v = u\rho' + \varepsilon\sigma_\varepsilon', \quad X_2 = X_0\xi + u\tau' \tag{5.2}$$

where X_0 is a $T \times k_2$ matrix of exogenous variables, ε is a $T \times 1$ vector i.i.d $N(0, 1)$ variables uncorrelated with u; $\rho = (\rho_j)_{1\leq j\leq G}$ and $\sigma_\varepsilon = (\sigma_{\varepsilon j})_{1\leq j\leq G}$ are $G \times 1$ vectors of unknown coefficients; $\tau = (\tau_i)_{1\leq i\leq k_2}$ is a $k_2 \times 1$ vector; and $\xi = [diag(\xi_i)]_{1\leq i\leq k_2}$ is a $k_2 \times k_2$ matrix of unknown coefficients. Assume that

$$\begin{pmatrix} u_t & v_t \end{pmatrix} \sim N(0, \Sigma) \tag{5.3}$$

where $\Sigma = \begin{pmatrix} \sigma_{11} & \rho' \\ \rho & \Sigma_{22} \end{pmatrix}$ and from (5.2), $\Sigma_{22} = \sigma_{11}\rho\rho' + \sigma_\varepsilon\sigma_\varepsilon'$

Finally, assume also that

$$\frac{1}{T}X_0'X_0 \to Q_{X_0X_0} \quad \text{and that}\, \xi \to \xi_0 \,\text{when}\, \tau \to 0 \tag{5.4}$$

From (5.2), we have $cov(u, v) = \rho$ and $cov(u, X_2) = \tau$. The parameter ρ controls the endogeneity of the variable y whereas τ controls instrument X_2 endogeneity. If $\tau_j = 0$ for some $j \in \{1, 2, ..., k_2\}$, then instrument X_{2j} is valid, otherwise, it is invalid (endogenous). If the model contains only one instrument and one endogenous variable ($G = k_2 = 1$), the three statistics AR, K and CLR are identical. So, in this case, Kleibergen's K-test statistic and Moreira's conditional likelihood ratio CLR-test statistic will be pivotal in finite sample if the instrument is valid ($\tau = 0$). When $k_2 > 1$, Kleibergen's K-test statistic is not pivotal in finite sample but is asymptotically pivotal whereas Moreira's CLR-test statistic is not pivotal even asymptotically.

Under conditions (5.1)-(5.4), the four statistics T_{2SLS}, AR, K and CLR for testing $H_0 : \beta = \beta_0$ given in (3.2) and (3.4) are now :

$$T_{2SLS}(\beta_{0j}) = \frac{\hat{\beta}_j - \beta_{0j}}{\hat{\sigma}_{jj}} \tag{5.5}$$

where $\hat{\beta}_j$, $j = 1, ..., G$ is the the jth element of the $G \times 1$ vector $[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}Y'X_2(X_2'X_2)^{-1}X_2'y$ and $\hat{\sigma}_{jj}$ is the squared root of the product of the $jjth$ element of the $G \times G$ matrix $[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}$ to

$\frac{1}{T-k_2}\left[y'y - y'X_2(X_2'X_2)^{-1}X_2'Y[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}Y'X_2(X_2'X_2)^{-1}X_2'y\right]$

$$AR(\beta_0) = \frac{(T-k_2)}{k_2}\frac{(y-Y\beta_0)'P_{X_2}(y-Y\beta_0)}{(y-Y\beta_0)'M_{X_2}(y-Y\beta_0)} \tag{5.6}$$

$$K(\beta_0) = \frac{(y-Y\beta_0)'P_{\tilde{Y}(\beta_0)}(y-Y\beta_0)}{\frac{1}{T-k_2}(y-Y\beta_0)'M_{X_2}(y-Y\beta_0)} \tag{5.7}$$

where
$\tilde{Y}(\beta_0) = X_2\tilde{\Pi}(\beta_0)$, $\tilde{\Pi}(\beta_0) = (X_2'X_2)^{-1}X_2'[Y - (y-Y\beta_0)\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}]$,
$S_{uu}(\beta_0) = \frac{1}{T-k_2}(y-Y\beta_0)'M_{X_2}(y-Y\beta_0)$, $S_{uv}(\beta_0) = \frac{1}{T-k_2}(y-Y\beta_0)'M_{X_2}Y$,
$P_B = B(B'B)^{-1}B'$ and $M_B = I - P_B$, for any matrix columns B.

$$CLR(\beta_0) = \hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\lambda}^{min} \tag{5.8}$$

where $\hat{\lambda}^{min}$ is the smallest eigenvalue of $(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}})$,
$\hat{\mathcal{S}} = (X_2'X_2)^{-1/2}X_2'\underline{Y}b_0(b_0'\hat{\Delta}b_0)^{-1/2}$, $\hat{\mathcal{J}} = (X_2'X_2)^{-1/2}X_2'\underline{Y}\hat{\Delta}^{-1}A_0(A_0'\hat{\Delta}^{-1}A_0)^{-1/2}$,
$\hat{\Delta} = \underline{Y}'M_{X_2}\underline{Y}/(T-k_2)$ is the estimation of covariance matrix of the reduced-form errors.
$\underline{Y} = [y, Y]$, $b_0 = [1, -\beta_0]'$ and $A_0 = [\beta_0, I_G]'$.
When $G = 1$, this statistic can be simplified as

$$CLR(\beta_0) = \frac{1}{2}\left[\hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\mathcal{J}}'\hat{\mathcal{J}} + \sqrt{(\hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\mathcal{J}}'\hat{\mathcal{J}})^2 - 4[\hat{\mathcal{S}}'\hat{\mathcal{S}}\hat{\mathcal{J}}'\hat{\mathcal{J}} - (\hat{\mathcal{S}}'\hat{\mathcal{J}})^2]}\right] \tag{5.9}$$

where $a_0 = [\beta_0, 1]'$.
In this section, we are interested in the asymptotic distributions of these four statistics under the null hypothesis. Now, let $W_{1T} = X_0'u$ and $W_{2T} = X_0'\varepsilon$. Clearly, since X_0 is exogenous, assumptions (5.1)-(5.4) imply that $\frac{1}{\sqrt{T}}W_{1T} \overset{a}{\sim} W_1$, $\frac{1}{\sqrt{T}}W_{2T} \overset{a}{\sim} W_2$; $W_1 \overset{i.i.d}{\sim} N(0, \sigma_{11}Q_{X_0X_0})$ and $W_2 \overset{i.i.d}{\sim} N(0, Q_{X_0X_0})$. The following proposition gives the asymptotic distributions under the null hypothesis of some statistics when the instrument is nearly invalid [9].

Proposition 5.1 *If assumptions* (5.1)-(5.4) *hold, under* H_0,

(i) if $\Pi_2 = \Pi_0/\sqrt{T}$ *and* $\tau = \tau_0/\sqrt{T}$
where Π_0 *and* τ_0 *are respectively a* $k_2 \times G$ *constant matrix and a* $k_2 \times 1$ *constant vector as* $T \to \infty$*, then*

[9] we emphasize two cases : the first case is about both instrument weakness and invalidity whereas the second treats both instrument strength and invalidity.

$$
\begin{array}{ccc}
Y'X_2(X_2'X_2)^{-1}X_2'Y & \overset{a}{\sim} & \Psi_0'Q_0\Psi_0 \\
Y'X_2(X_2'X_2)^{-1}X_2'y & \overset{a}{\sim} & \Psi_0'Q_0\Psi_0\beta_0 + \Psi_0'Q_0^{1/2}(\xi_0'W_1 + \sigma_{11}\tau_0) \\
\hat{\beta}_{2SLS} - \beta_0 & \overset{a}{\sim} & (\Psi_0'Q_0\Psi_0)^{-1}\Psi_0'(\xi_0'W_1 + \sigma_{11}\tau_0) \\
\frac{1}{T}\widehat{\Omega} & \overset{P}{\to} & \varphi_0 \\
(y - Y\beta_0)'P_{\tilde{Y}(\beta_0)}(y - Y\beta_0)(\beta_0) & \overset{a}{\sim} & (\xi_0'W_1 + \sigma_{11}\tau_0)'\Gamma_0(\Gamma_0'Q_0\Gamma_0)^{-1}\Gamma_0'(\xi_0'W_1 + \sigma_{11}\tau_0) \\
(y - Y\beta_0)'P_{X_2}(y - Y\beta_0) & \overset{P}{\to} & (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0) \\
\frac{1}{T}(y - Y\beta_0)'M_{X_2}(y - Y\beta_0) & \overset{a}{\sim} & \sigma_{11} \\
\hat{\mathcal{S}}'\hat{\mathcal{S}} & \overset{a}{\sim} & (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0) \\
\frac{1}{\sqrt{T}}X_2'\underline{Y} & \overset{a}{\sim} & \bar{Q}_0 \\
\hat{\Delta} & \overset{a}{\sim} & \begin{bmatrix} \varphi_0 & \phi_0' \\ \phi_0 & \chi_0 \end{bmatrix} \\
\hat{\mathcal{J}}'\hat{\mathcal{J}} & \overset{a}{\sim} & \Sigma_0 \\
\hat{\mathcal{S}}'\hat{\mathcal{J}} & \overset{a}{\sim} & \Theta_0' \\
(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}}) & \overset{a}{\sim} & \bar{\Lambda}_0
\end{array}
$$

where

$$\widehat{\Omega} = \left[y'y - y'X_2(X_2'X_2)^{-1}X_2'Y[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}Y'X_2(X_2'X_2)^{-1}X_2'y\right]$$

$$\begin{aligned}
\varphi_0 &= \beta_0'(\sigma_{11}\rho\rho' + \sigma_\varepsilon\sigma_\varepsilon')\beta_0 + \sigma_{11}\beta_0'\rho + \sigma_{11}\rho'\beta_0 + \sigma_{11}, \\
\Psi_0 &= \Pi_0 + Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon' \\
\Gamma_0 &= \Pi_0 + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon' \\
\bar{Q}_0 &= \left[\begin{array}{cc} Q_0\Psi_0\beta_0 + \xi_0'W_1 + \sigma_{11}\tau_0 & Q_0\Psi_0 \end{array}\right] \\
\phi_0 &= \chi_0'\beta_0 + \sigma_{11}\rho \\
\chi_0 &= \sigma_{11}\rho\rho' + \sigma_\varepsilon\sigma_\varepsilon' \\
\Sigma_0 &= (A_0'\Delta_0^{-1}A_0)^{-1/2}A_0'\Delta_0^{-1}\bar{Q}_0'Q_0^{-1}\bar{Q}_0\Delta_0^{-1}A_0(A_0'\Delta_0^{-1}A_0)^{-1/2} \\
\Theta_0 &= (A_0'\Delta_0^{-1}A_0)^{-1/2}A_0'\Delta_0^{-1}\bar{Q}_0'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0) \\
\bar{\Lambda}_0 &= \left[\begin{array}{cc} (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0) & \Theta_0' \\ \Theta_0 & \Sigma_0 \end{array}\right] \\
A_0 &= [\beta_0, I_G]' \\
Q_0 &= \xi_0'Q_{X_0X_0}\xi_0
\end{aligned}$$

(ii) if $\Pi_2 = \Pi_0$ *and* $\tau = \tau_0/\sqrt{T}$
where $\Pi_0 \neq 0$ *and* τ_0 *are respectively a* $k_2 \times G$ *constant matrix and a* $k_2 \times 1$ *constant vector as* $T \to \infty$, *then*

$$\begin{aligned}
\tfrac{1}{T}(Y'X_2(X_2'X_2)^{-1}X_2'Y) &\xrightarrow{P} \Pi_0'Q_0\Pi_0 \\
\tfrac{1}{T}(Y'X_2(X_2'X_2)^{-1}X_2'y) &\xrightarrow{P} \Pi_0'Q_0\Pi_0\beta_0 \\
\sqrt{T}(\hat{\beta}_{2SLS} - \beta_0) &\overset{a}{\sim} (\Pi_0'Q_0\Pi_0)^{-1}\Pi_0'(\xi_0'W_1 + \sigma_{11}\tau_0) \\
\tfrac{1}{T}\widehat{\Omega} &\xrightarrow{P} \bar{\varphi}_0 - \beta_0'\Pi_0'Q_0\Pi_0(\Pi_0'Q_0\Pi_0)^{-1}\Pi_0'Q_0\Pi_0\beta_0
\end{aligned}$$

$$
\begin{array}{ccc}
(y - Y\beta_0)' P_{\tilde{Y}(\beta_0)} (y - Y\beta_0) & \overset{a}{\sim} & (\xi_0' W_1 + \sigma_{11}\tau_0)' \Pi_0 (\Pi_0' Q_0 \Pi_0)^{-1} \Pi_0' (\xi_0' W_1 + \sigma_{11}\tau_0) \\
\hat{\mathcal{S}}'\hat{\mathcal{S}} & \overset{a}{\sim} & (\xi_0' W_1 + \sigma_{11}\tau_0)' Q_0^{-1} (\xi_0' W_1 + \sigma_{11}\tau_0) \\
(y - Y\beta_0)' P_{X_2} (y - Y\beta_0) & \overset{a}{\sim} & (\xi_0' W_1 + \sigma_{11}\tau_0)' Q_0^{-1} (\xi_0' W_1 + \sigma_{11}\tau_0) \\
\frac{1}{T}(y - Y\beta_0)' M_{X_2} (y - Y\beta_0) & \overset{a}{\sim} & \sigma_{11} \\
\frac{1}{T}(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}}) & \overset{P}{\to} & \widehat{\Lambda}_0 \\
\frac{1}{T} X_2' \underline{Y} & \overset{P}{\to} & \widehat{Q}_0 \\
\hat{\Delta} & \overset{P}{\to} & \left[\begin{array}{cc} \bar{\varphi}_0 - \beta_0' \Pi_0' Q_0 \Pi_0 \beta_0 & \bar{\phi}_0' \\ \bar{\phi}_0 & \bar{\Psi}_0 \end{array} \right] \\
\frac{\hat{\mathcal{J}}'\hat{\mathcal{J}}}{T} & \overset{P}{\to} & \widetilde{\Sigma}_0 \\
\frac{\hat{\mathcal{S}}'\hat{\mathcal{J}}}{T} & \overset{P}{\to} & \widehat{\Theta}_0'
\end{array}
$$

where

$$
\widehat{\Omega} = \left[y'y - y'X_2(X_2'X_2)^{-1}X_2'Y[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}Y'X_2(X_2'X_2)^{-1}X_2'y \right]
$$

$$
\begin{array}{rcl}
\bar{\varphi}_0 & = & \beta_0'(Q_0\Pi_0 + \chi_0)\beta_0 + \sigma_{11}\beta_0'\rho + \sigma_{11}\rho'\beta_0 + \sigma_{11}, \\
\widehat{Q}_0 & = & \left[\begin{array}{cc} Q_0\Pi_0\beta_0 & Q_0\Pi_0 \end{array} \right] \\
\widehat{\Delta}_0 & = & \left[\begin{array}{cc} \bar{\varphi}_0 - \beta_0'\Pi_0'Q_0\Pi_0\beta_0 & \bar{\phi}_0' \\ \bar{\phi}_0 & \bar{\Psi}_0 \end{array} \right] \\
\bar{\phi}_0' & = & \beta_0'(\Pi_0' Q_{X_0X_0}\Pi_0 + \chi_0) + \sigma_{11}\rho' - \beta_0'\Pi_0' Q_{X_0X_0}\Pi_0 \\
\chi_0 & = & \sigma_{11}\rho\rho' + \sigma_\varepsilon\sigma_\varepsilon' \\
\bar{\Psi}_0 & = & \beta_0'(\Pi_0'Q_0\Pi_0 + \chi_0) - \Pi_0'Q_0\Pi_0 \\
\widetilde{\Sigma}_0 & = & (A_0'\widehat{\Delta}_0^{-1}A_0)^{-1/2} A_0'\widehat{\Delta}_0^{-1}\widehat{Q}_0' Q_0^{-1}\widehat{Q}_0\widehat{\Delta}_0^{-1}A_0 (A_0'\widehat{\Delta}_0^{-1}A_0)^{-1/2}
\end{array}
$$

$$\widehat{\Theta}_0' = (\xi_0' W_1 + \sigma_{11}\tau_0)' Q_0^{-1} \widehat{Q}_0 \widehat{\Delta}_0^{-1} A_0 (A_0' \widehat{\Delta}_0^{-1} A_0)^{-1/2}$$

$$\widehat{\Lambda}_0 = \begin{bmatrix} 0 & \widehat{\Theta}_0' \\ \widehat{\Theta}_0 & \widetilde{\Sigma}_0 \end{bmatrix}$$

$$A_0 = [\beta_0, I_G]'$$

$$Q_0 = \xi_0' Q_{X_0 X_0} \xi_0$$

Proofs of Propositions and Theorems are given in Appendix.

The following theorem gives the asymptotic distributions of the four statistics T_{2SLS}, AR, K and CLR. Proposition 5.1.

THEOREM **5.1** *If proposition 5.1 hold and under H_0,*

(i) if $\Pi_2 = \Pi_0/\sqrt{T}$ and $\tau = \tau_0/\sqrt{T}$ where Π_0 and τ_0 are respectively a $k_2 \times G$ constant matrix and a $k_2 \times 1$ constant vector as $T \to \infty$, then

$$T_{2SLS}(\beta_{0j}) \overset{a}{\sim} \varphi_0^{-1/2} \bar{\Upsilon}_j$$

$$AR(\beta_0) \overset{a}{\sim} \tfrac{1}{k_2}\chi^2(k_2, \mu_1)$$

$$K(\beta_0)|_{W_2} \overset{a}{\sim} \chi^2(G, \mu_2)$$

$$CLR(\beta_0) \overset{a}{\sim} \tfrac{1}{k_2}\chi^2(k_2, \mu_1) - \hat{\lambda}_0^{min}$$

where $(\Psi_0' Q_0 \Psi_0)^{-1/2} \Psi_0' (\xi_0' W_1 + \sigma_{11}\tau_0) = [\bar{\Upsilon}_j]_{1 \le j \le G}$ and

φ_0, Ψ_0, Γ_0, $\bar{Q}_0$, χ_0, ϕ_0, Σ_0, Θ_0, $\bar{\Lambda}_0$, A_0, and Q_0 are defined in Proposition 5.1-(i).

$$\mu_1 = \sigma_{11} \tau_0' Q_0^{-1} \tau_0$$

$$\mu_2 = \sigma_{11} \tau_0' \Gamma_0 (\Gamma_0' Q_0 \Gamma_0)^{-1} \Gamma_0' \tau_0 \quad ;$$

$\hat{\lambda}_0^{min}$ is the smallest eigenvalue of $\bar{\Lambda}_0$, W_1 and W_2 are respectively i.i.d $N(0, \sigma_{11} Q_{X_0 X_0})$ and $N(0, Q_{X_0 X_0})$.

(ii) if $\Pi_2 = \Pi_0$ and $\tau = \tau_0/\sqrt{T}$ where $\Pi_0 \neq 0$ and τ_0 are respectively a $k_2 \times G$ constant matrix and a $k_2 \times 1$ constant vector as $T \to \infty$, then

$$T_{2SLS}(\beta_{0j}) \overset{a}{\sim} \bar{\varphi}_0^{-1/2}\widetilde{\Upsilon}_j$$

$$AR(\beta_0) \overset{a}{\sim} \frac{1}{k_2}\chi^2(k_2,\mu_1)$$

$$K(\beta_0) \overset{a}{\sim} \chi^2(G,\mu_3)$$

$$CLR(\beta_0) \overset{a}{\sim} \frac{1}{k_2}\chi^2(k_2,\mu_1) - \widetilde{\lambda}_0^{min}$$

where $(\Pi_0'Q_0\Pi_0)^{-1/2}\Pi_0'(\xi_0'W_1+\sigma_{11}\tau_0) = [\widetilde{\Upsilon}_j]_{1\leq j\leq G}$ *and*

$\bar{\varphi}_0$, $\widehat{\Delta}_0$, $\widetilde{\Sigma}_0$, $\bar{\phi}_0$, $\widehat{\Theta}_0$, $\widehat{\Lambda}_0$, A_0 *are defined in Proposition 5.1-(ii),* μ_1 *and* μ_2 *are given in Theorem 5.1-(i),*
$\mu_3 = \sigma_{11}\tau_0'\Pi_0(\Pi_0'Q_0\Pi_0)^{-1}\Pi_0'\tau_0$;
$\widetilde{\lambda}_0^{min}$ *is the smallest eigenvalue of* $\widetilde{\Lambda}_0$, W_1 *is i.i.d* $N(0,\sigma_{11}Q_{X_0X_0})$.

THEOREM 5.1 shows clearly that the asymptotic null distributions of the four statistics depend on the parameter τ_0 even if the instruments are strong. This means that like the statistics T_{2SLS} and CLR, the AR and K statistics are not pivotal even asymptotically when instruments are invalid. This is probably why the Monte Carlo experiments in Section 4 show serious size distortions. Moreover, the asymptotic null distribution of the AR-statistic does not depend on the instruments being weak or strong : it is distributed as a non central chi-square $[\frac{1}{k_2}\chi^2(k_2,\sigma_{11}\tau_0'Q_0^{-1}\tau_0)]$ and the non centrality parameter is $\sigma_{11}\tau_0Q_0^{-1}\tau_0'$. However, for the other three statistics, the asymptotic null distributions change according to whether instruments are weak or strong. When instruments are strong [THEOREM 5.1-(ii)], the K-statistic is distributed as a non central chi-square $[\chi^2(G,\sigma_{11}\tau_0\Pi_0(\Pi_0'Q_0\Pi_0)^{-1}\Pi_0'\tau_0')]$ and the non centrality parameter is $\sigma_{11}\tau_0\Pi_0(\Pi_0'Q_0\Pi_0)^{-1}\Pi_0'\tau_0'$. However, when the instruments are weak, unlike the result of Kleibergen (2002), the distribution of the K-statistic under the null hypothesis depends on W_2, but its conditional distribution is a noncentral chi-square with non centrality parameter $\sigma_{11}\tau_0'\Gamma_0(\Gamma_0'Q_0\Gamma_0)^{-1}\Gamma_0'\tau_0$, [i.e $K(\beta_0)|_{W_2} \overset{a}{\sim} \chi^2(G,\sigma_{11}\tau_0'\Gamma_0(\Gamma_0'Q_0\Gamma_0)^{-1}\Gamma_0'\tau_0)$].
When the instruments are endogenous, the conditional and unconditional distributions of the K-statistic depend not only on the endogeneity parameter τ_0, but also on σ_{11}, Π_0, σ_ε, Q_0 and W_2. It is worthwhile to note that the asymptotic null distributions of the statistics T_{2SLS} and CLR depend not only on τ_0, Q_0, Π_0, σ_{11}, σ_ε and W_2 but also on ρ and β_0 , as expected from the theory [Dufour (1997), Begger (1994), Phillips (1989), Wang and Zivot (1998), Kleibergen (2002), and Moreira (2003)].
When k_2=G=1, σ_{11}, Π_0, ρ, β_0, τ_0, σ_ε and Q_0 are scalars, we have :

COROLLARY **5.1** *Let* THEOREM 5.1 *holds, if k_2=G=1, then the three statistics $AR(\beta_0)$, $K(\beta_0)$, and $CLR(\beta_0)$ are identical and distributed as a non central chi-square with one degree of freedom and the non centrality parameter is* $\mu_0 = \frac{\sigma_{11}\tau_0^2}{Q_0}$, *i.e* $AR(\beta_0)= K(\beta_0) =CLR(\beta_0) \overset{a}{\sim} \chi^2(1,\mu_0)$.

COROLLARY **5.2** *Let* THEOREM 5.1 *holds, if* $\tau_0 = 0$, *then* $AR(\beta_0)\overset{a}{\sim} \frac{1}{k_2}\chi^2(k_2)$ *and* $K(\beta_0)\overset{a}{\sim} \chi^2(G)$ *irrespective of whether instruments are weak or not.*

When instruments are valid, i.e when $\tau_0 = 0$, the parameters of non centrality in the limiting distributions of the statistics AR(β_0) and K(β_0) vanishes as expected from the theory. Furthermore, if this assumption holds, AR(β_0) is exactly distributed as $F(k_2, T - k_2)$.

6 Conclusions

In this paper, we made comparative Monte Carlo experiments of the performance of the standard t-test, the Anderson and Rubin test , the Kleibergen's K-test, and the Moreira's conditional likelihood ratio test statistics when some of the instruments are invalid in the structural linear model. We show that when some of the instruments are endogenous, the inference procedures based on Anderson-Rubin statistic are globally more robust than the other procedures and that those based on the t-test statistic exhibit more size distortions. Furthermore, our results indicate that Kleibergen's K-statistic is globally more robust to invalid instruments than Moreira's conditional likelihood ratio statistic whereas the results reverse as far as it concerns the effect of missing instruments. Moreover, unlike the results of Dufour and Taamouti (2005), we find that when some of the instruments are endogenous the AR-test statistic is not anymore robust to missing instrument. We also find that the instrument invalidity is much more detrimental than the instrument weakness on the inference procedures based on AR, K and CLR statistics i.e. tests based on strong but invalid instruments have higher size distortions than those based on valid but weak instruments. This later result contrast with the findings of Kiviet and Niemcsyk (working paper). Finally, we derive explicit form for the asymptotic limiting null distributions of the four statistics and show that they depend on nuisance parameters, in particular, an additive parameter which characterizes the instrument endogeneity.

A Appendix

Proof of proposition 5.1

If assumptions (5.1)-(5.4) hold. Under the null hypothesis, we have

(i) if $\Pi_2 = \Pi_0/\sqrt{T}$ and $\tau = \tau_0/\sqrt{T}$
where Π_0, τ_0 are respectively a $k_2 \times G$ constant matrix and a $1 \times k_2$ constant vector as $T \to \infty$, then

$$X_2'X_2 = \xi' X_0' X_0 \xi + \xi' X_0' u \tau' + \tau u' X_0 \xi + \tau u' u \tau'$$

Since X_0 is exogenous, $u \overset{i.i.d}{\sim} N(0, \sigma_{11} I_T)$, then

$$\frac{1}{T} X_2'X_2 = \frac{1}{T}\xi' X_0' X_0 \xi + \frac{1}{T^{3/2}} \xi' X_0' u \tau_0' + \frac{1}{T^{3/2}} \tau_0 u' X_0 \xi + \frac{1}{T^2} \tau_0 u' u \tau_0' \xrightarrow{\mathbb{P}} \xi_0' Q_{X_0 X_0} \xi_0 = Q_0$$

$$\frac{1}{\sqrt{T}} X_2' u = \xi' \frac{X_0' u}{\sqrt{T}} + \tau_0 \frac{u'u}{T} \overset{a}{\sim} \xi_0' W_1 + \sigma_{11} \tau_0$$

$$\frac{1}{\sqrt{T}} X_2' \varepsilon = \xi' \frac{X_0' \varepsilon}{\sqrt{T}} + \tau_0' \frac{u' \varepsilon}{T} \overset{a}{\sim} \xi_0' W_2$$

$$\frac{1}{\sqrt{T}} X_2' Y = \frac{X_2' X_2}{T} \Pi_0 + \frac{X_2' u}{\sqrt{T}} \rho + \frac{X_2' \varepsilon}{\sqrt{T}} \sigma_\varepsilon \overset{a}{\sim} Q_0 \Psi_0 = Q_0 \Pi_0 + (\xi_0' W_1 + \sigma_{11} \tau_0) \rho' + \xi_0' W_2 \sigma_\varepsilon'$$

$$\frac{1}{\sqrt{T}} X_2' y = \frac{1}{\sqrt{T}} X_2' Y \beta_0 + \frac{1}{\sqrt{T}} X_2' u \overset{a}{\sim} [Q_0 \Pi_0 + (\xi_0' W_1 + \sigma_{11} \tau_0) \rho' + \xi_0' W_2 \sigma_\varepsilon'] \beta_0 + \xi_0' W_1 + \sigma_{11} \tau_0 = Q_0 \Psi_0 \beta_0 + \xi_0' W_1 + \sigma_{11} \tau_0$$

$$\frac{1}{T} v' v = \rho \frac{1}{T} u' u \rho' + \rho \frac{1}{T} u' \varepsilon \sigma_\varepsilon' + \frac{1}{T} \sigma_\varepsilon \varepsilon' u \rho' + \sigma_\varepsilon \frac{1}{T} \varepsilon' \varepsilon \sigma_\varepsilon' \xrightarrow{\mathbb{P}} \chi_0 = \sigma_{11} \rho \rho' + \sigma_\varepsilon \sigma_\varepsilon'$$

$$\frac{1}{T} Y' Y = \Pi_0' \frac{1}{T^2} X_2' X_2 \Pi_0 + \Pi_0' \frac{1}{T^{3/2}} X_2' v + \frac{1}{T^{3/2}} v' X_2 \Pi_0 + \frac{1}{T} v' v$$

$$\xrightarrow{\mathbb{P}} \chi_0 = \sigma_{11} \rho \rho' + \sigma_\varepsilon \sigma_\varepsilon'$$

$$\frac{1}{T} Y' u = \Pi_0' \frac{1}{T^{3/2}} X_2' u + \rho \frac{1}{T} u' u + \sigma_\varepsilon \frac{1}{T} \varepsilon' u \xrightarrow{\mathbb{P}} \sigma_{11} \rho$$

So,

$$\frac{1}{T} y' y = \beta_0' \frac{1}{T} Y' Y \beta_0 + \beta_0' \frac{1}{T} Y' u + \frac{1}{T} u' Y \beta_0 + \frac{1}{T} u' u \xrightarrow{\mathbb{P}} \beta_0' (\sigma_{11} \rho \rho' + \sigma_\varepsilon \sigma_\varepsilon') \beta_0 + \sigma_{11} \beta_0' \rho + \sigma_{11} \rho' \beta_0 + \sigma_{11} = \varphi_0$$

$$\frac{1}{T} X_2' Y = \frac{1}{\sqrt{T^3}} X_2' X_2 \Pi_0 + \frac{1}{T} X_2' v \xrightarrow{\mathbb{P}} 0$$

$$\frac{1}{T} X_2' y = \frac{1}{T} X_2' Y \beta_0 + \frac{1}{T} X_2' u \xrightarrow{\mathbb{P}} 0$$

$Y'X_2(X_2'X_2)^{-1}X_2'Y = (\frac{1}{\sqrt{T}}Y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{\sqrt{T}}X_2'Y) \overset{a}{\sim} [Q_0\Pi_0 + (\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + \xi_0'W_2\sigma_\varepsilon']'Q_0^{-1}[Q_0\Pi_0 + (\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + \xi_0'W_2\sigma_\varepsilon'] =$

$[\Pi_0 + Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon']'Q_0[\Pi_0 + Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon']$

$= \Psi_0'Q_0\Psi_0$

$Y'X_2(X_2'X_2)^{-1}X_2'y = (\frac{1}{\sqrt{T}}Y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{\sqrt{T}}X_2'y) \overset{a}{\sim} [Q_0\Pi_0 + (\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + \xi_0'W_2\sigma_\varepsilon']'Q_0^{-1}\left[\{Q_0\Pi_0 + (\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + \xi_0'W_2\sigma_\varepsilon'\}\beta_0 + \xi_0'W_1 + \sigma_{11}\tau_0\right] =$

$\Psi_0'Q_0\Psi_0\beta_0 + \Psi_0'Q_0^{1/2}(\xi_0'W_1 + \sigma_{11}\tau_0)$

$\hat{\beta}_{2SLS} - \beta_0 \overset{a}{\sim} (\Psi_0'Q_0\Psi_0)^{-1}\Psi_0'(\xi_0'W_1 + \sigma_{11}\tau_0) =$

$\left\{[\Pi_0 + Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon']'Q_0^{-1}[\Pi_0 + Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)\rho'Q_0^{-1}\xi_0'W_2\sigma_\varepsilon']\right\}^{-1} \times$

$[Q_0\Pi_0 + (\xi_0'W_1 + \sigma_{11}\tau_0)\rho' + \xi_0'W_2\sigma_\varepsilon']'Q_0^{-1}[\xi_0'W_1 + \sigma_{11}\tau_0]$

$\frac{1}{T-k_2}\left[y'y - y'X_2(X_2'X_2)^{-1}X_2'Y[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}Y'X_2(X_2'X_2)^{-1}X_2'y\right] =$

$\frac{T}{T-k_2}\left[\frac{y'y}{T} - \frac{y'X_2}{\sqrt{T}}(\frac{X_2'X_2}{T})^{-1}\frac{X_2'Y}{T}[\frac{Y'X_2}{\sqrt{T}}(\frac{X_2'X_2}{T})^{-1}\frac{X_2'Y}{\sqrt{T}}]^{-1}\frac{Y'X_2}{T}(\frac{X_2'X_2}{T})^{-1}\frac{X_2'y}{\sqrt{T}}\right] \xrightarrow{\mathbb{P}}$

$\varphi_0 = \beta_0'(\sigma_{11}\rho\rho' + \sigma_\varepsilon\sigma_\varepsilon')\beta_0 + \sigma_{11}\beta_0'\rho + \sigma_{11}\rho'\beta_0 + \sigma_{11}$, because $\frac{X_2'Y}{T} \xrightarrow{\mathbb{P}} 0$

$(y - Y\beta_0)'P_{X_2}(y - Y\beta_0) = (\frac{1}{\sqrt{T}}u'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{\sqrt{T}}X_2'u)$

$(y - Y\beta_0)'P_{X_2}(y - Y\beta_0) \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)$

$\frac{1}{T}(y - Y\beta_0)'M_{X_2}(y - Y\beta_0) = \frac{1}{T}u'u - (\frac{1}{T}u'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'u) \xrightarrow{\mathbb{P}} \sigma_{11}$

$(y - Y\beta_0)'P_{\tilde{Y}(\beta_0)}(y - Y\beta_0) = u'\tilde{Y}(\beta_0)[\tilde{Y}(\beta_0)'\tilde{Y}(\beta_0)]^{-1}\tilde{Y}(\beta_0)'u$

Since,
$\tilde{\Pi}(\beta_0) = (X_2'X_2)^{-1}X_2'Y - (X_2'X_2)^{-1}X_2'u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}$
and $S_{uu}(\beta_0) = \frac{1}{T-k_2}(y - Y\beta_0)'M_{X_2}(y - Y\beta_0) \xrightarrow{\mathbb{P}} \sigma_{11}$,

$S_{uv}(\beta_0) = \frac{1}{T-k_2}(y - Y\beta_0)'M_{X_2}Y$
$= \frac{T}{T-k_2}[\frac{1}{T}u'Y - (\frac{1}{T}u'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'Y)] \xrightarrow{\mathbb{P}} \sigma_{11}\rho'$

then,
$\tilde{Y}(\beta_0)'\tilde{Y}(\beta_0) = (\frac{X_2'Y}{\sqrt{T}} - \frac{X_2'u}{\sqrt{T}}\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)})'(\frac{X_2'X_2}{T})^{-1}(\frac{X_2'Y}{\sqrt{T}} - \frac{X_2'u}{\sqrt{T}}\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}) \overset{a}{\sim}$

$\Gamma_0'Q_0\Gamma_0$, where $\Gamma_0 = \Pi_0 + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon'$

$u'\tilde{Y}(\beta_0) = (\frac{u'X_2}{\sqrt{T}})(\frac{X_2'X_2}{T})^{-1}(\frac{X_2'Y}{\sqrt{T}} - \frac{X_2'u}{\sqrt{T}}\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}) \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'\Gamma_0$

So,
$(y - Y\beta_0)'P_{\tilde{Y}(\beta_0)}(y - Y\beta_0) \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'(\Pi_0 + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon')\times$
$[(\Pi_0 + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon')'Q_0(\Pi_0 + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon']^{-1}(\Pi_0 + Q_0^{-1}\xi_0'W_2\sigma_\varepsilon')'(\xi_0'W_1 + \sigma_{11}\tau_0)$

$= (\xi_0'W_1 + \sigma_{11}\tau_0)'\Gamma_0(\Gamma_0'Q_0\Gamma_0)^{-1}\Gamma_0'(\xi_0'W_1 + \sigma_{11}\tau_0)$

$\hat{\mathcal{S}} = (X_2'X_2)^{-1/2}X_2'\underline{Y}b_0(b_0'\hat{\Delta}b_0)^{-1/2} = (X_2'X_2)^{-1/2}X_2'u(S_{uu}(\beta_0))^{-1/2}$, then
$\hat{\mathcal{S}}'\hat{\mathcal{S}} = \frac{u'X_2(X_2'X_2)^{-1}X_2'u}{S_{uu}(\beta_0)} \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)$

$\hat{\mathcal{J}} = (X_2'X_2)^{-1/2}X_2'\underline{Y}\hat{\Delta}^{-1}A_0(A_0'\hat{\Delta}^{-1}A_0)^{-1/2}$,

$\hat{\mathcal{J}}'\hat{\mathcal{J}} = (A_0'\hat{\Delta}^{-1}A_0)^{-1/2}A_0'\hat{\Delta}^{-1}\underline{Y}'X_2(X_2'X_2)^{-1}X_2'\underline{Y}\hat{\Delta}^{-1}A_0(A_0'\hat{\Delta}^{-1}A_0)^{-1/2}$ $\frac{1}{\sqrt{T}}X_2'\underline{Y} =$
$[\frac{1}{\sqrt{T}}X_2'y, \frac{1}{\sqrt{T}}X_2'Y]$

$\overset{a}{\sim} \bar{Q}_0 = \begin{bmatrix} Q_0\Psi_0\beta_0 + \xi_0'W_1 + \sigma_{11}\tau_0 & Q_0\Psi_0 \end{bmatrix}$

$\hat{\Delta} = \underline{Y}'M_{X_2}\underline{Y}/(T - k_2) = \frac{1}{T-k_2}\begin{bmatrix} y'M_{X_2}y & y'M_{X_2}Y \\ Y'M_{X_2}y & Y'M_{X_2}Y \end{bmatrix}$

$\frac{1}{T-k_2}y'M_{X_2}y = \frac{T}{T-k_2}[\frac{1}{T}y'y - (\frac{1}{T}y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'y)]$

$\overset{\mathbb{P}}{\to} = \varphi_0$

$\frac{1}{T-k_2}y'M_{X_2}Y = \frac{T}{T-k_2}[\frac{1}{T}(\beta_0'Y'Y + u'Y) - (\frac{1}{T}y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'Y)]$
$\overset{\mathbb{P}}{\to} \phi_0' = \beta_0'\chi_0 + \sigma_{11}\rho'$

$\frac{1}{T-k_2}Y'M_{X_2}Y = \frac{T}{T-k_2}[\frac{1}{T}(Y'Y) - (\frac{1}{T}Y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'Y)]$

$\overset{\mathbb{P}}{\to} \chi_0 = \sigma_{11}\rho\rho' + \sigma_\varepsilon\sigma_\varepsilon'$

So, we have
$\hat{\Delta} \overset{\mathbb{P}}{\to} \Delta_0 = \begin{bmatrix} \varphi_0 & \phi_0' \\ \phi_0 & \chi_0 \end{bmatrix}$

Then we have
$\hat{\mathcal{J}}'\hat{\mathcal{J}} \overset{a}{\sim} \Sigma_0 = (A_0'\Delta_0^{-1}A_0)^{-1/2}A_0'\Delta_0^{-1}\bar{Q}_0'Q_0^{-1}\bar{Q}_0\Delta_0^{-1}A_0(A_0'\Delta_0^{-1}A_0)^{-1/2}$

$\hat{\mathcal{S}}'\hat{\mathcal{J}} = (b_0'\hat{\Delta}b_0)^{-1/2}b_0'\underline{Y}'X_2(X_2'X_2)^{-1}X_2'\underline{Y}\hat{\Delta}^{-1}A_0(A_0'\hat{\Delta}^{-1}A_0)^{-1/2}$

$\hat{\mathcal{S}}'\hat{\mathcal{J}} \overset{a}{\sim} \Theta_0' = (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}\bar{Q}_0\Delta_0^{-1}A_0(A_0'\Delta_0^{-1}A_0)^{-1/2}$, where $A_0 = [\beta_0, I_G]'$

$(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}}) = \begin{bmatrix} \hat{\mathcal{S}}'\hat{\mathcal{S}} & \hat{\mathcal{S}}'\hat{\mathcal{J}} \\ \hat{\mathcal{J}}'\hat{\mathcal{S}} & \hat{\mathcal{J}}'\hat{\mathcal{J}} \end{bmatrix}$,
So,
$(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}}) \overset{a}{\sim} \bar{\Lambda}_0$, where
$\bar{\Lambda}_0 = \begin{bmatrix} (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0) & \Theta_0' \\ \Theta_0 & \Sigma_0 \end{bmatrix}$

(ii) if $\Pi_2 = \Pi_0$ and $\tau = \tau_0/\sqrt{T}$
where $\Pi_0 \neq 0$ and τ_0 are respectively a $k_2 \times G$ constant matrix and a $k_2 \times 1$ constant vector as $T \to \infty$, then

$X_2'X_2 = \xi'X_0'X_0\xi + \xi'X_0'u\tau' + \tau u'X_0\xi + \tau u'u\tau'$

Since X_0 is exogenous, $u \overset{i.i.d}{\sim} N(0, \sigma_{11}I_T)$, then

$\frac{1}{T}X_2'X_2 = \frac{1}{T}\xi'X_0'X_0\xi + \frac{1}{T^{3/2}}\xi'X_0'u\tau_0' + \frac{1}{T^{3/2}}\tau_0 u'X_0\xi + \frac{1}{T^2}\tau_0 u'u\tau_0' \overset{\mathbb{P}}{\to} \xi_0'Q_{X_0X_0}\xi_0 = Q_0$

$\frac{1}{\sqrt{T}}X_2'u = \xi'\frac{X_0'u}{\sqrt{T}} + \tau_0\frac{u'u}{T} \overset{a}{\sim} \xi_0'W_1 + \sigma_{11}\tau_0$

$\frac{1}{T}X_2'u = \xi'\frac{X_0'u}{T} + \tau_0\frac{u'u}{T^{3/2}} \overset{\mathbb{P}}{\to} 0$

$\frac{1}{\sqrt{T}}X_2'\varepsilon = \xi'\frac{X_0'\varepsilon}{\sqrt{T}} + \tau_0\frac{u'\varepsilon}{T} \overset{a}{\sim} \xi_0'W_2$

$\frac{1}{T}X_2'\varepsilon = \xi'\frac{X_0'\varepsilon}{T} + \tau_0\frac{u'\varepsilon}{T^{3/2}} \overset{\mathbb{P}}{\to} 0$

$\frac{1}{T}X_2'Y = \frac{X_2'X_2}{T}\Pi_0 + \frac{X_2'u}{T}\rho' + \frac{X_2'\varepsilon}{T}\sigma_\varepsilon' \overset{\mathbb{P}}{\to} Q_0\Pi_0$

$\frac{1}{T}X_2'y = \frac{1}{T}X_2'Y\beta_0 + \frac{1}{T}X_2'u \overset{\mathbb{P}}{\to} Q_0\Pi_0\beta_0$

$\frac{1}{T}v'v = \rho\frac{1}{T}u'u\rho' + \rho\frac{1}{T}u'\varepsilon\sigma_\varepsilon' + \frac{1}{T}\sigma_\varepsilon\varepsilon'u\rho' + \sigma_\varepsilon\frac{1}{T}\varepsilon'\varepsilon\sigma_\varepsilon' \overset{\mathbb{P}}{\to} \chi_0 = \sigma_{11}\rho\rho' + \sigma_\varepsilon\sigma_\varepsilon'$

$\frac{1}{T}Y'Y = \Pi_0'\frac{X_2'X_2}{T}\Pi_0 + \Pi_0'\frac{X_2'v}{T} + \frac{v'X_2}{T}\Pi_0 + \frac{v'v}{T} \overset{\mathbb{P}}{\to} \Pi_0'Q_0\Pi_0 + \chi_0$

$\frac{1}{T}Y'u = \Pi_0' \frac{1}{T} X_2' u + \rho \frac{1}{T} u'u + \sigma_\varepsilon \frac{1}{T} \varepsilon' u \xrightarrow{\mathbb{P}} \sigma_{11}\rho$

So,
$\frac{1}{T}y'y = \beta_0' \frac{1}{T} Y'Y\beta_0 + \beta_0' \frac{1}{T} Y'u + \frac{1}{T} u'Y\beta_0 + \frac{1}{T} u'u \xrightarrow{\mathbb{P}} \bar{\varphi}_0$, where
$\bar{\varphi}_0 = \beta_0'(Q_0\Pi_0 + \chi_0)\beta_0 + \sigma_{11}\beta_0'\rho + \sigma_{11}\rho'\beta_0 + \sigma_{11}$

$\frac{Y'X_2(X_2'X_2)^{-1}X_2'Y}{T} = (\frac{1}{T}Y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'Y) \xrightarrow{\mathbb{P}} \Pi_0' Q_0 \Pi_0$

$\frac{Y'X_2(X_2'X_2)^{-1}X_2'y}{T} = (\frac{1}{T}Y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'y) \xrightarrow{\mathbb{P}} \Pi_0' Q_0 \Pi_0 \beta_0$

$\frac{Y'X_2(X_2'X_2)^{-1}X_2'u}{\sqrt{T}} = (\frac{1}{T}Y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{\sqrt{T}}X_2'u) \overset{a}{\sim} \Pi_0'(\xi_0' W_1 + \sigma_{11}\tau_0)$

So,
$\sqrt{T}(\hat{\beta}_{2SLS} - \beta_0) = [\frac{Y'X_2(X_2'X_2)^{-1}X_2'Y}{T}]^{-1} \frac{Y'X_2(X_2'X_2)^{-1}X_2'u}{\sqrt{T}} \overset{a}{\sim}$
$(\Pi_0' Q_0 \Pi_0)^{-1} \Pi_0' (\xi_0' W_1 + \sigma_{11}\tau_0)$

$\frac{1}{T-k_2}\left[y'y - y'X_2(X_2'X_2)^{-1}X_2'Y[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}Y'X_2(X_2'X_2)^{-1}X_2'y\right] =$

$\frac{T}{T-k_2}\left[\frac{y'y}{T} - \frac{y'X_2}{T}(\frac{X_2'X_2}{T})^{-1}\frac{X_2'Y}{T}[\frac{Y'X_2}{T}(\frac{X_2'X_2}{T})^{-1}\frac{X_2'Y}{T}]^{-1}\frac{Y'X_2}{T}(\frac{X_2'X_2}{T})^{-1}\frac{X_2'y}{T}\right] \xrightarrow{\mathbb{P}}$

$\bar{\varphi}_0 - \beta_0'\Pi_0' Q_0 Q_0^{-1} Q_0 \Pi_0 (\Pi_0' Q_0 Q_0^{-1} Q_0 \Pi_0)^{-1} \Pi_0' Q_0 Q_0^{-1} Q_0 \Pi_0 \beta_0 =$

$\bar{\varphi}_0 - \beta_0'\Pi_0' Q_0 \Pi_0 (\Pi_0' Q_0 \Pi_0)^{-1} \Pi_0' Q_0 \Pi_0 \beta_0$

$(y - Y\beta_0)' P_{X_2} (y - Y\beta_0) = (\frac{1}{\sqrt{T}}u'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{\sqrt{T}}X_2'u)$

$(y - Y\beta_0)' P_{X_2} (y - Y\beta_0) \overset{a}{\sim} (\xi_0' W_1 + \sigma_{11}\tau_0)' Q_0^{-1} (\xi_0' W_1 + \sigma_{11}\tau_0)$

$\frac{1}{T}(y - Y\beta_0)' M_{X_2} (y - Y\beta_0) = \frac{1}{T}u'u - (\frac{1}{T}u'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'u) \xrightarrow{\mathbb{P}} \sigma_{11}$

$(y - Y\beta_0)' P_{\tilde{Y}(\beta_0)} (y - Y\beta_0) = u'\tilde{Y}(\beta_0)[\tilde{Y}(\beta_0)'\tilde{Y}(\beta_0)]^{-1}\tilde{Y}(\beta_0)'u$

Since, $\tilde{\Pi}(\beta_0) = (X_2'X_2)^{-1}X_2'Y - (X_2'X_2)^{-1}X_2'u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}$

and $S_{uu}(\beta_0) = \frac{1}{T-k_2}(y - Y\beta_0)' M_{X_2} (y - Y\beta_0) \xrightarrow{\mathbb{P}} \sigma_{11}$,

$S_{uv}(\beta_0) = \frac{1}{T-k_2}(y - Y\beta_0)'M_{X_2}Y$
$= \frac{T}{T-k_2}[\frac{1}{T}u'Y - (\frac{1}{T}u'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'Y)] \xrightarrow{\mathbb{P}} \sigma_{11}\rho'$

then,
$\tilde{Y}(\beta_0)'\tilde{Y}(\beta_0) = (Y - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)})'X_2(X_2'X_2)^{-1}X_2'(Y - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)})$

$\frac{1}{\sqrt{T}}X_2'(Y - X_2\Pi_0 - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}) = \frac{1}{\sqrt{T}}X_2'(v - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}) \overset{a}{\sim} \xi_0'W_2\sigma_\varepsilon'$

So,
$\frac{1}{\sqrt{T}}\tilde{Y}(\beta_0)'u =$
$\frac{1}{\sqrt{T}}\{Y - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\}'X_2(X_2'X_2)^{-1}X_2'u = \frac{1}{\sqrt{T}}\{Y - X_2\Pi_0 - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\}'X_2(\frac{X_2'X_2}{T})^{-1}\frac{X_2'u}{T} +$
$\Pi_0'\frac{1}{\sqrt{T}}X_2'u \overset{a}{\sim} \Pi_0'(\xi_0'W_1 + \sigma_{11}\tau_0)$, because $\frac{X_2'u}{T} \to 0$

Similarly, we have

$\frac{1}{T}(\tilde{Y}(\beta_0)'\tilde{Y}(\beta_0)) = \frac{1}{T}\{Y - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\}'X_2(X_2'X_2)^{-1}X_2'\{Y - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\} =$

$\frac{1}{T}\{Y - X_2\Pi_0 + X_2\Pi_0 - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\}'X_2(X_2'X_2)^{-1}X_2'\{Y - X_2\Pi_0 + X_2\Pi_0 - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\} =$
$\frac{1}{T}\{v - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\}'X_2(X_2'X_2)^{-1}X_2'\{v - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\} + \frac{1}{T}\{v - u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\}'X_2\Pi_0 + \frac{1}{T}\Pi_0'X_2'\{v -$
$u\frac{S_{uv}(\beta_0)}{S_{uu}(\beta_0)}\} + \Pi_0'(\frac{X_2'X_2}{T})\Pi_0 \xrightarrow{\mathbb{P}} \Pi_0'Q_0\Pi_0$

and

$(y - Y\beta_0)'P_{\tilde{Y}(\beta_0)}(y - Y\beta_0) \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'\Pi_0(\Pi_0'Q_0\Pi_0)^{-1}\Pi_0'(\xi_0'W_1 + \sigma_{11}\tau_0)$

From (i),

$\hat{\mathcal{S}}'\hat{\mathcal{S}} = \frac{u'X_2(X_2'X_2)^{-1}X_2'u}{S_{uu}(\beta_0)} \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)$,
and

$\frac{\hat{\mathcal{S}}'\hat{\mathcal{S}}}{T} = (\frac{u'X_2}{T})(\frac{X_2'X_2}{T})^{-1}(\frac{X_2'u}{T})[S_{uu}(\beta_0)]^{-1} \xrightarrow{\mathbb{P}} 0$

$\hat{\mathcal{J}}'\hat{\mathcal{J}} = (A_0'\hat{\Delta}^{-1}A_0)^{-1/2}A_0'\hat{\Delta}^{-1}\underline{Y}'X_2(X_2'X_2)^{-1}X_2'\underline{Y}\hat{\Delta}^{-1}A_0(A_0'\hat{\Delta}^{-1}A_0)^{-1/2}$ $\frac{1}{T}X_2'\underline{Y} =$
$[\frac{1}{T}X_2'y, \frac{1}{T}X_2'Y] \xrightarrow{\mathbb{P}} \widehat{Q}_0$, where

$\widehat{Q}_0 = \begin{bmatrix} Q_0\Pi_0\beta_0 & Q_0\Pi_0 \end{bmatrix}$

$$\hat{\Delta} = \underline{Y}' M_{X_2} \underline{Y}/(T-k_2) = \frac{1}{T-k_2} \begin{bmatrix} y'M_{X_2}y & y'M_{X_2}Y \\ Y'M_{X_2}y & Y'M_{X_2}Y \end{bmatrix}$$

$$\frac{1}{T-k_2} y'M_{X_2}y = \frac{T}{T-k_2}[\frac{1}{T}y'y - (\frac{1}{T}y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'y)]$$

$$\xrightarrow{\mathbb{P}} \bar{\varphi}_0 - \beta_0'\Pi_0'Q_0\Pi_0\beta_0$$

$$\frac{1}{T-k_2} y'M_{X_2}Y = \frac{T}{T-k_2}[\frac{1}{T}(\beta_0'Y'Y + u'Y) - (\frac{1}{T}y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'Y)]$$

$\xrightarrow{\mathbb{P}} \bar{\phi}_0'$, such that $\bar{\phi}_0' = \beta_0'(\Pi_0'Q_0\Pi_0 + \chi_0) + \sigma_{11}\rho' - \beta_0'\Pi_0'Q_0\Pi_0$

$$\frac{1}{T-k_2} Y'M_{X_2}Y = \frac{T}{T-k_2}[\frac{1}{T}(Y'Y) - (\frac{1}{T}Y'X_2)(\frac{1}{T}X_2'X_2)^{-1}(\frac{1}{T}X_2'Y)]$$

$$\xrightarrow{\mathbb{P}} \bar{\Psi}_0 = \beta_0'(\Pi_0'Q_0\Pi_0 + \chi_0) - \Pi_0'Q_0\Pi_0$$

So, we have $\hat{\Delta} \xrightarrow{\mathbb{P}} \widehat{\Delta}_0 = \begin{bmatrix} \bar{\varphi}_0 - \beta_0'\Pi_0'Q_0\Pi_0\beta_0 & \bar{\phi}_0' \\ \bar{\phi}_0 & \bar{\Psi}_0 \end{bmatrix}$

Then we have

$$\frac{\hat{\mathcal{J}}'\hat{\mathcal{J}}}{T} \xrightarrow{\mathbb{P}} \widetilde{\Sigma}_0 = (A_0'\widehat{\Delta}_0^{-1}A_0)^{-1/2}A_0'\widehat{\Delta}_0^{-1}\widehat{Q}_0'Q_0^{-1}\widehat{Q}_0\widehat{\Delta}_0^{-1}A_0(A_0'\widehat{\Delta}_0^{-1}A_0)^{-1/2}$$

$$\hat{\mathcal{S}}'\hat{\mathcal{J}} = (b_0'\hat{\Delta}^{-1}b_0)^{-1/2}b_0'\underline{Y}'X_2(X_2'X_2)^{-1}X_2'\underline{Y}\hat{\Delta}^{-1}A_0(A_0'\hat{\Delta}^{-1}A_0)^{-1/2}$$

$\frac{\hat{\mathcal{S}}'\hat{\mathcal{J}}}{T} \xrightarrow{\mathbb{P}} \widehat{\Theta}_0' = (\xi_0'W_1 + \sigma_{11}\tau_0)')'Q_0^{-1}\widehat{Q}_0\widehat{\Delta}_0^{-1}A_0(A_0'\widehat{\Delta}_0^{-1}A_0)^{-1/2}$, where $A_0 = [\beta_0, I_G]'$

$$\frac{1}{T}(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}}) = \begin{bmatrix} \frac{\hat{\mathcal{S}}'\hat{\mathcal{S}}}{T} & \frac{\hat{\mathcal{S}}'\hat{\mathcal{J}}}{T} \\ \frac{\hat{\mathcal{J}}'\hat{\mathcal{S}}}{T} & \frac{\hat{\mathcal{J}}'\hat{\mathcal{J}}}{T} \end{bmatrix},$$

So, $\frac{1}{T}(\hat{\mathcal{S}}, \hat{\mathcal{J}})'(\hat{\mathcal{S}}, \hat{\mathcal{J}}) \xrightarrow{\mathbb{P}} \sim\widehat{\Lambda}_0$, where

$$\widehat{\Lambda}_0 = \begin{bmatrix} 0 & \widehat{\Theta}_0' \\ \widehat{\Theta}_0 & \widetilde{\Sigma}_0 \end{bmatrix}$$

B Appendix

Proof of theorem 5.1

Let Proposition 5.1 hold. We know that, under the null hypothesis H_0,

$$T_{2SLS}(\beta_{0j}) = \frac{\hat{\beta}_j - \beta_{0j}}{\hat{\sigma}} \tag{B.1}$$

where $\hat{\beta}_j - \beta_{0j}$, $j = 1, ..., G$ is the the jth element of the vector

$[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1/2}Y'X_2(X_2'X_2)^{-1}X_2'u$ and

$\hat{\sigma}^2 = \frac{1}{T-k_2}\left[y'y - y'X_2(X_2'X_2)^{-1}X_2'Y[Y'X_2(X_2'X_2)^{-1}X_2'Y]^{-1}Y'X_2(X_2'X_2)^{-1}X_2'y\right]$

$$AR(\beta_0) = \frac{T-k_2}{k_2}\frac{(y-Y\beta_0)'P_{X_2}(y-Y\beta_0)}{(y-Y\beta_0)'M_{X_2}(y-Y\beta_0)} \tag{B.2}$$

$$K(\beta_0) = \frac{(y-Y\beta_0)'P_{\tilde{Y}(\beta_0)}(y-Y\beta_0)}{\frac{1}{T-k_2}(y-Y\beta_0)'M_{X_2}(y-Y\beta_0)} \tag{B.3}$$

$$CLR(\beta_0) = \hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\lambda}^{min} \tag{B.4}$$

(i) if $\Pi_2 = \Pi_0/\sqrt{T}$ and $\tau = \tau_0/\sqrt{T}$
where Π_0, τ_0 are respectively a $k_2 \times G$ constant matrix and a $k_2 \times 1$ constant vector as $T \to \infty$, then
from Proposition 5.1, it easy to see that
$T_{2SLS}(\beta_{0j}) \overset{a}{\sim} \varphi_0^{-1/2}\bar{\Upsilon}_j$
where $\bar{\Upsilon}_j$ is such that $(\Psi_0'Q_{X_0}\Psi_0)^{-1/2}\Psi_0'(\xi_0'W_1 + \sigma_{11}\tau_0) = [\bar{\Upsilon}_j]_{1\leq j\leq G}$
Similarly, we have

$AR(\beta_0) \overset{a}{\sim} \frac{1}{k_2}(\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0)$

$K(\beta_0) \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'\Gamma_0(\Gamma_0'Q_0\Gamma_0)^{-1}\Gamma_0'(\xi_0'W_1 + \sigma_{11}\tau_0)$

$CLR(\beta_0) \overset{a}{\sim} (\xi_0'W_1 + \sigma_{11}\tau_0)'Q_0^{-1}(\xi_0'W_1 + \sigma_{11}\tau_0) - \hat{\lambda}_0^{min}$

where
φ_0, Ψ_0, Γ_0, $\bar{\Lambda}_0$ are defined in Proposition 5.1 -(i); $\hat{\lambda}_0^{min}$ is the smallest eigenvalue of $\bar{\Lambda}_0$; W_1, W_2 are respectively i.i.d $N(0, \sigma_{11}Q_{X_0X_0})$ and $N(0, Q_{X_0X_0})$.
Since, $(\xi_0'W_1+\sigma_{11}\tau_0) \overset{a}{\sim} N(\sigma_{11}\tau_0, \sigma_{11}Q_0)$ and $\Gamma_0'(\xi_0'W_1+\sigma_{11}\tau_0)|_{W_2} \overset{a}{\sim} N(\sigma_{11}\Gamma_0'\tau_0, \sigma_{11}\Gamma_0'Q_0\Gamma_0)$,
hence, $\frac{1}{\sqrt{\sigma_{11}}}Q_0^{-1/2}(\xi_0'W_1 + \sigma_{11}\tau_0) \overset{a}{\sim} N(\sqrt{\sigma_{11}}Q_0^{-1/2}\tau_0, I_{k_2})$
and $\frac{1}{\sqrt{\sigma_{11}}}(\Gamma_0'Q_0\Gamma_0)^{-1/2}\Gamma_0'(\xi_0'W_1 + \sigma_{11}\tau_0)|_{W_2} \overset{a}{\sim} N(\sqrt{\sigma_{11}}(\Gamma_0'Q_0\Gamma_0)^{-1/2}\Gamma_0'\tau_0, I_G)$.
So, we then have $AR(\beta_0) \overset{a}{\sim} \frac{1}{k_2}\chi^2(k_2, \sigma_{11}\tau_0'Q_0^{-1}\tau_0)$ and $K(\beta_0)|_{W_2} \overset{a}{\sim} \chi^2(G, \tau_0'\Gamma_0(\Gamma_0'Q_0\Gamma_0)^{-1}\Gamma_0'\tau_0)$.
It is clear that the conditional distribution of $K(\beta_0)$ given W_2 depends on W_2, so, the conditional distribution is different to the unconditional distribution.

(ii) Similarly, if $\Pi_2 = \Pi_0$ and $\tau = \tau_0/\sqrt{T}$
where $\Pi_0 \neq 0$ and τ_0 are respectively a $k_2 \times G$ constant matrix and a $k_2 \times 1$ constant vector as $T \to \infty$, then

$$T_{2SLS}(\beta_{0j}) \overset{a}{\sim} \bar{\varphi}_0^{-1/2}\widetilde{\Upsilon}_j$$

$$AR(\beta_0) \overset{a}{\sim} \frac{1}{k_2}\chi^2(k_2, \sigma_{11}\tau_0' Q_0^{-1}\tau_0)$$

$$K(\beta_0) \overset{a}{\sim} \chi^2(G, \sigma_{11}\tau_0'\Pi_0(\Pi_0' Q_0 \Pi_0)^{-1}\Pi_0'\tau_0)$$

$$CLR(\beta_0) \overset{a}{\sim} \frac{1}{k_2}\chi^2(k_2, \sigma_{11}\tau_0' Q_0^{-1}\tau_0) - \widetilde{\lambda}_0^{min}$$

where $\widetilde{\Upsilon}_j$, $\bar{\varphi}_0$, $\widehat{\Lambda}_0$ are defined in Proposition 5.1 -(ii); $\tilde{\lambda}_0^{min}$ is the smallest eigenvalue of $\widehat{\Lambda}_0$, W_1, W_2 are respectively i.i.d $N(0, \sigma_{11}Q_{X_0X_0})$ and $N(0, Q_{X_0X_0})$.
Here, the conditional distribution of $K(\beta_0)$ given W_2 do not depend on W_2, so, it is equal to the unconditional distribution.

Proof of corollary 5.1
Assume that Theorem 5.1 hold under H_0 and suppose $G = k_2 = 1$.
Since $k_2 = 1$, hence $P_{\tilde{Y}(\beta_0)} = P_{X_2} = 1$ and from (5.6)-(5.7) $AR(\beta_0) = K(\beta_0)$. Because also $G = 1$, by (5.9),
$CLR(\beta_0) = \frac{1}{2}(\hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\mathcal{J}}'\hat{\mathcal{J}} + \sqrt{(\hat{\mathcal{S}}'\hat{\mathcal{S}} - \hat{\mathcal{J}}'\hat{\mathcal{J}})^2 - 4[\hat{\mathcal{S}}'\hat{\mathcal{S}}\hat{\mathcal{J}}'\hat{\mathcal{J}} - (\hat{\mathcal{S}}'\hat{\mathcal{J}})^2]})$.
When $K_2 = 1$, $\hat{\mathcal{S}}$ and $\hat{\mathcal{J}}$ are scalars, and the $CLR(\beta_0)$ statistic collapses to the statistic $\hat{\mathcal{S}}'\hat{\mathcal{S}} = AR(\beta_0)$. So, we have $AR(\beta_0) = K(\beta_0) = CLR(\beta_0)$. And we know by Theorem 5.1 that $AR(\beta_0) \overset{a}{\sim} \chi^2(1, \frac{\sigma_{11}\tau_0^2}{Q_0})$ because for $k_2 = 1$, Q_0 and τ_0 are scalars.

Proof of corollary 5.2
Assume that Theorem 5.1 hold and set $\tau_0 = 0$, then we have $AR(\beta_0) \overset{a}{\sim} \frac{1}{k_2}\chi^2(k_2)$ and $K(\beta_0) \overset{a}{\sim} \chi^2(G)$ independently of the conditions (i) and (ii) of Theorem 5.1.

References

ANDERSON, T. W. (2003). *An Intrododuction to Multivariate Statistical Analysis*. John Wiley.

ANDREW, D. and STOCK, J. (2005). Inference with weak instruments. Invited Paper for the 2005 World Congress of the Econometric Society, London. Cowles Foundation Discussion Paper N0. 1530.

BEKKER, P. A. and KLEIBERGEN, F. (2003). Finite-sample instrumental variables inference using an asymptotically pivotal statistic. *Econometric Theory*, 19:744–753.

BLUNDELL, R. and HOROWITZ, J. (2005). A nonparametric test of exogeneity. *Working Paper.*

BOUND, J., JAEGER, D. A., and BAKER, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of American Statistical Association*, 90:443–450.

DUFOUR, J.-M. (1987). Linear wald methods for inference on covariances and weak exogeneity tests in structural equations. *Time Series and Econometric Modelling*, pages 317–338.

DUFOUR, J.-M. (1997). Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica*, 65:1365–1388.

DUFOUR, J.-M. (2003). Identification, weak instruments and statistical inference in econometrics. *anadian Journal of Economics*, 36:767–808.

DUFOUR, J.-M. and JASIAK, J. (2001). Finite sample limited information inference methods for structural equations and models with generated regressors. *International Economic Review*, 42:815–843.

DUFOUR, J.-M. and TAAMOUTI, M. (2006). Further results on projection-based inference in iv regression with weak, collinear or missing instruments. *Journal of Econometrics.*

ENGLE, R. F. (1982). A general approach to lagrange multiplier diagnostics. *Journal of Econometrics*, 20:83–104.

FAREBROTHER, R. W. (1976). A remark on the wu test. *Econometrica*, 44:475–477.

HAUSMAN, J. (1978). Specification tests in econometrics. *Econometrica*, 46:1251–1272.

HOLLY, A. (1982). A remark on hausman's test. *Econometrica*, 50:749–760.

HOLLY, A. and SARGAN, D. Testing of exogeneity within a linited information framework. Cahier 8204, Université de Lausanne.

HWANG, H.-S. (1980). Test of independence between a subset of stochastic regressors and disturbances. *International Economic Review*, 21:749–760.

KARIYA, T. and HODOSHIMA, H. (1980). Finite-sample properties of the tests for independence in structural systems and LRT. *The Econometric Studies Quaterly*, 31:45–56.

KIVET, J. F. and NIEMCZYK, J. (2006). On the limiting and empirical distribution of iv estimators when some of the instruments are invalid. *Working Paper*.

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

MOREIRA, J. M. (2001). Test with correct size when instruments can be arbitrairy weak. *Working Paper*, 37. Center for Labor Economics Working Paper Series, UC Berkeley.

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

NAKAMURA, A. and NAKAMURA, M. (1981). On the relationships among several specification error tests presented by DURBIN, WU and HAUSMAN. *Econometrica*, 49:1583–1588.

NELSON, C. R. and STARTZ, R. (1990). Some further results on th exact small sample properties of the intrumental variable estimator. *Econometrica*, 58:967–976.

REVANKAR, N. S. (1978). Asymptotic relative efficiency analysis of certain tests in structural sysytems. *International Economic Review*, 19:165–179.

REVANKAR, N. S. and 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. and STOCK, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65:557–586.

STOCK, J. H. and WRIGHT, J. H. (2000). GMM with weak identification. *Econometrica*, 68:1097–1126.

WANG, J. and ZIVOT, E. (1998). Inference on a structural parameter in intrumental variables regression with weak instruments. *Econometrica*, 66:1389–1404.

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 smaple results. *Econometrica*, 42:529–546.

ZIVOT, E., STARTZ, R., and NELSON, C. R. (1998). Valid confidence intervals and inference in the presence of weak instruments. *International Economic Review*, 39:1119–1144.