SCHOOL OF ECONOMICS AND FINANCE

Discussion Paper 2011-03

Subset Hypotheses Testing and Instrument Exclusion in the Linear IV Regression

Firmin Doko Tchatoka

Subset hypotheses testing and instrument exclusion in the linear IV regression [*]

Firmin Doko Tchatoka[†]

School of Economics and Finance

University of Tasmania

February 7, 2011

ABSTRACT

This paper investigates the asymptotic size properties of robust subset tests when instruments are left out of the analysis. Recently, robust subset procedures have been developed for testing hypotheses which are specified on the subsets of the structural parameters or on the parameters associated with the included exogenous variables. It has been shown that they never over-reject the true parameter values even when nuisance parameters are not identified. However, their robustness to instrument exclusion has not been investigated. Instrument exclusion is an important problem in econometrics and there are at least two reasons to be concerned. Firstly, it is difficult in practice to assess whether an instrument has been omitted. For example, some components of the "identifying" instruments that are excluded from the structural equation may be quite uncertain or "left out" of the analysis. Secondly, in many instrumental variable (IV) applications, an infinite number of instruments are available for use in large sample estimation. This is particularly the case with most time series models. If a given variable, say X_t, is a legitimate instrument, so too are its lags X_{t-1}, X_{t-2}, ... Hence, instrument exclusion seems highly likely in most practical situations. After formulating a general asymptotic framework which allows one to study this issue in a convenient way, I consider two main setups: (1) the missing instruments are (possibly) relevant, and, (2) they are asymptotically weak. In both setups, I show that all subset procedures studied are in general consistent against instrument inclusion (hence asymptotically invalid for the subset hypothesis of interest). I characterize cases where consistency may not hold, but the asymptotic distribution is modified in a way that would lead to size distortions in large samples. I propose a "rule of thumb" which allows to practitioners to know whether a missing instrument is detrimental or not to subset procedures. I present a Monte Carlo experiment confirming that the subset procedures are unreliable when instruments are missing.

Key words: Robust subset tests; LIML estimator; missing instruments; weak identification; non robust; consistency; projection-based method.

JEL classification: C12; C13; C30.

1. Introduction

Subset hypotheses testing is likely to be prevalent in applied work. Many interesting hypotheses are specified on subsets of the structural parameters or on the parameters associated with the included exogenous variables. Several models where subset hypotheses are often specified include but are not restricted to: (1) forward-looking models, such as the new Keynesian Phillips curve [Mavroeidis (2004, 2005), Dufour, Khalaf and Kichian (2006), Kleibergen and Mavroeidis (2008)]; (2) stochastic discount factor models, in particular the linear factor model [Kocherlakota (1990), Hansen, Heaton and Yaron (1996), Kleibergen (2005, 2009)]; and (3) models of unemployment [Bean (1994), Malcomson and Mavroeidis (2006)].

The literature concerned with subset hypotheses testing falls globally into two categories. The first is the projection method based on identification-robust statistics [see Dufour (1997), Dufour and Jasiak (2001), Dufour and Taamouti (2005, 2007)]. This method consists of inverting robust statistics to build a confidence set for the full set of parameters, and then uses projection techniques to obtain a confidence set for the subset of parameters of interest. The projection method is robust to weak instruments and other statistical difficulties, such as instrument exclusion. However, it has often been criticized for: (1) being overly conservative and (2) having low power when too many instruments are used. Recently, Chaudhuri and Zivot (2010) have suggested a new projection procedure based on the K-statistic, namely EPK (efficient projection based on the K-statistic) which exhibits more power than the standard projection method. Nevertheless, the robustness of the EPK to instrument exclusion has not been investigated. The second category is the robust subset procedures proposed by Stock and Wright (2000) and recently developed by Kleibergen (2004, 2005, 2008), Startz, Nelson and Zivot (2006). These procedures, commonly called the conventional plug-in based tests, consist of replacing the nuisance parameters that are not specified by the hypothesis of interest by estimators. Several studies have shown that the plug-in based tests never over-reject the true parameter values even when nuisance parameters are not, or weakly, identified [see Startz et al. (2006) and Kleibergen (2008, 2009)].

However, the robustness of these tests to instrument exclusion has not been explored. Instrument exclusion is an important problem in econometrics and there are at least two reasons to be concerned. Firstly, it is difficult in practice to assess whether an instrument has been omitted. For example, in model (2.1)-(2.2), some components of the "identifying" instruments Z that are excluded from the structural equation (2.1) may be quite uncertain or "left out" of the analysis. Secondly, in many

instrumental variable (IV) applications, an infinite number of instruments are available for use in large sample estimation. This is particularly the case with most time series models. If a given variable, say X_t, is a legitimate instrument, so too are its lags X_{t-1}, X_{t-2}, ... Hence, instrument exclusion seems highly likely in most practical situations.

Although several authors have worked on proposing optimal instrumental variable estimations [see Hayashi and Christopher (1983), Hansen (1985), Hansen, Heaton and Masao (1988), West (2001)] or proposing optimal IV selection [Donald (1999), Donald and Newey (2001), Hall and Peixe (2003), Dufour and Taamouti (2003)], there are still issues related to inference. For example, use of the entire set of available instruments can result in large asymptotic efficiency gains relative to the use of a small set of instruments [see Hansen and Singleton (1991), West and Wilcox (1996)]. More recently, Doko and Dufour (2008) have shown that when identification is deficient or weak (weak instruments) and endogeneity moderate, the use of additional instruments (even invalid) improves the efficiency of partial IV estimators and pretest-estimators based on Generalized Wald-type exogeneity tests.

This paper has two main goals. The first is to explore the asymptotic size properties of robust subset tests when instruments are omitted. The second is to provide a tool which allows to practitioners to know whether or not a set of instruments omitted is detrimental to subset inference.

Dufour and Taamouti (2007) have investigated the robustness of the Anderson and Rubin (1949, AR), Kleibergen (2002, K), and Moreira (2003, CLR) tests to instrument exclusion. They found that the K-and CLR-tests are seriously size distorted when relevant instruments are omitted. However, the AR-test is typically robust to instrument exclusion. One of the difficulties with Dufour and Taamouti (2007) is that the null hypothesis is specified on the full set rather than a subset of structural parameters. It is natural to conjecture that with missing instruments, even the validity of the AR-test is not guaranteed when testing subset hypotheses. The reason is that all subset procedures studied here are plug-in based methods that require the available instruments to be valid (uncorrelated with the structural errors). However, when some instruments are omitted, they become part of the residuals (see Section 2) and the exogeneity assumption is generally violated. Hence, it is not clear how subset tests behave when subject to instrument exclusion.

This paper focuses on the linear IV model and investigates the asymptotic size of robust subset tests when instruments are missing. An extension to the GMM framework is studied in Doko (2010). Four subset tests are considered in this study: Anderson and Rubin (1949, AR-test), Kleibergen (2005, KLM-test), Moreira (2003, MQLR-test), and the J-statistic (JKLM) that tests the miss-

specification of the model. After formulating a general asymptotic framework which allows one to study this issue in a convenient way, I consider two main setups: (1) the missing instruments are (possibly) relevant, and (2) they are asymptotically weak, *i.e.* the parameter which controls their quality in the first step regression that identifies the nuisance parameters, converges to zero at rate $[T^{-\frac{1}{2}}]$ when the sample size T increases.

In both setups, I show that all subset procedures are in general consistent against instrument inclusion (hence asymptotically invalid for the subset hypothesis of interest). I characterize cases where consistency may not hold, but the asymptotic distribution is modified in a way that would lead to size distortions in large samples. I also provide conditions under which the asymptotic distribution of the statistics remains the same as in the case where any instrument is omitted (despite the exclusion of some of them). I present a Monte Carlo experiment which confirms the theoretical results. I find that when no instrument is missing, all the test procedures studied are valid whether identification is strong or weak. In particular, they are conservative when identification is deficient or weak (weak instruments) [similar to Kleibergen (2008, 2009) and Kleibergen and Mavroeidis (2009)]. However, all tests are seriously size distorted with maximal size distortion as great as 100 percent if instruments are missing, while the projection method remains valid (level is controlled).

Overall, my results underscore the importance of using the projection techniques [see Dufour (1997), Dufour and Jasiak (2001), Dufour and Taamouti (2005, 2007)] which do not exhibit such a problem.

The paper is organized as follows. Section 2 formulates the model. Section 3 briefly describes the robust subset statistics. Section 4 studies the asymptotic distribution of the statistics (under the null hypothesis) when potential relevant instruments are omitted. Section 5 deals with the case where the omitted instruments are weak. Section 6 presents the Monte Carlo experiment, and conclusions are drawn in Section 7. Proofs are presented in the Appendix.

2. Framework

Consider the following simplified simultaneous equations model:

$$y \;=\; Y_1\theta_1 + Y_2\theta_2 + \varepsilon\,, \tag{2.1}$$

$$(Y_1\,,\,Y_2) \;=\; Z(\Pi\,,\,\Gamma) + (V_1^*\,,\,V_2^*)\,, \tag{2.2}$$

where $y \in \mathbb{R}^T$ is a vector of observations on the dependent variable, $Y_1 \in \mathbb{R}^{T \times G_1}$ and $Y_2 \in \mathbb{R}^{T \times G_2}$ are matrices of explanatory (supposedly) endogenous variables, $Z \in \mathbb{R}^{T \times k_z}$ is a matrix of excluded exogenous variables, $\varepsilon \in \mathbb{R}^T$ is a vector of structural disturbances, $V_1^* \in \mathbb{R}^{T \times G_1}$ and $V_2^* \in \mathbb{R}^{T \times G_2}$ are matrices of reduced form disturbances. $\theta_1 \in \mathbb{R}^{G_1 \times 1}$ and $\theta_2 \in \mathbb{R}^{G_2 \times 1}$ are unknown structural coefficients, $\Pi \in \mathbb{R}^{k_z \times G_1}$ and $\Gamma \in \mathbb{R}^{k_z \times G_2}$ are matrices of unknown reduced form of coefficients. Let $\theta = (\theta_1', \, \theta_2')'$, and $G = G_1 + G_2$ with $G \geq 1$.

We are interested in making inference on the first component θ_1, treating the second component θ_2 as a nuisance parameter. More formally, we want to test the hypothesis

$$\mathrm{H}_0 : \theta_1 = \theta_{01} \,, \tag{2.3}$$

where $\theta_{01} \in \mathbb{R}^{G_1 \times 1}$ is fixed.

The projection method proposed by Dufour (1997), Dufour and Jasiak (2001), Dufour and Taamouti (2005, 2007), and Doko and Dufour (2006), provides a two-step weak instrument robust procedure for H_0. The first step consists of testing the joint hypothesis $\mathrm{H}_{\theta_0} : \theta = \theta_0$. Since the reduced form for model (2.1)-(2.2) can be written as:

$$y - Y\theta_0 \;\; = \;\; Z\pi + \varepsilon^* \,, \tag{2.4}$$

where $\pi = \Pi(\theta_1 - \theta_{01}) + \Gamma(\theta_2 - \theta_{02})$ and $\varepsilon^* = \varepsilon + V_1^*(\theta_1 - \theta_{01}) + V_2^*(\theta_2 - \theta_{02})$, H_{θ_0} can be assessed by testing $\mathrm{H}_\pi : \pi = 0$ from a F-type statistic in (2.4), namely AR-statistic [see Anderson and Rubin (1949)]. Hence, a confidence set for the true value θ_0 can be obtained by inverting this statistic. The second step consists of applying projection techniques to obtain a confidence set with correct coverage probability for θ_1, whether identification is deficient or weak (weak instruments). More interestingly, the confidence set for θ_1 obtained by projection is robust to instrument exclusion [see Dufour and Taamouti (2007)]. A common criticism is that the projection method is overly conservative and has low power when too many instruments are used. Robust subset procedures have then been suggested to assess subset hypotheses [see Stock and Wright (2000), Kleibergen (2004, 2005, 2008, 2009), Kleibergen and Mavroeidis (2008, 2009), and Startz et al. (2006)]. However, not much is known about their behavior when instruments are omitted in model specification. The main goal of this paper is to investigate the asymptotic size properties of the plug-in based subset tests when such problems arise.

To better circumscribe the problem, suppose that Y_1 and Y_2 depend on a second set of instru-

ments W which are not used within the subset procedures, *i.e.*:

$$y = Y_1\theta_1 + Y_2\theta_2 + \varepsilon, \tag{2.5}$$

$$(Y_1, Y_2) = Z(\Pi, \Gamma) + W(\Phi_w^1, \Phi_w^2) + (V_1, V_2), \tag{2.6}$$

where $W \in \mathbb{R}^{T \times k_w}$ is a matrix of explanatory variables (possibly correlated with u). Note that W may include any variable that could be viewed as independent of the structural disturbance u in (2.1), and could be unobservable; $\Phi_w^1 \in \mathbb{R}^{k_w \times G_1}$ and $\Phi_w^2 \in \mathbb{R}^{k_w \times G_2}$ are unknown coefficients associated with W in each first step regression. The link between (V_1, V_2) in (2.6) and (V_1^*, V_2^*) in (2.2) is:

$$(V_1^*, V_2^*) = (V_1, V_2) + W(\Phi_w^1, \Phi_w^2). \tag{2.7}$$

Under H$_0$, the reduced form for $y - Y_1\theta_{01}$ from (2.5)-(2.6) can be written as:

$$y - Y_1\theta_{01} = Z\pi^* + W\phi_w + V_2\theta_2 + \varepsilon = Z\pi^* + V_2\theta_2 + u, \tag{2.8}$$

where $\pi^* = \Gamma\theta_2$, $\phi_w = \Phi_w^2\theta_2$, and the errors u are given by:

$$u = \varepsilon + W\phi_w. \tag{2.9}$$

In this paper, I examine the size property of robust subset procedures when u is used in the computation of the statistics instead of the true errors ε. More precisely, what happens to subset procedures if instruments W are ignored while they are part of the true DGP given by (2.5)-(2.6)?

Let $\tilde{\theta}_2$ be the limited information maximum likelihood (LIML) estimator of θ_2 under H$_0$. From (2.5)-(2.8), the variable $y - Y_1\theta_{01} - Y_2\tilde{\theta}_2$ used as the dependent variable in robust subset procedures satisfies:

$$y - Y_1\theta_{01} - Y_2\tilde{\theta}_2 = Z\tilde{\pi} + \tilde{u}, \tag{2.10}$$

where $\tilde{u} = W\tilde{\phi}_w + \tilde{\varepsilon}$ and

$$\tilde{\pi} = \Pi(\theta_1 - \theta_{01}) - \Gamma(\tilde{\theta}_2 - \theta_2), \quad \tilde{\phi}_w = \Phi_w^1(\theta_1 - \theta_{01}) - \Phi_w^2(\tilde{\theta}_2 - \theta_2)$$

$$\tilde{\varepsilon} = \varepsilon + V_1(\theta_1 - \theta_{01}) - V_2(\tilde{\theta}_2 - \theta_2). \tag{2.11}$$

5

The robust subset procedures consider that the errors $\tilde{u}$ in (2.10) are asymptotically uncorrelated with Z. However, we can see from (2.10)-(2.11) that this is true only if Z and W are uncorrelated or $\tilde{\theta}_2$ is consistent to θ_2. Furthermore, even if $\theta_1 = \theta_{01}$, H_0 cannot be assessed from (2.10) *via* $H^*_\pi : \tilde{\pi} = 0$ even asymptotically. Indeed, suppose that we have $\tilde{\theta}_2 - \theta_2 \xrightarrow{p} \Delta$ under H_0, thus, $\tilde{\pi} \xrightarrow{p} \pi = \Gamma\Delta$ and $\tilde{\phi}_w \xrightarrow{p} \phi^*_w = \Phi^2_w\Delta$. If $\Delta \neq 0$, it is clear that H_0 does not entail $H^{As}_\pi : \pi = 0$ unless $\Gamma\Delta = 0$. This suggests that the omission of the instruments W may have negative impact in term of size control.

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

$$\frac{1}{T} (\varepsilon : V_1 : V_2)' (\varepsilon : V_1 : V_2) \xrightarrow{p} \Sigma > 0, \tag{2.12}$$

$$\Sigma = \begin{bmatrix} \sigma_{\varepsilon\varepsilon} & \sigma_{\varepsilon 1} & \sigma_{\varepsilon 2} \\ \sigma_{1\varepsilon} & \Sigma_{11} & \Sigma_{12} \\ \sigma_{2\varepsilon} & \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \tag{2.13}$$

where $\sigma_{\varepsilon\varepsilon} : 1 \times 1$, $\sigma_{\varepsilon 1} = \sigma'_{\varepsilon 1} : 1 \times G_1$, $\sigma_{\varepsilon 2} = \sigma'_{\varepsilon 2} : 1 \times G_2$, $\Sigma_{11} : G_1 \times G_1$, $\Sigma_{12} = \Sigma'_{12} : G_1 \times G_2$, $\Sigma_{22} : G_2 \times G_2$.

$$\frac{1}{T} (Z, W)' (Z, W) \xrightarrow{p} \Omega_{ZW} = \begin{bmatrix} \Sigma_Z & \Sigma_{ZW} \\ \Sigma'_{ZW} & \Sigma_W \end{bmatrix}, \tag{2.14}$$

$$\frac{1}{T} (Z, W)' (\varepsilon, V_1, V_2) \xrightarrow{p} \begin{bmatrix} 0 & 0 & 0 \\ \delta_{w\varepsilon} & \Sigma_{W1} & \Sigma_{W2} \end{bmatrix}, \tag{2.15}$$

where $\Sigma_Z : k_z \times k_z$, $\Sigma_{ZW} : k_z \times k_w$, $\Sigma_W : k_w \times k_w$, $\Sigma_{W1} : k_w \times G_1$, $\Sigma_{W2} : k_w \times G_2$ and $\delta_{w\varepsilon} : k_w \times 1$.

$$\frac{1}{\sqrt{T}} Z' (\varepsilon, V_1, V_2) \xrightarrow{d} (\mathscr{S}_{z\varepsilon}, \mathscr{S}_{z,1}, \mathscr{S}_{z,2}), \tag{2.16}$$

where

$$vec(\mathscr{S}_{z\varepsilon}, \mathscr{S}_{z,1}, \mathscr{S}_{z,2}) \sim N[0, \Sigma \otimes \Sigma_Z] \tag{2.17}$$

with $\mathscr{S}_{z\varepsilon} : k_z \times 1$, $\mathscr{S}_{z,1} : k_z \times G_1$, $\mathscr{S}_{z,2} : k_z \times G_2$, $\mathscr{S}_{z\varepsilon} \sim \mathrm{N}[0, \sigma_{\varepsilon\varepsilon}\Sigma_Z]$, $vec(\mathscr{S}_{z,1}) \sim \mathrm{N}[0, \Sigma_{11} \otimes \Sigma_Z]$ and $vec(\mathscr{S}_{z,2}) \sim \mathrm{N}[0, \Sigma_{22} \otimes \Sigma_Z]$.

Note that Σ_W may be singular, *i.e.* the columns of W may be redundant or asymptotically dependent. In addition, Z and W may be asymptotically correlated so that $\Sigma_{ZW} \neq 0$. Furthermore, we allow a dependence between W (the missing instruments) and ε, *i.e.* $\delta_{W\varepsilon} \neq 0$. Clearly, the above framework is quite general and allows W to include potential endogenous regressors. An important question that is often addressed is the validity of the set of available instruments. However, Doko and Dufour (2008) have shown that when identification is deficient or weak, and endogeneity moderate, using invalid instruments may result in an asymptotic efficiency gain relative to not including them in the set of available instruments. Strictly speaking, there is no reason to assume $\delta_{W\varepsilon} = 0$.

Now, from (2.12)-(2.17), we have:

$$\operatorname*{plim}_{T\to\infty} \left(\frac{Z'Y_2}{T} \right) = \Sigma_{ZY_2} = \Sigma_Z\Gamma + \Sigma_{ZW}\Phi_w^2, \ \operatorname*{plim}_{T\to\infty} \left(\frac{Z'W}{T} \right) = \Sigma_{ZW}, \quad (2.18)$$

$$\operatorname*{plim}_{T\to\infty} \left(\frac{Y_2'Y_2}{T} \right) = \Sigma_{Y_2}, \ \operatorname*{plim}_{T\to\infty} \left(\frac{Y_2'(I_T - \tilde{k}M)Y_2}{T} \right) = \tilde{\Sigma}_{Y_2}, \quad (2.19)$$

$$\operatorname*{plim}_{T\to\infty} \left(\frac{Y_2'(I_T - \tilde{k}M)W}{T} \right) = \Delta_{ZW}. \quad (2.20)$$

I will assume that

$$\mathrm{rank}(\tilde{\Sigma}_{Y_2}) = G_2 \,. \quad (2.21)$$

Assumption (2.21) replaces the usual high-level assumption of identification of θ_2. However, the results obtained can easily be extended to weak and complete non identification of θ_2. Of course, in these later two cases the concentrated LIML estimator of θ_2 under $\mathrm{H}_0 : \theta_1 = \theta_{01}$, say $\tilde{\theta}_2 = \tilde{\theta}_2(\theta_{01})$, converges to a non degenerate random variable [see Staiger and Stock (1997, Theorem 1) in the context of k-class estimators]. Strictly speaking, this assumption facilitates the exposition of the results but can be relaxed without changing their meaning. Following Startz et al. (2006), the LILM estimator $\tilde{\theta}_2$ of θ_2 computed from model (2.1)-(2.2) is given by:

$$\tilde{\theta}_2 = [Y_2'(I_T - \tilde{k}M_Z)Y_2]^{-1}Y_2'(I_T - \tilde{k}M_Z)(y - Y_1\theta_{01}), \quad (2.22)$$

where $\tilde{k} = \tilde{k}_{LILM}$ is obtained by minimizing

$$k(\theta_{01}, \theta_2) = \frac{(y - Y_1\theta_{01} - Y_2\theta_2)'(y - Y_1\theta_{01} - Y_2\theta_2)}{(y - Y_1\theta_{01} - Y_2\theta_2)'M_Z(y - Y_1\theta_{01} - Y_2\theta_2)} \tag{2.23}$$

with respect to θ_2 and $\tilde{k} = arg\min_{\theta_2}[k(\theta_{01}, \theta_2)]$. Stock and Wright (2000) have shown that if rank$(\Gamma) = G_2$ (*i.e.* the hight-level assumption for identification of θ_2 holds), $\plim_{T\to\infty}(\tilde{k}) = 1$ and $\tilde{\theta}_2$ is consistent to θ_2 , *i.e.*:

$$\plim_{T\to\infty} \tilde{\theta}_2 = \theta_2. \tag{2.24}$$

Of course, (2.24) holds under no instrument exclusion in (2.1)-(2.2). Consistency may not hold if some instruments are left out of the analysis.

Now, define:

$$
\begin{aligned}
D_Y &= Y_2[Y_2'(I_T - \tilde{k}M_Z)Y_2]^{-1}Y_2'(I_T - \tilde{k}M_Z) \\
&\equiv [d_{tj}]_{1 \le t,j \le T}, \\
\tilde{\Sigma}_{ZW} &= \plim_{T\to\infty}\left(\frac{Z'(I_T - D_Y)W}{T}\right)|_{k^*=1} \\
&= [I_{k_z,k_w} - \Sigma_{ZY_2}(\Sigma_{ZY_2}'\Sigma_Z^{-1}\Sigma_{ZY_2})^{-1}\Sigma_{ZY_2}\Sigma_Z^{-1}]\Sigma_{ZW},
\end{aligned}
\tag{2.25}
$$
$$\tag{2.26}$$

where $\Sigma_{ZY_2} = \plim_{T\to\infty}\left(\frac{Z'Y_2}{T}\right)$ and d_{tj} is the (t, j)-th element of D_Y. Let $\tilde{\varepsilon}_t = \sum_{j=1}^{T} \tilde{d}_{tj}\varepsilon_t$, where

$$
\tilde{d}_{tj} = \begin{cases}
-d_{tj} & \text{if } j \neq t \\
\\
1 - d_{tj} & \text{if } j = t.
\end{cases}
\tag{2.27}
$$

I assume that the following distributional limit holds jointly:

$$
\begin{bmatrix}
\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[Z_t'\tilde{\varepsilon}_t - \plim_{T\to\infty}\left(\frac{Z'\tilde{\varepsilon}}{T}\right)\right] \\
\frac{1}{\sqrt{T}}\left[Z'(I_T - D_Y)W - \tilde{\Sigma}_{ZW}\right]\phi_w
\end{bmatrix}
\xrightarrow{d} \mathscr{S} \equiv
\begin{bmatrix}
\mathscr{S}_d \\
\mathscr{S}_{\phi_w}
\end{bmatrix},
\tag{2.28}
$$

where $\mathscr{S} \sim N[0, \Sigma_{\mathscr{S}}]$, $\mathscr{S}_d \sim N[0, \Sigma_d]$ and $\mathscr{S}_{\phi_w} \sim N[0, \Sigma_{\phi_w}]$. Note that from the above notations, we can write $Z'(I_T - D_Y)\varepsilon = \sum_{t=1}^{T} Z_t'\tilde{\varepsilon}_t$. It is also worthwhile noting that the normality assumption can be relaxed. In particular, $\mathscr{S}$, $\mathscr{S}_d$ and $\mathscr{S}_{\phi_w}$ could have any distribution. Furthermore,

$\mathscr{S}_d$ and $\mathscr{S}_{\phi_w}$ may also be correlated. That should be the when W and ε are correlated.

Before characterizing the asymptotic behavior of $\tilde{\theta}_2$, I first present the subset statistics studied here.

3. Test statistics

This section introduces the four robust subset statistics which are studied.

(a) The AR subset statistic to test $H_0 : \theta_1 = \theta_{01}$ reads

$$\text{AR}(\theta_{01}) = \frac{1}{k_z \sigma_{\varepsilon\varepsilon}(\theta_{01})}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)'P_Z(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2), \tag{3.1}$$

where $\sigma_{\varepsilon\varepsilon}(\theta_{01}) = \frac{1}{T-k_z}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)'M_Z(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)$.

(b) Kleibergen's (2002) Lagrange multiplier (KLM) statistic to test $H_0 : \theta_1 = \theta_{01}$ reads [see Kleibergen (2004)],

$$\text{KLM}(\theta_{01}) = \frac{1}{\sigma_{\varepsilon\varepsilon}(\theta_{01})}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)'P_{Z(\tilde{\Pi}(\theta_{01}):\tilde{\Gamma}(\theta_{01}))}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2), \tag{3.2}$$

where

$$\tilde{\Pi}(\theta_{01}) = (Z'Z)^{-1}Z'[Y_1 - (y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)\frac{\sigma_{\varepsilon Y_1}(\theta_{01})}{\sigma_{\varepsilon\varepsilon}(\theta_{01})}], \tag{3.3}$$

$$\tilde{\Gamma}(\theta_{01}) = (Z'Z)^{-1}Z'[Y_2 - (y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)\frac{\sigma_{\varepsilon Y_2}(\theta_{01})}{\sigma_{\varepsilon\varepsilon}(\theta_{01})}], \tag{3.4}$$

and $\sigma_{\varepsilon Y_i}(\theta_{01}) = \frac{1}{T-k_z}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)'M_ZY_i$, $i = 1, 2$.

(c) A J-statistic that tests miss-specification under H_0, $H_M : E[Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)] = 0$, reads,

$$\text{JKLM}(\theta_{01}) = \text{AR}(\theta_{01}) - \text{KLM}(\theta_{01}). \tag{3.5}$$

(d) And a subset extension of the conditional likelihood ratio statistic to test $H_0 : \theta_1 = \theta_{01}$ reads [see Moreira (2003)]:

$$\text{MQLR}(\theta_{01}) = \frac{1}{2}(\text{AR}(\theta_{01}) - \tau_m(\theta_{01})) +$$
$$\frac{1}{2}\sqrt{[\text{AR}(\theta_{01}) + \tau_m(\theta_{01})]^2 - 4[\text{AR}(\theta_{01}) - \text{KLM}(\theta_{01})]\tau_m(\theta_{01})}, \quad (3.6)$$

where $\tau_m(\theta_{01})$ is the smallest eigenvalue of

$$\hat{\Sigma}_{MQLR}(\theta_{01}) = [\mathbb{T}(\theta_{01})]'[\mathbb{T}(\theta_{01})], \quad (3.7)$$

$$\mathbb{T}(\theta_{01}) = (Z'Z)^{\frac{1}{2}}[\tilde{\Pi}(\theta_{01}) \vdots \tilde{\Gamma}(\theta_{01})]\hat{\Sigma}^{-\frac{1}{2}}_{(Y_1:Y_2)(Y_1:Y_2).\varepsilon}, \quad (3.8)$$

$$\hat{\Sigma}^{-\frac{1}{2}}_{(Y_1:Y_2)(Y_1:Y_2).\varepsilon} = \begin{bmatrix} \hat{\Sigma}^{-\frac{1}{2}}_{Y_1Y_1.(\varepsilon:Y_2)} & 0 \\ -\hat{\Sigma}^{-1}_{Y_2Y_2.\varepsilon}\hat{\Sigma}_{Y_2Y_1.\varepsilon}\hat{\Sigma}^{-\frac{1}{2}}_{Y_1Y_1.(\varepsilon:Y_2)} & \hat{\Sigma}^{-\frac{1}{2}}_{Y_2Y_2.\varepsilon} \end{bmatrix}, \quad (3.9)$$

$$\hat{\Sigma}_{Y_1Y_1.(\varepsilon:Y_2)} = \frac{1}{T-k}Y_1'M_{(Z:Y_2:\hat{\varepsilon})}Y_1, \ \hat{\Sigma}_{Y_2Y_1.\varepsilon} = \frac{1}{T-k}Y_2'M_{(Z:\hat{\varepsilon})}Y_1,$$

$$\hat{\Sigma}_{Y_2Y_2.\varepsilon} = \frac{1}{T-k}Y_2'M_{(Z:\hat{\varepsilon})}Y_2, \ \hat{\varepsilon} = y - Y_1\theta_{01} - Y_2\tilde{\theta}_2. \quad (3.10)$$

If $\text{rank}(\Gamma) = G_2$, and no instrument is omitted, the null (conditional) limiting distribution of the above statistics are standard chi-squares [see Kleibergen (2008, Theorem 1)]. Furthermore, Kleibergen and Mavroeidis (2009) extend their validity to a weak identification setup. More precisely, the authors show that the null limiting distribution of the test statistics when $\text{rank}(\Gamma) = G_2$, provides an upper bound for their limiting distribution when the nuisance parameter is not identified. So, all subset procedures are typically identification-robust if no instrument is omitted, but may not be if instruments are missing.

Let

$$\tilde{\sigma}_{2\varepsilon} = (1 - k^*)(\Phi_w^{2'}\delta_{w\varepsilon} + \sigma_{2\varepsilon}), \ \Delta_{\theta_2}(k^*, \phi_w) = \tilde{\Sigma}^{-1}_{Y_2}(\tilde{\sigma}_{2\varepsilon} + \Delta_{ZW}\phi_w), \quad (3.11)$$

$$\tilde{\sigma}_{\varepsilon\varepsilon} = \sigma_{\varepsilon\varepsilon} + 2\phi_w'\delta_{w\varepsilon} + \phi_w'\Sigma_W\phi_w + \Delta_{\theta_2}(k^*, \phi_w)'\Sigma_{Y_2}\Delta_{\theta_2}(k^*, \phi_w)$$
$$-2[\sigma_{2\varepsilon}' + \phi_w'(\Sigma_{ZW}'\Gamma + \Sigma_W\Phi_w^2 + \Sigma_{W2})]\Delta_{\theta_2}(k^*, \phi_w), \quad (3.12)$$

$$\tilde{\sigma}_{\varepsilon, w} = [\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w)]'\Sigma_Z^{-1}[\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w)], \quad (3.13)$$

$$\sigma_{\varepsilon\varepsilon}^* = \sigma_{\varepsilon\varepsilon} + 2\phi_w'\delta_{w\varepsilon} + \phi_w'\Sigma_W\phi_w + \Delta_{\theta_2}(1, \phi_w)'\Sigma_{Y_2}\Delta_{\theta_2}(1, \phi_w)$$

$$-2[\sigma'_{2\varepsilon} + \phi'_w(\Sigma'_{ZW}\Gamma + \Sigma_W\Phi_w^2 + \Sigma_{W2})]\Delta_{\theta_2}(1, \phi_w), \tag{3.14}$$

$$k^* = \frac{\tilde{\sigma}_{\varepsilon\varepsilon}}{\tilde{\sigma}_{\varepsilon\varepsilon} - \tilde{\sigma}_{\varepsilon,w}} = 1 + \frac{\tilde{\sigma}_{\varepsilon,w}}{\tilde{\sigma}_{\varepsilon\varepsilon} - \tilde{\sigma}_{\varepsilon,w}}. \tag{3.15}$$

where $\tilde{\sigma}_{\varepsilon\varepsilon} - \tilde{\sigma}_{\varepsilon,w} \geq 0$, *i.e.* $k^* \geq 1$. Lets also define

$$\Gamma_{\phi_w} = \Gamma + \Sigma_Z^{-1}\Sigma_{ZW}\Phi_{0w}^2 - \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 2}}{\bar{\sigma}_{\varepsilon\varepsilon}}, \tag{3.16}$$

$$\Theta_{\phi_w} = \Pi_{\phi_w} - \Gamma_{\phi_w}(\Gamma'_{\phi_w}\Sigma_Z\Gamma_{\phi_w})^{-1}\Gamma'_{\phi_w}\Sigma_Z\Pi_{\phi_w}, \tag{3.17}$$

$$\Pi_{\phi_w} = \Pi_0 + \Sigma_Z^{-1}\Sigma_{ZW}\Phi_{0w}^2 - \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 1}}{\bar{\sigma}_{\varepsilon\varepsilon}}, \tag{3.18}$$

$$\Theta_b^* = \Pi_{\phi_w}^* - \Gamma_{\phi_w}(\Gamma'_{\phi_w}\Sigma_Z\Gamma_{\phi_w})^{-1}\Gamma'_{\phi_w}\Sigma_Z\Pi_{\phi_w}^*, \tag{3.19}$$

$$\Pi_{\phi_w}^* = \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 1}}{\bar{\sigma}_{\varepsilon\varepsilon}}, \tag{3.20}$$

where $\bar{\sigma}_{\varepsilon i}$, $i = 1, 2$ are given by

$$\bar{\sigma}_{\varepsilon i} = (\sigma'_{i\varepsilon} + \phi'_w\Sigma_{WY_i})(I_{G_2} - \tilde{\Sigma}_{Y_2}^{-1}\Sigma_{Y_2Y_i}) - (\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))'\Sigma_Z^{-1}\Sigma_{ZY_i}.$$

And finally, let

$$\Pi_{\phi_w}^* = \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 1}}{\bar{\sigma}_{\varepsilon\varepsilon}}, \tag{3.21}$$

$$\Theta_{\phi_w}^* = \Pi_{\phi_w}^* - \Gamma_{\phi_w}(\Gamma'_{\phi_w}\Sigma_Z\Gamma_{\phi_w})^{-1}\Gamma'_{\phi_w}\Sigma_Z\Pi_{\phi_w}^*, \tag{3.22}$$

$$\psi_{\phi_w} = \Pi_0 + \Sigma_Z^{-1}\Sigma_{ZW}\Phi_{0w}^2 + \Sigma_Z^{-1}\mathscr{S}_{z,1} - \Sigma_Z^{-1}(\mathscr{S}_d + \mathscr{S}_{\phi_w})\frac{\bar{\sigma}_{\varepsilon 1}}{\bar{\sigma}_{\varepsilon\varepsilon}}, \tag{3.23}$$

$$\Gamma_{\phi_w}^0 = \Gamma + \Sigma_Z^{-1}\Sigma_{ZW}\Phi_w^2, \quad \Lambda_{\phi_w} = \psi_{\phi_w} - \Gamma_{\phi_w}^0(\Gamma_{\phi_w}^{0'}\Sigma_Z\Gamma_{\phi_w}^0)^{-1}\Gamma_{\phi_w}^{0'}\Sigma_Z\psi_{\phi_w}. \tag{3.24}$$

Lemma **3.1** characterizes the asymptotic behavior of the concentrated LIML estimator $\tilde{\theta}_2$ of θ_2.

Lemma 3.1 ASYMPTOTIC BEHAVIOR OF $\tilde{\theta}_2$ WITH MISSING INSTRUMENTS. *Suppose the assumptions* (2.12) - (2.17) *and* (2.21) *hold, and let* $\theta_1 = \theta_{01}$. *If model* (2.1) - (2.2) *is used in the computation of* $\tilde{\theta}_2$ *instead of the true model* (2.5) - (2.6), *we have*

$$\operatorname*{plim}_{T\to\infty}(\tilde{\theta}_2 - \theta_2) = \Delta_{\theta_2}(k^*, \phi_w), \tag{3.25}$$

where $\Delta_{\theta_2}(k^*, \phi_w)$ *is defined by* (3.11).

(i) We observe that the above lemma requires $\operatorname{rank}(\tilde{\Sigma}_{Y_2}) = G_2$, which is the high-level assumption of identification of θ_2. However, the result extends to a weak identification setup from

similar assumptions, as in Staiger and Stock (1997, Theorem 1). In which case $\Delta_{\theta_2}(k^*, \phi_w)$ is a non degenerate random variable rather than a fixed vector.

(ii) We note that $\Delta_{\theta_2}(k^*, \phi_w)$ is not necessarily equal to zero even if $k^* = 1$. Hence, $k^* = 1$ is neither a necessary nor sufficient condition for consistency when instruments are missing.

(iii) We can decompose $\Delta_{\theta_2}(k^*, \phi_w)$ as:

$$\Delta_{\theta_2}(k^*, \phi_w) = \underbrace{(1 - k^*)\tilde{\Sigma}_{Y_2}^{-1}\sigma_{2\varepsilon}}_{\text{bias if any missing instruments}} + \underbrace{\tilde{\Sigma}_{Y_2}^{-1}[(1 - k^*)\Phi_w^{2\prime}\delta_{w\varepsilon} + \Delta_{ZW}\phi_w]}_{\text{bias due to missing instruments}}. \tag{3.26}$$

If any instrument is missing, the second term in (3.26) vanishes and the asymptotic bias of $\tilde{\theta}_2$ reduces to the first term. Since when no instrument is omitted we have $k^* = 1$, the first term also disappears so that $\Delta_{\theta_2}(1, \phi_w) = 0$. Hence, $\tilde{\theta}_2$ is consistent to θ_2 as expected. However, if instruments are missing, both terms may not vanish and $\Delta_{\theta_2}(k^*, \phi_w)$ may not be zero except when the sum is zero. Which means that $\tilde{\theta}_2$ is generally inconsistent with missing instruments. So, it is natural to conjecture that this inconsistency affects the asymptotic distributions of subset tests, as shown in Sections 4-5.

Now, let $\mathcal{V}(v)$ be the affine hyperplane given by:

$$\mathcal{V}(v) = \left\{ v \in \mathbb{R}^{k_w} : \mathbf{B}_w v - b_w = 0 \right\}, \tag{3.27}$$

where $\mathbf{B}_w = \Sigma_{ZW} - \Sigma_{ZY_2}\tilde{\Sigma}_{Y_2}^{-1}\Delta_{ZW}$ is a $k_z \times k_w$ matrix and $b_w = (1 - k^*)\Sigma_{ZY_2}\tilde{\Sigma}_{Y_2}^{-1}(\sigma_{2\varepsilon} + \Phi_w^{2\prime}\delta_{w\varepsilon})$ is a $k_z \times 1$ vector. We observe that if $\phi_w \in \mathcal{V}(v)$, we have $\tilde{\sigma}_{\varepsilon, w} = \Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w) = 0$ and $k^* = 1$. Let $\mathbf{B}_w^-$ be any generalized inverse of $\mathbf{B}_w$. We then have

$$v_1 \in \mathcal{V}(v) \Leftrightarrow v_1 = \mathbf{B}_w^- b_w + (I_{k_w} - \mathbf{B}_w^- \mathbf{B}_w)v_0, \tag{3.28}$$

where v_0 is any arbitrary vector in $\mathbb{R}^{k_w}$.

To give an intuitive regression interpretation of $\mathcal{V}(v)$, assume that $\tilde{k} = 1$, and define $\tilde{Y}_2 = P_Z Y_2$. Then, we have

$$Z'(I_T - D_Y)W = Z'(W - P_{\tilde{Y}_2}W) = Z'\hat{u}_W, \tag{3.29}$$

where $\hat{u}_W$ are the residuals from the regression of W on $\tilde{Y}_2$ and $P_{\tilde{Y}_2} = \tilde{Y}_2(\tilde{Y}_2'\tilde{Y}_2)^{-1}\tilde{Y}_2'$. So, under

the assumptions of Lemma **3.1**, if further $k^* = 1$, then we have:

$$\operatorname*{plim}_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) = \operatorname*{plim}_{T\to\infty}\left(\frac{Z'(I_T - D_Y)W}{T}\right) = \mathbf{B}_w\phi_w\,. \qquad (3.30)$$

Hence, since $b_w = 0$ when $k^* = 1$, we have $\phi_w \in \mathcal{V}(v) \Leftrightarrow \operatorname*{plim}_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) = 0$, *i.e.* Z is asymptotically uncorrelated with the residuals $\hat{u}_W$. Clearly, $\mathcal{V}(v)$ characterizes the set of parameters in which Z is asymptotically uncorrelated with the residuals $\hat{u}_W$. We will see in the next section that (3.30) provides a necessary condition for subset procedures to be valid (in terms of size control). Moreover, even if $k^* = 1$, we may have $\mathbf{B}_w\phi_w - b_w \neq 0$. Hence, $k^* = 1$ is not a sufficient condition for the subset tests to be valid.

To study the effect of instrument exclusion on subset statistics, I distinguish two setups, related to the strength of the missing instruments W: (I) $\Phi_w^2 = \Phi_{0w}^2$ is fixed and (II) $\Phi_w^2 = \Phi_{0w}^2/\sqrt{T}$. Setup (I) describes the case where the missing instruments are (possibly) relevant. For example, W is strong if $\operatorname{rank}(\Phi_{0w}^2) = G_2$ and weak if $0 \leq \operatorname{rank}(\Phi_{0w}^2) < G_2$. In the latter case, some linear combination of the columns of W are strong. Setup (II) represents asymptotically weak missing instruments. It is assumed here that all the columns of W are asymptotically weak [similar to Staiger and Stock (1997)]. Of course, an extension of (II) that is more relevant for practical purposes arises when at least one component of W is strong. However, the meaning of the results will not change from such an extension. Section 4 presents the asymptotic distributions of subset statistics when Φ_w^2 is fixed.

4. Asymptotic behavior with possibly relevant missing instruments

In this section, I characterize the limiting distributions of robust subset statistics when potential relevant instruments are left out of the analysis. Theorem **4.1** characterizes the limiting distributions of the subset AR statistic.

Theorem 4.1 DISTRIBUTION OF SUBSET AR STATISTIC WITH MISSING IV. *Suppose the assumptions* (2.13) - (2.17) *and* (2.28) *hold. Assume that model* (2.1) - (2.2) *is used in the computation of the statistics instead of the true model* (2.5) - (2.6). *Assume furthermore that* (2.21) *is satisfied. Let* $\Phi_w^2 = \Phi_{0w}^2$ *and* $\theta_1 = \theta_{01}$, *where* $\Phi_{0w}^2 \in \mathbb{R}^{k_w \times G_2}$ *and* $\theta_{01} \in \mathbb{R}^{G_1}$ *are fixed. If*

$\phi_w \notin \mathcal{V}(v)$:

$$AR(\theta_{01}) \quad \xrightarrow{d} \quad +\infty. \tag{4.1}$$

If $\phi_w \in \mathcal{V}(v)$:

$$AR(\theta_{01}) \xrightarrow{d} \xi^1_{d,\phi_w} = \frac{1}{k_z \sigma^*_{\varepsilon\varepsilon}} (\mathscr{S}_d + \mathscr{S}_{\phi_w})' \Sigma_Z^{-1} (\mathscr{S}_d + \mathscr{S}_{\phi_w}), \tag{4.2}$$

where $\mathscr{S}_d$ *and* $\mathscr{S}_{\phi_w}$ *are defined in* (2.28).

First, the results of Theorem **4.1** holds irrespective of whether θ_1 is identified or not, *i.e.* Π and Φ^1_w may have fixed, zero or weak (local to zero) values. Second, we note clearly that when instruments are missing, subset AR-test may be seriously size distorted in large samples. Indeed, if $\phi_w \notin \mathcal{V}(v)$, $AR(\theta_{01})$ diverges. Hence, the maximal asymptotic size distortion of this statistic is as high as 100 % when we use standard chi-square critical values. Furthermore, asymptotic size correction is infeasible since the limiting distributions of the statistics diverge. However, if $\phi_w \in \mathcal{V}(v)$, the asymptotic distribution of the AR-statistic is finite but is modified in a way that may lead to size distortions in large samples. Corollary **4.2** characterizes the necessary and sufficient condition under which subset AR-test is valid (level is controlled).

Corollary 4.2 VALIDITY OF SUBSET AR-TEST. *Suppose the assumptions of Theorem* **4.1** *hold. A necessary and sufficient condition for the subset AR test to be valid is:*

$$\underset{T \to \infty}{\text{plim}} \left(\frac{Z' \hat{u}_W}{T} \right) \quad = \quad 0 \quad and \quad \phi_w = 0, \tag{4.3}$$

where $\hat{u}_W$ *is defined in* (3.29). *More precisely, if* (4.3) *is satisfied, we have:*

$$AR(\theta_{01}) \xrightarrow{d} \frac{1}{k_z} \chi^2(k_z - G_2). \tag{4.4}$$

The above corollary indicates that the validity of the subset AR-test requires the set of instruments Z that is used the inference to be asymptotically uncorrelated with the residuals which result from the regression of the missing instruments W on the fitted value $\tilde{Y}_2 = P_Z Y_2$ in the first step regression. If W is known, this condition can be used as a rule of thumb to check whether the AR subset test is valid or not. Unfortunately, the missing instruments could be unobservable, leaving

14

this rule of thumb infeasible. It is important to observe that neither $\plim_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) = 0$ nor $\phi_w = 0$ provide the sufficient condition for the validity of the test. The condition $\phi_w = 0$ is sufficient only when W is uncorrelated with ε, *i.e.* W is exogenous. In this case, we have $\Delta_{\theta_2}(k^*, 0) = 0$ and the AR-test is valid.

I now focus on the subset KLM, JKLM and MQLR statistics. In what follows, the notation $\text{MQLR}(\theta_{01}) \overset{d}{\to} X$ reads $\text{MQLR}(\theta_{01})|_{\tau_m(\theta_{01})} \overset{d}{\to} X$, for any random variable X. Theorem **4.3** present the results.

Theorem 4.3 DISTRIBUTIONS OF KLM, JKLM AND MQLR SUBSET STATISTICS WITH MISS-ING IV. *Suppose the assumptions* (2.13) - (2.17) *and* (2.28) *hold. Assume that model* (2.1) - (2.2) *is used in the computation of the statistics instead of the true model* (2.5) - (2.6). *Assume further-more that* $\Phi_w^2 = \Phi_{0w}^2$ *and* $\theta_1 = \theta_{01}$, *where* $\Phi_{0w}^2 \in \mathbb{R}^{k_w \times G_2}$ *and* $\theta_{01} \in \mathbb{R}^{G_1}$ *are fixed. If* $\phi_w \notin \mathcal{V}(v)$, *then:*

$$KLM(\theta_{01}) \overset{d}{\to} +\infty, \ JKLM(\theta_{01}) \overset{d}{\to} +\infty, \ MQLR(\theta_{01}) \overset{d}{\to} +\infty \tag{4.5}$$

when at least one of the following conditions holds: (A) $\Pi = \Pi_0$, $\Phi_w^1 = \Phi_w^{10}$, *with* $rank(\Gamma_{\phi_w}) = G_2$ *and* $rank(\Theta_{\phi_w}) = G_1$ *or* (B) $\Pi = \Pi_0/\sqrt{T}$, $\Phi_w^1 = \Phi_w^{10}/\sqrt{T}$, *with* $rank(\Gamma_{\phi_w}) = G_2$ *and* $rank(\Theta_{\phi_w}^*) = G_1$. *If* $\phi_w \in \mathcal{V}(v)$, *then:*

$$KLM(\theta_{01}) \overset{d}{\to} \zeta_{d,\phi_w}^2, \ JKLM(\theta_{01}) \overset{d}{\to} k_z\xi_{d,\phi_w}^1 - \zeta_{d,\phi_w}^2, \tag{4.6}$$

$$\begin{aligned} MQLR(\theta_{01}) \overset{d}{\to} \ &\frac{1}{2}[k_z\xi_{d,\phi_w}^1 - \tau_m(\theta_{01})] + \\ &\frac{1}{2}\left[\sqrt{(k_z\xi_{d,\phi_w}^1 + \tau_m(\theta_{01}))^2 - 4(k_z\xi_{d,\phi_w}^1 - \zeta_{d,\phi_w}^2)\tau_m(\theta_{01})}\right], \end{aligned} \tag{4.7}$$

where

$$\zeta_{d,\phi_w}^2 = \xi_{d,\phi_w}^2 = \frac{1}{\sigma_{\varepsilon\varepsilon}^*}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'P_{\Theta_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}), \tag{4.8}$$

$$k_z\xi_{d,\phi_w}^1 - \zeta_{d,\phi_w}^2 = \frac{1}{\sigma_{\varepsilon\varepsilon}^*}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'M_{\Theta_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}), \tag{4.9}$$

15

when (A) $\Pi = \Pi_0$, $\Phi_w^1 = \Phi_w^{1^0}$ *with* $rank(\Gamma_{\phi_w}) = G_2$ *and* $rank(\Theta_{\phi_w}) = G_1$ *and:*

$$\zeta^2_{d,\phi_w} = \xi^{*2}_{d,\phi_w} = \frac{1}{\sigma^*_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'P_{\Lambda_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}), \tag{4.10}$$

$$k_z\xi^1_{d,\phi_w} - \zeta^{*2}_{d,\phi_w} = \frac{1}{\sigma^*_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'M_{\Lambda_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}), \tag{4.11}$$

when (B) $\Pi = \Pi_0/\sqrt{T}$, $\Phi_w^1 = \Phi_w^{1^0}/\sqrt{T}$, *with* $rank(\Gamma^0_{\phi_w}) = G_2$, ξ^1_{d,ϕ_w} *is defined in Theorem* **4.1**.

The above theorem is similar to Theorem **4.1**. When $\phi_w \notin \mathcal{V}(v)$, the subset KLM, JKLM and MQLR procedures are seriously size distorted in large samples, hence invalid for the hypothesis of interest. Even when $\phi_w \in \mathcal{V}(v)$, the asymptotic distribution of the statistics can be modified in a way that could lead to size distortions. Unlike Theorem **4.1**, the results of Theorem **4.3** additionally require the full rank assumption of the matrices Γ_{ϕ_w}, Θ_{ϕ_w}, and $\Gamma^0_{\phi_w}$. An extension to cases where these assumptions break down is more complex and not covered by this paper.

Corollary **4.4** characterizes the case where the subset KLM, JKLM and MQLR procedures are valid.

Corollary 4.4 VALIDITY OF SUBSET KLM, JKLM AND MQLR TESTS. *Suppose the assumptions of Theorem* **4.3** *hold. A necessary and sufficient condition for the subset KLM, JKLM and MQLR tests to be valid, is:*

$$\operatorname*{plim}_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) = 0 \quad and \quad \phi_w = 0, \tag{4.12}$$

where $\hat{u}_W$ *is defined in* (3.29). *More precisely, if this condition is satisfied, we have*

$$KLM(\theta_{01}) \xrightarrow{d} \xi^1_d \sim \chi^2(G_1), \ JKLM(\theta_{01}) \xrightarrow{d} \xi^2_d \sim \chi^2(k_z - G), \tag{4.13}$$

$$\begin{aligned} MQLR(\theta_{01}) &\xrightarrow{d} \frac{1}{2}[k_z\xi^1_d - \tau_m(\theta_{01})] + \\ &\frac{1}{2}\left[\sqrt{(k_z\xi^1_d + \tau_m(\theta_{01}))^2 - 4(k_z\xi^1_d - \xi^2_d)\tau_m(\theta_{01})}\right] \end{aligned} \tag{4.14}$$

where ξ^1_d *and* ξ^2_d *are independent distributed random variables.*

The proof of the above corollary is similar to Corollary **4.2** and is omitted. The following section focuses on the setup where the missing instruments are asymptotically weak.

5. Behavior with asymptotically weak missing instruments

In this section, I consider the missing asymptotically weak instruments setup. I characterize the asymptotically weak instruments setup by the assumption:

$$\Phi_w^2 = \Phi_{0w}^2 / \sqrt{T}, \tag{5.1}$$

where $\Phi_{0w}^2 : k_w \times G_2$ is a fixed matrix. Assumption (5.1) is similar to the weak instrument asymptotic of Staiger and Stock (1997). Theorem **5.1** characterizes the limiting distributions of the subset AR statistic.

Theorem 5.1 DISTRIBUTION OF SUBSET AR STATISTIC WHEN THE MISSING IVS ARE ASYMPTOTICALLY WEAK. *Suppose the assumptions* (2.13) - (2.17), (2.28) *and* (2.28) *hold. Assume that model* (2.1) - (2.2) *is used in the computation of the statistics instead of the true model* (2.5) - (2.6). *Let* $\Phi_w^2 = \Phi_{0w}^2 / \sqrt{T}$ *and* $H_0 : \theta_1 = \theta_{01}$, *where* Φ_{0w}^2 *is a* $k_w \times G_2$ *fixed matrix. If* $k^* \neq 1$, *then:*

$$AR(\theta_{01}) \xrightarrow{d} +\infty. \tag{5.2}$$

If $k^* = 1$, *then*

$$AR(\theta_{01}) \xrightarrow{d} \xi_{d,0}^1 = \frac{1}{k_z \sigma_{\varepsilon\varepsilon}} (\mathscr{S}_d + \mu_{\phi_w})' \Sigma_Z^{-1} (\mathscr{S}_d + \mu_{\phi_w}), \tag{5.3}$$

where $\mu_{\phi_w} = \tilde{\Sigma}_{ZW} \phi_w^0$, $\tilde{\Sigma}_{ZW}$ *is given by* (2.26) *and* $\mathscr{S}_d$ *is defined in* (2.28).

We note from the above theorem that the subset AR procedure may be invalid even if the omitted instruments are weak. In particular, if $k^* \neq 0$, $AR(\theta_{01}) \xrightarrow{d} +\infty$ and the size distortions of the AR-test are as great as 100 %. Even when $k^* = 0$, the asymptotic distribution of $AR(\theta_{01})$ is not necessarily a standard chi-square. The subset AR procedure is only valid when $\mu_{\phi_w} = 0$, as indicated in Corollary **5.2**.

Corollary 5.2 VALIDITY OF SUBSET AR-TEST WITH WEAK OMITTED INSTRUMENTS. *Suppose the assumptions of Theorem* **5.1** *hold. If* $k^* = 1$ *and* $\tilde{\Sigma}_{ZW} \phi_w^0 = 0$, *then*

$$AR(\theta_{01}) \xrightarrow{d} \frac{1}{k_z} \chi^2(k_z - G_2). \tag{5.4}$$

The proof of Corollary **5.2** is straightforward. Under the assumptions of Theorem **5.1**, if further $k^* = 1$ and $\tilde{\Sigma}_{ZW} \phi_w^0 = 0$, we have $\mu_{\phi_w} = 0$. Putting this in Corollary **4.2** yields the result. Observe that the condition $\tilde{\Sigma}_{ZW} \phi_w^0 = 0$, which insures the validity of the AR-test is a special case. Hence, the subset AR-test is generally invalid unless this condition holds. If $k_z \neq k_w$, the kernel of $\tilde{\Sigma}_{ZW}$ is different from $\{0\}$. Thus, there exists $z_0 \neq 0$ such that $\tilde{\Sigma}_{ZW} z_0 = 0$. In that case, one could have $\phi_w^0 \neq 0$ satisfying $\tilde{\Sigma}_{ZW} \phi_w^0 = 0$ and the subset AR-test procedure will still be invalid. The next theorem deals with the KLM, JKLM and MQLR subset statistics.

Theorem 5.3 DISTRIBUTIONS OF SUBSET KLM, JKLM AND MQLR STATISTICS WHEN THE MISSING IVS ARE WEAK. *Suppose the assumptions* (2.13) - (2.17), (2.28) *and* (2.28) *hold. Assume that model* (2.1) - (2.2) *is used in the computation of the statistics instead of the true model* (2.5) - (2.6). *Let* $\Phi_w^2 = \Phi_{0w}^2/\sqrt{T}$ *and* $H_0 : \theta_1 = \theta_{01}$, *where* Φ_{0w}^2 *is a* $k_w \times G_2$ *fixed matrix. If* $k^* \neq 1$, *then:*

$$KLM(\theta_{01}) \xrightarrow{d} +\infty, \ JKLM(\theta_{01}) \xrightarrow{d} +\infty, \ MQLR(\theta_{01}) \xrightarrow{d} +\infty \tag{5.5}$$

when at least one of the following conditions holds: (A) $\Pi = \Pi_0$ *with* $rank(\Gamma_0) = G_2$ *and* $rank(\Theta_0) = G_1$ *or* (B) $\Pi = \Pi_0/\sqrt{T}$ *with* $rank(\Gamma_0) = G_2$ *and* $rank(\Theta_0^*) = G_1$, *where* Γ_0, Θ_0 *and* Θ_0^* *are obtained by setting* $\phi_w = 0$ *in the expressions of* Γ_{ϕ_w}, Θ_{ϕ_w} *and* $\Theta_{\phi_w}^*$ *defined in* (3.21)-(3.24). *If* $k^* = 1$, *then:*

$$KLM(\theta_{01}) \xrightarrow{d} \zeta_{d,0}^2, \ JKLM(\theta_{01}) \xrightarrow{d} k_z \xi_{d,0}^1 - \zeta_{d,0}^2, \tag{5.6}$$

$$MQLR(\theta_{01}) \xrightarrow{d} \frac{1}{2}[k_z \xi_{d,0}^1 - \tau_m(\theta_{01})] + \frac{1}{2}\left[\sqrt{(k_z \xi_{d,0}^1 + \tau_m(\theta_{01}))^2 - 4(k_z \xi_{d,0}^1 - \zeta_{d,0}^2)\tau_m(\theta_{01})}\right], \tag{5.7}$$

where

$$\zeta_{d,0}^2 = \xi_{d,0}^2 = \frac{1}{\sigma_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mu_{\phi_w})' P_{\Theta_0}(\mathscr{S}_d + \mu_{\phi_w}), \tag{5.8}$$

$$k_z \xi_{d,0}^1 - \zeta_{d,0}^2 = \frac{1}{\sigma_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mu_{\phi_w})' M_{\Theta_0}(\mathscr{S}_d + \mu_{\phi_w}), \tag{5.9}$$

18

when (A) $\Pi = \Pi_0$ *with* $rank(\Gamma_0) = G_2$ *and* $rank(\Theta_0) = G_1$ *and*

$$\zeta_{d,0}^2 = \xi_{d,0}^{*2} = \frac{1}{\sigma_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mu_{\phi_w})' P_{\Lambda_0}(\mathscr{S}_d + \mu_{\phi_w}), \tag{5.10}$$

$$k_z \xi_{d,0}^1 - \zeta_{d,0}^{*2} = \frac{1}{\sigma_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mu_{\phi_w})' M_{\Lambda_0}(\mathscr{S}_d + \mu_{\phi_w}), \tag{5.11}$$

when (B) $\Pi = \Pi_0/\sqrt{T}$ *with* $rank(\Gamma_0) = G_2$,

$$\Lambda_0 = \psi_0 - \Gamma_0(\Gamma_0' \Sigma_Z \Gamma_0)^{-1} \Gamma_0' \Sigma_Z \psi_0,$$

$$\psi_0 = \Pi_0 + \Sigma_Z^{-1} \mathscr{S}_{z,1} - \Sigma_Z^{-1}(\mathscr{S}_d + \mu_{\phi_w})\frac{\bar{\sigma}_{\varepsilon 1}^0}{\sigma_{\varepsilon\varepsilon}}, \tag{5.12}$$

$\xi_{d,0}^1$, μ_{ϕ_w} *are defined in Theorem* **4.1**, $\bar{\sigma}_{\varepsilon i}^0$, $i = 1, 2$, *are obtained by setting* $\phi_w = 0$ *in the expression of* $\bar{\sigma}_{\varepsilon i}$, *defined in* (3.16)-(3.20).

The proof follows from Theorem **4.3** and Theorem **5.1**. Again, note that the subset KLM, JKLM and MQLR procedures may be seriously size distorted even when the missing instruments are weak. I now characterize cases where the subset KLM, JKLM and MQLR are valid.

Corollary 5.4 VALIDITY OF SUBSETKLM, JKLM AND MQLR TESTS WITH WEAK OMITTED INSTRUMENTS. *Suppose the assumptions of Theorem* **5.3** *hold. If* $k^* = 1$ *and* $\tilde{\Sigma}_{ZW}\phi_w^0 = 0$, *then:*

$$KLM(\theta_{01}) \xrightarrow{d} \xi_d^1 \sim \chi^2(G_1), \; JKLM(\theta_{01}) \xrightarrow{d} \xi_d^2 \sim \chi^2(k_z - G), \tag{5.13}$$

$$MQLR(\theta_{01}) \xrightarrow{d} \frac{1}{2}[k_z \xi_d^1 - \tau_m(\theta_{01})] +$$
$$\frac{1}{2}\left[\sqrt{(k_z \xi_d^1 + \tau_m(\theta_{01}))^2 - 4(k_z \xi_d^1 - \xi_d^2)\tau_m(\theta_{01})}\right], \tag{5.14}$$

where ξ_d^1 *and* ξ_d^2 *are independent distributed random variables.*

The proof is obtained by setting $\mu_{\phi_w} = 0$ and $k^* = 1$ in Theorem **5.3**. Section 6 presents the Monte Carlo experiment.

6. Monte Carlo simulations

In this section I explore the effect of missing instruments on the size of robust subset statistics through a Monte Carlo experiment. I consider the model with two endogenous variables described

by the following data generating process:

$$y = Y_1\theta_1 + Y_2\theta_2 + u, \quad (Y_1, Y_2) = Z\left(\Pi, \Gamma\right) + W\left(\delta_1, \delta_2\right) + (V_1, V_2). \tag{6.1}$$

The instrument $T \times k_z$ matrix Z is such that Z_t follows $i.i.d$ $N(0, I_{k_z})$ for $t = 1, \ldots, T$. The errors u, V_1 and V_2 are $T \times 1$ vectors drawn as:

$$(u_t, V_{1t}, V_{2t})' \overset{i.i.d}{\sim} N\left(0, \begin{bmatrix} 1 & .8 & .8 \\ .8 & 1 & 0 \\ .8 & 0 & 1 \end{bmatrix}\right) \quad \text{for all} \quad t = 1, \ldots, T. \tag{6.2}$$

W is a $T \times 1$ omitted instrument vector which is not taken into account when computing the different subset statistics. The sample correlation between W_t and Z_{jt} $cov(W_t, Z_{jt}) = 0.4$, for each $j = 1, \ldots, k_z$. I define:

$$\Pi = \eta_1 \Pi_0, \ \Gamma = \eta_2 \Gamma_0, \ (\delta_1, \delta_2) = \lambda\left(1, 1\right) \tag{6.3}$$

where η_1 and η_2 take the value 0 (design of complete non-identification), 0.01 (design of weak identification) or 1 (design of strong identification), $[\Pi_0, \Gamma_0]$ is a $k_z \times 2$ matrix obtained by taking the first two columns of the identity matrix of order k_z, λ takes the value 0, 0.01 and 1. If $\lambda = 0$, there is no omitted instrument. If $\lambda = 0.01$, the omitted instrument is weak and finally if $\lambda = 1$, it is relevant. The correlation coefficient between u and V_i ($i = 1, 2$) is 0.8, hence the variables Y_1 and Y_2 are endogenous and the IVs Z are necessary. I want to test the hypothesis

$$H_0 : \theta_1 = \theta_{01}. \tag{6.4}$$

The number of instruments k_z varies in $\{5, 10, 40\}$ and the true value of θ_1 and θ_2 are set at $\theta_{01} = 2$, $\theta_{02} = 5$. The simulations are run with the sample size $T = 300$, and the number of replications is $N = 10,000$. The nominal level of the tests is 5%.

The results are presented in Table 1. In the first column of the table, I report the statistics (including the projection-based, say AR_{pj}, which uses the AR-test that tests the joint Hypothesis $\theta_1 = \theta_{01}$ and $\theta_2 = \theta_{02}$.). In the second column, I report the values of k_z (number of excluded instruments). The other columns report (for each value of λ and instrument qualities η_1 and η_2), the rejection frequencies of the statistics at nominal level 5%. Except for the critical value of the subset MQLR statistic which is computed in the simulations [see Moreira (2003)], I use standard

20

chi-squares critical values for the other statistics. The main findings can be summarized as follows:

(i) the projection-based AR-test (here AR_{pj}) is typically robust to instrument exclusion whether identification is strong or weak [similar to Dufour and Taamouti (2007)],

(ii) if no instrument is missing, all plug-in based subset tests are valid (size is controlled) even if identification is deficient (weak IVs). In particular, they are conservative when IVs are weak [similar to Kleibergen (2008, 2009) and Kleibergen and Mavroeidis (2009)],

(iii) all plug-in based subset tests are seriously size distorted, with empirical rejection frequencies as high as 100 % (rather than 5 %). The distortion persists even when the omitted instrument is not very strong. However, the more relevant the missing instrument is, the larger the distortions. Furthermore, we observe that $AR(\theta_{01})$ is less size distorted than MQLR which itself is more robust than KLM. The JKLM statistic which tests the miss-specification under H_0 exhibits less distortions than the AR, KLM and MQLM tests,

(iv) overall, these results underscore the importance of using the projection techniques [see Dufour (1997), Dufour and Jasiak (2001), and Dufour and Taamouti (2005, 2007)] which do not exhibit such a problem.

Table 1. Size of robust subset tests at nominal level 5%

First block (η₁ = 0)

	k₂	λ = 0			λ = .01			λ = .1			λ = 1			λ = 10		
		η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1
AR	5	0.1	0	2.6	0.5	1.2	6.4	38.9	38.1	100	42.9	42.8	100	42.2	44.9	100
KLM	5	0	1.3	5.3	7.4	7.3	8.3	50.5	48.5	100	54.1	53.1	99.4	55	55.7	96.7
JKLM	5	1.4	0.3	5.2	0.8	0.4	11.1	4	2.8	32.3	11	9.1	96.9	9.8	10.3	99.8
MQLR	5	0.1	0	3.1	5.5	5.2	0	43.9	42.8	100	43.6	43.2	100	42	44.8	100
AR_pj	5	4.8	4.9	4.8	4.5	5	4.7	4.9	4.6	3.2	5.7	5	4.9	4.6	5.4	4.9
AR	10	0.1	0	3.3	1.1	0.7	5.8	46.3	49.1	100	54.5	50.8	100	49.3	50.3	100
KLM	10	0	2.2	5.1	8.8	9	6.4	60.8	63.1	100	68.3	67.3	98.9	65.4	66.4	94.6
JKLM	10	0.6	0.3	5.4	0.7	0.5	9	5	4.2	49.9	15.7	14	99.9	10.7	11.1	100
MQLR	10	0.1	0	3.4	7.8	7.6	0	54.3	56.6	100	55.2	51	100	49.4	50.1	100
AR_pj	10	4.2	5.2	4.5	4.3	4.5	5.2	4.2	4.9	5.5	4.4	5.5	4.4	5.8	5.6	4.2
AR	40	0.3	0.2	4.8	0.6	0.2	6.9	55.8	57.7	100	66.9	65.8	100	67.3	67	100
KLM	40	0	2.1	8.9	12.6	13	7.5	75	77.4	100	83.8	80.4	98.2	83.1	83.3	89.1
JKLM	40	0.5	0.3	4.8	0.5	0.3	7.5	6.1	7.1	75.6	31	33.3	100	30.9	30.8	100
MQLR	40	0.3	0.2	5	12.2	13.1	0	67.3	71.9	100	67.2	65.9	100	67.6	66.9	100
AR_pj	40	5.1	5.9	4.2	4.5	5.5	4	4.3	4.1	4.9	5.8	4.9	4.5	2.8	5.6	6

Second block (η₁ = 1)

	k₂	λ = 0			λ = .01			λ = .1			λ = 1			λ = 10		
		η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1	η₂=0	η₂=.01	η₂=1
AR	5	0	0	2.6	0.7	1	5.8	39.2	38.7	100	44.3	44.2	100	42.8	40.7	100
KLM	5	0.8	0.9	5	7.3	7	16.4	51.9	50	100	56.3	55.3	99.8	55.8	53.9	100
JKLM	5	0.1	0.2	4.5	0.4	0.5	4.7	6.1	5.5	11.5	8.5	10.1	85.4	6	6.6	53.4
MQLR	5	0	0	2.6	5.4	4.8	2.8	42.2	41.7	95.7	44.7	44.9	99.6	43.3	40.8	100
AR_pj	5	4.5	4.2	3.7	4.2	5	4.9	4.9	5.6	5.1	4	4.9	4.8	4.8	5.6	4.6
AR	10	0.1	0	2.4	0.3	0.6	5.7	46.4	47.4	100	53.7	53.8	100	52.4	50.6	100
KLM	10	1.4	1.8	5.7	9.1	10.3	14.1	63.2	61.8	100	68.9	68.9	100	68.3	66.4	100
JKLM	10	0.1	0.3	4.5	0.1	0.5	5.1	7.1	7.3	17.4	11.2	12.4	97.6	9.4	11.4	81.4
MQLR	10	0.1	0	2.4	8	8.7	2.9	51.3	53.7	95.9	54.4	54.6	99.9	52.3	50.9	100
AR_pj	10	5.2	5	6.3	5	5.5	5.4	4.1	5.4	5	4.8	5.4	5.3	5.9	4.9	5.1
AR	40	0.2	0.1	5.8	0.4	0.7	7.2	55.6	56.3	99.8	71	68.4	100	67.3	68.7	100
KLM	40	3.1	3.1	4.4	13.9	13	14.9	78.3	76.8	100	86.8	83.8	99.7	84.8	86	100
JKLM	40	0.3	0	6.6	0.4	0.5	7.1	17.5	16.6	37.7	30	28.1	100	24.4	25.8	99.7
MQLR	40	0.3	0.1	5.8	13.4	12.7	3.5	63.2	64	86.4	71.2	68.8	99.8	67.7	68.9	100
AR_pj	40	5.5	5.3	5.6	4.5	4.9	4.3	4	4.3	4.9	4.5	5.1	4.9	5.6	4.5	4.9

7. Conclusion

In this paper, I focus on the linear IV regression model and study the asymptotic size properties of the subset tests proposed by Kleibergen (2004, 2008), namely AR-test, KLM-test, JKLM-test and MQLR-test. I consider two main setups: the first where the missing instruments are (possibly) relevant and the second where they are asymptotically weak. In both setups, I show that the asymptotic distributions of all statistics diverges generally when instruments are omitted. Hence, all subset procedures are in general consistent against instrument inclusion (hence asymptotically invalid for the subset hypothesis of interest) I characterize cases where consistency may not hold, but the asymptotic distribution is modified in a way that would lead to size distortions in large samples. I provide a necessary and sufficient conditions under which the asymptotic distribution of the statistics remains the same as in the case where any instrument is missing (despite the exclusion of some of them). And finally, I propose a "rule of thumb" which allows to practitioners to know whether a missing instrument is detrimental or not to subset inference. I present a Monte Carlo experiment which confirms the theoretical results. I find that when no instrument is missing, all the test procedures studied are valid whether identification is strong or weak. In particular, they are conservative when identification is deficient or weak (weak instruments) [similar to Kleibergen (2008, 2009) and Kleibergen and Mavroeidis (2009)]. However, all tests are seriously size distorted with maximal size distortion as great as 100 percent if instruments are missing, while the projection method remains valid (level is controlled).

Overall, my results underscore the importance of using the projection techniques [see Dufour (1997), Dufour and Jasiak (2001), and Dufour and Taamouti (2005)Dufour-Taamouti(2005),Dufour-Taamouti(2007)] which do not exhibit such a problem.

A. Proofs

PROOF OF LEMMA **3.1**   Assume that $H_0 : \theta_1 = \theta_{01}$. From (2.1), we have

$$y - Y_1\theta_{01} \quad = \quad Y_2\theta_2 + u_0. \tag{A.1}$$

Substituting (A.1) in (2.22) gives

$$\tilde{\theta}_2 \quad = \quad \theta_2 + [Y_2'(I_T - \tilde{k}M_Z)Y_2]^{-1}Y_2'(I_T - \tilde{k}M_Z)u_0. \tag{A.2}$$

By Using the first equality in (**??**) and the notation $D_Y = [Y_2'(I_T - \tilde{k}M_Z)Y_2]^{-1}Y_2'(I_T - \tilde{k}M_Z)$, we have

$$\tilde{\theta}_2 - \theta_2 \quad = \quad D_Y\varepsilon + D_Y W\phi_w. \tag{A.3}$$

From the expression of Y_2 given by (2.6), we have under the assumptions (2.12) - (2.17)

$$\frac{Z'Y_2}{T} \quad \xrightarrow{p} \quad \Sigma_{ZY_2} = \Sigma_Z\Gamma + \Sigma_{ZW}\Phi_w^2, \tag{A.4}$$

$$\begin{aligned}
\frac{Y_2'Y_2}{T} &= \frac{(Z\Gamma + W\Phi_w^2 + V_2)'(Z\Gamma + W\Phi_w^2 + V_2)}{T} \\
&= \Gamma'\left(\frac{Z'Z}{T}\right)\Gamma + \Gamma'\left(\frac{Z'W}{T}\right)\Phi_w^2 + \Gamma'\left(\frac{Z'V_2}{T}\right) + \Phi_w^{2'}\left(\frac{W'Z}{T}\right)\Gamma \\
&\quad + \Phi_w^{2'}\left(\frac{W'W}{T}\right)\Phi_w^2 + \Phi_w^{2'}\left(\frac{W'V_2}{T}\right) + \left(\frac{V_2'Z}{T}\right)\Gamma + \left(\frac{V_2'W}{T}\right)\Phi_w^2 \\
&\quad + \left(\frac{V_2'V_2}{T}\right) \xrightarrow{p} \Sigma_{Y_2},
\end{aligned} \tag{A.5}$$

$$\begin{aligned}
\Sigma_{Y_2} &= \Gamma'\Sigma_Z\Gamma + \Gamma'\Sigma_{ZW}\Phi_w^2 + \Phi_w^{2'}\Sigma_{ZW}'\Gamma + \Phi_w^{2'}\Sigma_W\Phi_w^2 + \Phi_w^{2'}\Sigma_{W2} \\
&\quad + \Sigma_{W2}'\Phi_w^2 + \Sigma_{22}
\end{aligned} \tag{A.6}$$

$$\begin{aligned}
\frac{Y_2'(I_T - \tilde{k}M_Z)Y_2}{T} &= (1 - \tilde{k})\frac{Y_2'Y_2}{T} + \tilde{k}\frac{Y_2'Z}{T}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'Y_2}{T} \\
&\xrightarrow{p} \tilde{\Sigma}_{Y_2} = (1 - k^*)\Sigma_{Y_2} + k^*\Sigma_{ZY_2}'\Sigma_Z^{-1}\Sigma_{ZY_2},
\end{aligned} \tag{A.7}$$

$$\frac{Y_2'(I_T - \tilde{k}M_Z)\varepsilon}{T} \xrightarrow{p} (1 - k^*)(\Phi_w^{2'}\delta_{w\varepsilon} + \sigma_{2\varepsilon}) = \tilde{\sigma}_{2\varepsilon}. \tag{A.8}$$

$$\frac{Y_2'(I_T - \tilde{k}M_Z)W}{T} \xrightarrow{p} \Delta_{ZW} = \Phi_w^{2'}\Sigma_W + \Gamma'\Sigma_{ZW} + (1 - k^*)\Sigma_{W2}' +$$

$$-k^*\Phi_w^{2'}(\Sigma_W - \Sigma_{ZW}'\Sigma_Z^{-1}\Sigma_{ZW}). \tag{A.9}$$

Hence, provided $\text{rank}(\tilde{\Sigma}_{Y_2}) = G_2$, we have

$$\frac{D_Y\varepsilon}{T} \xrightarrow{p} \tilde{\Sigma}_{Y_2}^{-1}\tilde{\sigma}_{2\varepsilon}, \quad \frac{D_Y W}{T}\phi_w \xrightarrow{p} \tilde{\Sigma}_{Y_2}^{-1}\Delta_{ZW}\phi_w, \tag{A.10}$$

$$\tilde{\theta}_2 - \theta_2 \xrightarrow{p} \Delta_{\theta_2}(k^*, \phi_w) = \tilde{\Sigma}_{Y_2}^{-1}(\tilde{\sigma}_{2\varepsilon} + \Delta_{ZW}\phi_w), \tag{A.11}$$

where k^* is defined by $\plim_{T\to\infty}(\tilde{k}) = k^*$. We note that even if $k^* = 1$, $\Delta_{\theta_2}(1, \phi_w)$ may be different from zero. Hence, the condition $k^* = 1$ is not sufficient for identifying θ_2 in presence of missing instruments. The formula that determines k^* (implicitly) is given by (3.15).

$\square$

PROOF OF THEOREM **4.1**   Under $H_0 : \theta_1 = \theta_{01}$, we have

$$
\begin{aligned}
\varepsilon(\theta_{01}) &= y - Y_1\theta_{01} - Y_2\tilde{\theta}_2 = y - Y_1\theta_{01} - Y_2\theta_2 - Y_2(\tilde{\theta}_2 - \theta_2) \\
&= u_0 - Y_2(\tilde{\theta}_2 - \theta_2) = (I_T - D_Y)u_0,
\end{aligned}
\tag{A.12}
$$

where $D_Y = Y_2[Y_2'(I_T - \tilde{k}M)Y_2]^{-1}Y_2'(I_T - \tilde{k}M)$. So, we can write the numerator of the AR-statistic (times k) as

$$\varepsilon(\theta_{01})'P_Z\varepsilon(\theta_{01}) = T\left[\frac{u_0'(I_T - D_Y)'P_Z(I_T - D_Y)u_0}{T}\right] =$$

$$T\left[\frac{\varepsilon'(I_T - D_Y)'P_Z(I_T - D_Y)\varepsilon}{T} - \frac{\phi_w'W'(I_T - D_Y)'P_Z(I_T - D_Y)W\phi_w}{T}\right], \tag{A.13}$$

where the last equality in (A.13) holds because $u = \varepsilon + W\phi_w$. Let focus on each term of (A.13) (without times T). The first is such that

$$\frac{\varepsilon'(I_T - D_Y)'P_Z(I_T - D_Y)\varepsilon}{T} = \frac{\varepsilon'(I_T - D_Y)'Z}{T}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'(I_T - D_Y)\varepsilon}{T}, \tag{A.14}$$

where

$$\frac{Z'(I_T - D_Y)\varepsilon}{T} = \frac{Z'\varepsilon}{T} - \frac{Z'Y_2}{T}\left[\frac{Y_2(I_T - \tilde{k}M)'Y_2}{T}\right]^{-1}\frac{Y_2(I_T - \tilde{k}M)'\varepsilon}{T}$$

$$\xrightarrow{p} -\Sigma_{ZY_2}\tilde{\Sigma}_{Y_2}^{-1}\tilde{\sigma}_{2\varepsilon}. \tag{A.15}$$

Thus we have

$$\frac{\varepsilon'(I_T - D_Y)'P_Z(I_T - D_Y)\varepsilon}{T} \xrightarrow{p} \tilde{\sigma}_{2\varepsilon}'\tilde{\Sigma}_{Y_2}^{-1}\Sigma_{ZY_2}'\Sigma_Z^{-1}\Sigma_{ZY_2}\tilde{\Sigma}_{Y_2}^{-1}\tilde{\sigma}_{2\varepsilon}. \tag{A.16}$$

By the same way, we get for the second term

$$\frac{\phi_w'W'(I_T - D_Y)'P_Z(I_T - D_Y)W\phi_w}{T} \xrightarrow{p} \sigma_{\phi_w} \tag{A.17}$$

$$\sigma_{\phi_w} = \phi_w'(\Sigma_{ZW} - \Sigma_{ZY_2}\tilde{\Sigma}_{Y_2}^{-1}\Delta_{ZW})'\Sigma_Z^{-1}(\Sigma_{ZW} - \Sigma_{ZY_2}\tilde{\Sigma}_{Y_2}^{-1}\Delta_{ZW})\phi_w. \tag{A.18}$$

So, we find

$$\left[\frac{u_0'(I_T - D_Y)'P_Z(I_T - D_Y)u_0}{T}\right] \xrightarrow{p} \tilde{\sigma}_{\varepsilon,w} = [\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w)]'\Sigma_Z^{-1} \times$$

$$[\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w)]. \tag{A.19}$$

Furthermore, using the third equality in (A.12), *i.e.*, $y - Y_1\theta_{01} - Y_2\tilde{\theta}_2 = u_0 - Y_2(\tilde{\theta}_2 - \theta_2)$, we have

$$\frac{\varepsilon(\theta_{01})'\varepsilon(\theta_{01})}{T} = \frac{u_0'u_0}{T} - 2\frac{u_0'Y_2}{T}(\tilde{\theta}_2 - \theta_2) + (\tilde{\theta}_2 - \theta_2)'\frac{Y_2'Y_2}{T}(\tilde{\theta}_2 - \theta_2), \tag{A.20}$$

where

$$\frac{u_0'u_0}{T} = \frac{(\varepsilon + W\phi_w)'(\varepsilon + W\phi_w)}{T} \xrightarrow{p} \sigma_{\varepsilon\varepsilon} + 2\phi_w'\delta_{w\varepsilon} + \phi_w'\Sigma_W\phi_w, \tag{A.21}$$

$$\frac{u_0'Y_2}{T} = \frac{(\varepsilon + W\phi_w)'Y_2}{T} \xrightarrow{p} \sigma_{2\varepsilon}' + \phi_w'(\Sigma_{ZW}'\Gamma + \Sigma_W\Phi_w^2 + \Sigma_{W2}). \tag{A.22}$$

Hence, we get

$$\frac{\varepsilon(\theta_{01})'\varepsilon(\theta_{01})}{T} \xrightarrow{p} \tilde{\sigma}_{\varepsilon\varepsilon} = \sigma_{\varepsilon\varepsilon} + 2\phi_w'\delta_{w\varepsilon} + \phi_w'\Sigma_W\phi_w$$

$$-2[\sigma_{2\varepsilon}' + \phi_w'(\Sigma_{ZW}'\Gamma + \Sigma_W\Phi_w^2 + \Sigma_{W2})]\Delta_{\theta_2}(k^*, \phi_w) + \Delta_{\theta_2}(k^*, \phi_w)'\Sigma_{Y_2}\Delta_{\theta_2}(k^*, \phi_w), \tag{A.23}$$

26

and

$$\frac{\varepsilon(\theta_{01})'M_Z\varepsilon(\theta_{01})}{T} \quad \xrightarrow{p} \quad \tilde{\sigma}_{\varepsilon\varepsilon} - \tilde{\sigma}_{\varepsilon,w} \geq 0, \tag{A.24}$$

where

$$k^* = \frac{\tilde{\sigma}_{\varepsilon\varepsilon}}{(\tilde{\sigma}_{\varepsilon\varepsilon} - \tilde{\sigma}_{\varepsilon,w})}. \tag{A.25}$$

If $b \notin \mathcal{V}(v)$, we have $\tilde{\sigma}_{\varepsilon,w} > 0$ and from (A.13), we have

$$\varepsilon(\theta_{01})'P_Z\varepsilon(\theta_{01}) \quad \xrightarrow{p} \quad +\infty. \tag{A.26}$$

Since the denominator of the AR statistic, say $\frac{\varepsilon(\theta_{01})'M_Z\varepsilon(\theta_{01})}{T-k}$ converges to $\tilde{\sigma}_{\varepsilon\varepsilon} - \tilde{\sigma}_{\varepsilon,w} \geq 0$, it is clear that

$$\mathrm{AR}(\theta_{01}) \xrightarrow{d} +\infty. \tag{A.27}$$

Suppose now that $\phi_w \in \mathcal{V}(v)$, we have $\tilde{\sigma}_{\varepsilon,w} = 0$, *i.e.* $\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*,\phi_w) = 0$. From (A.25), this is equivalent to $k^* = 1$ and $\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*,\phi_w) = 0$ becomes $\tilde{\Sigma}_{ZW}\phi_w = 0$, where $\tilde{\Sigma}_{ZW} = [I_{k_z,k_w} - \Sigma_{ZY_2}\tilde{\Sigma}_{Y_2}^{-1}(\Gamma' + \Phi_w^{2'}\Sigma_{ZW}'\Sigma_Z^{-1}]\Sigma_{ZW}$. We can then write $\frac{1}{\sqrt{T}}Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)$ as

$$\begin{aligned}
\frac{Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)}{\sqrt{T}} &= \frac{1}{\sqrt{T}}[Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2) - \tilde{\Sigma}_{ZW}\phi_w] \\
&= \frac{Z'(I_T - D_Y)\varepsilon}{\sqrt{T}} + \frac{[Z'(I_T - D_Y)W - \tilde{\Sigma}_{ZW}]\phi_w}{\sqrt{T}}, \tag{A.28}
\end{aligned}$$

where we recall $D_Y = Y_2[Y_2'(I_T - \tilde{k}M_Z)Y_2]^{-1}Y_2'(I_T - \tilde{k}M_Z)$. Note that the presence of the term $\frac{[Z'(I_T - D_Y)W - \Sigma_{ZW}]\phi_w}{\sqrt{T}}$ in (A.28) is due to missing instrument W. If no instrument is missing, $\phi_w = 0$ and this term vanishes from (A.28).

Now, from (2.28), we have

$$\frac{1}{\sqrt{T}}Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2) \quad \xrightarrow{d} \quad \mathscr{S}_d + \mathscr{S}_{\phi_w} \tag{A.29}$$

and

$$\varepsilon(\theta_{01})' P_Z \varepsilon(\theta_{01}) \xrightarrow{d} (\mathscr{S}_d + \mathscr{S}_{\phi_w})' \Sigma_Z^{-1} (\mathscr{S}_d + \mathscr{S}_{\phi_w}). \tag{A.30}$$

Furthermore, we have

$$\frac{1}{T-k}\varepsilon(\theta_{01})' M_Z \varepsilon(\theta_{01}) \xrightarrow{d} \sigma_{\varepsilon\varepsilon} + 2\phi_w' \delta_{w\varepsilon} + \phi_w' \Sigma_W \phi_w + \Delta_{\theta_2}(1,\phi_w)' \Sigma_{Y_2} \Delta_{\theta_2}(1,\phi_w)$$
$$-2[\sigma_{2\varepsilon}' + \phi_w'(\Sigma_{ZW}'\Gamma + \Sigma_W \Phi_w^2 + \Sigma_{W2})]\Delta_{\theta_2}(1,\phi_w)$$
$$= \sigma_{\varepsilon\varepsilon}^*.$$

Thus

$$\mathrm{AR}(\theta_{01}) \xrightarrow{d} \xi_{d,\phi_w}^1 = \frac{1}{k_z \sigma_{\varepsilon\varepsilon}^*}(\mathscr{S}_d + \mathscr{S}_{\phi_w})' \Sigma_Z^{-1} (\mathscr{S}_d + \mathscr{S}_{\phi_w}). \tag{A.31}$$

$\square$

PROOF OF COROLLARY **4.2** (A) NC: Suppose that (4.3) is not satisfied. It is useful to distinguish the following three cases: (i) $\plim_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) \neq 0$ and $\phi_w = 0$; (ii) $\plim_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) = 0$ and $\phi_w \neq 0$; and (iii) $\plim_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) \neq 0$ and $\phi_w \neq 0$. By noting that $\phi_w \notin \mathcal{V}(v) \Leftrightarrow \plim_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) \neq 0$, from the first part of Theorem **4.1**, $\mathrm{AR}(\theta_{01})$ diverges under cases (i) and (iii) hold. Now, assume that (ii) is satisfied. From the second part of Theorem **4.1**, we have

$$\mathrm{AR}(\theta_{01}) \xrightarrow{d} \xi_{d,\phi_w}^1 = \frac{1}{k_z \sigma_{\varepsilon\varepsilon}^*}(\mathscr{S}_d + \mathscr{S}_{\phi_w})' \Sigma_Z^{-1} (\mathscr{S}_d + \mathscr{S}_{\phi_w}). \tag{A.32}$$

where $\mathscr{S}_{\phi_w}$ is a non degenrate random variable. Hence, ξ_{d,ϕ_w}^1 cannot follow standard chi-square with $k_z - G_2$ degrees of freedom. So, the AR subset test is invalid. So, the condition (4.3) is necessary for the validity of the subset AR-test.

(B) SC: Suppose that

$$\plim_{T\to\infty}\left(\frac{Z'\hat{u}_W}{T}\right) = 0 \quad \text{and} \quad \phi_w = 0. \tag{A.33}$$

Since $\phi_w \in \mathcal{V}(v) \Leftrightarrow \plim_{T \to \infty} \left(\frac{Z' \hat{u}_W}{T} \right) = 0$, from Theorem **4.1**, we have

$$\text{AR}(\theta_{01}) \xrightarrow{d} \xi^1_{d,\phi_w} = \frac{1}{k_z \sigma^*_{\varepsilon\varepsilon}} (\mathscr{S}_d + \mathscr{S}_{\phi_w})' \Sigma_Z^{-1} (\mathscr{S}_d + \mathscr{S}_{\phi_w}). \tag{A.34}$$

If further $\phi_w = 0$, then $\mathscr{S}_{\phi_w} \equiv 0$ so that $\Delta_{\theta_2}(1,0) = 0$ and $\sigma^*_{\varepsilon\varepsilon} = \sigma_{\varepsilon\varepsilon}$. Hence, we get

$$\text{AR}(\theta_{01}) \xrightarrow{d} \xi^1_{d,0} \equiv \xi^1_d = \frac{1}{k_z \sigma_{\varepsilon\varepsilon}} \mathscr{S}_d' \Sigma_Z^{-1} \mathscr{S}_d. \tag{A.35}$$

Furthermore, we can show that

$$\begin{aligned}
\mathscr{S}_d &= [I_{k_z} - \Sigma_{ZY_2}(\Sigma'_{ZY_2} \Sigma_Z^{-1} \Sigma_{ZY_2})^{-1} \Sigma'_{ZY_2} \Sigma_Z^{-1}] \mathscr{S}_{z,\varepsilon} \\
\xi^1_d &= \mathscr{S}'_{z,\varepsilon} \Sigma_Z^{-\frac{1}{2}} M_{\Sigma_Z^{-\frac{1}{2}} \Sigma_{ZY_2}} \Sigma_Z^{-\frac{1}{2}} \mathscr{S}_{z,\varepsilon},
\end{aligned} \tag{A.36}$$

where $M_B = I_{k_z} - B(B'B)^{-1}B'$, $B = \Sigma_Z^{-\frac{1}{2}} \Sigma_{ZY_2}$. Since $M_{\Sigma_Z^{-\frac{1}{2}} \Sigma_{ZY_2}}$ is idempotent of rank G_2 and $\Sigma_Z^{-\frac{1}{2}} \mathscr{S}_{z,\varepsilon} \sim \text{N}[0, \sigma_{\varepsilon\varepsilon} I_{k_z}]$, we have

$$\xi^1_d \sim \frac{1}{k_z} \chi^2(k_z - G_2). \tag{A.37}$$

$\square$

PROOF OF THEOREM **4.3**   Let $\text{H}_0 : \theta_1 = \theta_{01}$. For both setups $\phi_w \notin \mathcal{V}(v)$ and $\phi_w \in \mathcal{V}(v)$, we shall distinguish two cases: (A) $\Pi = \Pi_0$, where Π_0 is a $k_z \times G_1$ fixed matrix and (B) $\Pi = \Pi_0/\sqrt{T}$, where Π_0 is a $k_z \times G_1$ fixed matrix ($\Pi_0 = 0$ is allowed).

Suppose first that $\phi_w \notin \mathcal{V}(v)$.

Let (A) $\Pi = \Pi_0$ and $\Phi^1_w = \Phi^{1^0}_w$, where Π_0 and $\Phi^{1^0}_w$ are $k_z \times G_1$ and $k_w \times G_1$ fixed matrices. We first prove the result for KLM. This statistic can be written as

$$\begin{aligned}
\text{KLM}(\theta_{01}) &= \frac{1}{\sigma_{\varepsilon\varepsilon}(\theta_{01})} \varepsilon(\theta_{01})' P_{M_{Z\tilde{\Gamma}(\theta_{01})} Z\tilde{\Pi}(\theta_{01})} \varepsilon(\theta_{01}) & (A.38) \\
&= \frac{1}{\sigma_{\varepsilon\varepsilon}(\theta_{01})} \varepsilon(\theta_{01})' Z' \tilde{\Theta} [\tilde{\Theta}'(Z'Z)\tilde{\Theta}]^{-1} \tilde{\Theta} Z \varepsilon(\theta_{01}) & (A.39) \\
&= \frac{T}{\sigma_{\varepsilon\varepsilon}(\theta_{01})} \left(\frac{\varepsilon(\theta_{01})'Z'}{T} \right) \tilde{\Theta} \left[\tilde{\Theta}' \left(\frac{Z'Z}{T} \right) \tilde{\Theta} \right]^{-1} \tilde{\Theta} \left(\frac{Z\varepsilon(\theta_{01})}{T} \right), & (A.40)
\end{aligned}$$

where $\tilde{\Theta} \equiv \tilde{\Theta}(\theta_{01}) = \tilde{\Pi}(\theta_{01}) - \tilde{\Gamma}(\theta_{01}) \left[\tilde{\Gamma}(\theta_{01})' \left(\frac{Z'Z}{T} \right) \tilde{\Gamma}(\theta_{01}) \right]^{-1} \tilde{\Gamma}(\theta_{01}) \left(\frac{Z'Z}{T} \right) \tilde{\Pi}(\theta_{01})$.

Since

$$
\begin{aligned}
\sigma_{\varepsilon\varepsilon}(\theta_{01}) &\overset{p}{\to} \bar{\sigma}_{\varepsilon\varepsilon} > 0, \\
\sigma_{\varepsilon Y_i}(\theta_{01}) \overset{p}{\to} \bar{\sigma}_{\varepsilon i} &= \sigma'_{i\varepsilon} + \phi'_w \Sigma_{WY_i} - (\sigma'_{i\varepsilon} + \phi'_w \Sigma_{WY_2}) \tilde{\Sigma}_{Y_2}^{-1} \Sigma_{Y_2 Y_i} + \\
&\quad -(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))' \Sigma_Z^{-1} \Sigma_{ZY_i} \\
\tilde{\Pi}(\theta_{01}) &\overset{p}{\to} \Pi_{\phi_w} = \Pi + \Sigma_Z^{-1}\Sigma_{ZW}\Phi^1_w - \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 1}}{\bar{\sigma}_{\varepsilon\varepsilon}}, \\
\tilde{\Gamma}(\theta_{01}) &\overset{p}{\to} \Gamma_{\phi_w} = \Gamma + \Sigma_Z^{-1}\Sigma_{ZW}\Phi^2_w - \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 2}}{\bar{\sigma}_{\varepsilon\varepsilon}},
\end{aligned}
$$

provided $\text{rank}(\Gamma_{\phi_w}) = G_2$, we have

$$
\tilde{\Theta}(\theta_{01}) \overset{p}{\to} \Theta_{\phi_w} = \Pi_{\phi_w} - \Gamma_{\phi_w}(\Gamma'_{\phi_w}\Sigma_Z\Gamma_{\phi_w})^{-1}\Gamma'_{\phi_w}\Sigma_Z\Pi_{\phi_w} \tag{A.41}
$$

If further Θ_{ϕ_w} has full columns rank, then

$$
\begin{aligned}
\frac{Z\varepsilon(\theta_{01})}{T} &\overset{p}{\to} \omega(k^*, \phi_w) = \Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w) \neq 0, \\
\left(\frac{\varepsilon(\theta_{01})'Z'}{T}\right)\tilde{\Theta}\left[\tilde{\Theta}'\left(\frac{Z'Z}{T}\right)\tilde{\Theta}\right]^{-1}\tilde{\Theta}\left(\frac{Z\varepsilon(\theta_{01})}{T}\right) &\overset{p}{\to} \omega(k^*, \phi_w)'\Theta_{\phi_w}(\Theta'_{\phi_w}\Sigma_Z\Theta_{\phi_w})^{-1}\Theta_{\phi_w} \times \\
&\quad \omega(k^*, \phi_w) > 0
\end{aligned}
$$

so that we get

$$
\text{KLM}(\theta_{01}) \overset{d}{\to} +\infty. \tag{A.42}
$$

By the same way, we can show that

$$
\text{JKLM}(\theta_{01}) = \frac{1}{\sigma_{\varepsilon\varepsilon}(\theta_{01})}\varepsilon(\theta_{01})'M_{M_{Z\tilde{\Gamma}(\theta_{01})}Z\tilde{\Pi}(\theta_{01})}\varepsilon(\theta_{01}) \overset{d}{\to} +\infty. \tag{A.43}
$$

Now, let focus on $\text{MQLR}(\theta_{01})$. Following Kleibergen (2008), we have

$$
\begin{aligned}
\frac{\partial \text{MQLR}(\theta_{01})}{\partial \text{KLM}(\theta_{01})} &= \frac{1}{2}[1 + \frac{\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) + \tau_m(\theta_{01})}{\sqrt{[\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) + \tau_m(\theta_{01})]^2 - 4\text{JKLM}(\theta_{01})\tau_m(\theta_{01})}}] \\
&\geq 0, \tag{A.44}
\end{aligned}
$$

and

$$\frac{\partial \text{MQLR}(\theta_{01})}{\partial \text{JKLM}(\theta_{01})} = \frac{1}{2}[1 + \frac{\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) - \tau_m(\theta_{01})}{\sqrt{[\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) + \tau_m(\theta_{01})]^2 - 4\text{JKLM}(\theta_{01})\tau_m(\theta_{01})}}]$$
$$\geq 0. \tag{A.45}$$

Note that (A.44) is obvious. For the condition (A.45), we can see that: (1) if $\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) - \tau_m(\theta_{01}) \geq 0$, the result is obvious; (2) if $\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) - \tau_m(\theta_{01}) \leq 0$, we have

$$\frac{1}{2}[1 + \frac{\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) - \tau_m(\theta_{01})}{\sqrt{[\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) + \tau_m(\theta_{01})]^2 - 4\text{JKLM}(\theta_{01})\tau_m(\theta_{01})}}]$$
$$= \frac{1}{2}[1 + \frac{\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) - \tau_m(\theta_{01})}{\sqrt{[\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) - \tau_m(\theta_{01})]^2 + 4\text{KLM}(\theta_{01})\tau_m(\theta_{01})}}]$$
$$= \frac{1}{2}\left[1 - \frac{1}{\sqrt{1 + \frac{4\text{KLM}(\theta_{01})\tau_m(\theta_{01})}{[\text{KLM}(\theta_{01}) + \text{JKLM}(\theta_{01}) - \tau_m(\theta_{01})]^2}}}\right] \geq 0. \tag{A.46}$$

So, the derivatives of $\text{MQLR}(\theta_{01})$ with respect to both $\text{KLM}(\theta_{01})$ and $\text{JKLM}(\theta_{01})$ are non-negative. Since when $\phi_w \notin \mathcal{V}(v)$, we have $\text{KLM}(\theta_{01}) \xrightarrow{d} +\infty$ and $\text{JKLM}(\theta_{01}) \xrightarrow{d} +\infty$, (A.44)-(A.45) entail that

$$\text{MQLR}(\theta_{01}) \xrightarrow{d} +\infty. \tag{A.47}$$

(B) Now, let $\Pi = \Pi_0/\sqrt{T}$ and $\Pi = \Phi_w^{1^0}/\sqrt{T}$, where ($\Pi_0 = 0$ and $\Phi_w^{1^0} = 0$ are allowed). Consider the KLM statistic defined in (A.40). We have

$$\sigma_{\varepsilon\varepsilon}(\theta_{01}) \xrightarrow{p} \bar{\sigma}_{\varepsilon\varepsilon},$$
$$\sigma_{\varepsilon Y_i}(\theta_{01}) \xrightarrow{p} \bar{\sigma}_{\varepsilon i} = \sigma'_{i\varepsilon} + \phi'_w \Sigma_{WY_i} - (\sigma'_{i\varepsilon} + \phi'_w \Sigma_{WY_2})\tilde{\Sigma}_{Y_2}^{-1}\Sigma_{Y_2 Y_i} +$$
$$- (\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))'\Sigma_Z^{-1}\Sigma_{ZY_i}, \ i = 1, 2$$
$$\tilde{\Pi}(\theta_{01}) \xrightarrow{p} \Pi^*_{\phi_w} = \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 1}}{\bar{\sigma}_{\varepsilon\varepsilon}},$$
$$\tilde{\Gamma}(\theta_{01}) \xrightarrow{p} \Gamma_{\phi_w} = \Gamma + \Sigma_Z^{-1}\Sigma_{ZW}\Phi_w^2 - \Sigma_Z^{-1}(\Sigma_{ZW}\phi_w - \Sigma_{ZY_2}\Delta_{\theta_2}(k^*, \phi_w))\frac{\bar{\sigma}_{\varepsilon 2}}{\bar{\sigma}_{\varepsilon\varepsilon}}.$$

Hence, if $\text{rank}(\Gamma_{\phi_w}) = G_2$, we have

$$\tilde{\Theta}(\theta_{01}) \xrightarrow{p} \Theta^*_{\phi_w} = \Pi^*_{\phi_w} - \Gamma_{\phi_w}(\Gamma'_{\phi_w}\Sigma_Z\Gamma_{\phi_w})^{-1}\Gamma'_{\phi_w}\Sigma_Z\Pi^*_{\phi_w}. \tag{A.48}$$

If further $\phi_w \notin \mathcal{V}(v)$ and $\Theta^*_{\phi_w}$ has full columns rank, then

$$\left(\frac{\varepsilon(\theta_{01})'Z'}{T}\right)\tilde{\Theta}\left[\tilde{\Theta}'\left(\frac{Z'Z}{T}\right)\tilde{\Theta}\right]^{-1}\tilde{\Theta}\left(\frac{Z\varepsilon(\theta_{01})}{T}\right) \quad \overset{p}{\to} \quad \omega^{*\prime}(k^*,\phi_w)\Theta^*_{\phi_w}(\Theta^{*\prime}_{\phi_w}\Sigma_Z\Theta^*_{\phi_w})^{-1}\Theta^{*\prime}_{\phi_w} \times$$
$$\omega^*(k^*,\phi_w) > 0 \tag{A.49}$$

so that

$$\mathrm{KLM}(\theta_{01}) \overset{d}{\to} +\infty. \tag{A.50}$$

By the same arguments, we get

$$\mathrm{JKLM}(\theta_{01}) \overset{d}{\to} +\infty, \tag{A.51}$$

and from (A.44)-(A.47), we also have

$$\mathrm{MQLR}(\theta_{01}) \overset{d}{\to} +\infty. \tag{A.52}$$

Suppose now that $\phi_w \in \mathcal{V}(v)$.

As before, assume first that (A) $\Pi = \Pi_0$ and $\Phi^1_w = \Phi^{1^0}_w$. From the proof of Theorem **4.1**, we have

$$\frac{Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)}{\sqrt{T}} \quad \overset{d}{\to} \quad \mathscr{S}_d + \mathscr{S}_{\phi_w}. \tag{A.53}$$

The numerator of KLM can be written as

$$\varepsilon(\theta_{01})'P_{M_{Z\tilde{\Gamma}(\theta_{01})}Z\tilde{\Pi}(\theta_{01})}\varepsilon(\theta_{01}) = \frac{\varepsilon(\theta_{01})'Z'}{\sqrt{T}}\tilde{\Theta}(\theta_{01})\left[\tilde{\Theta}'(\frac{Z'Z}{T})\tilde{\Theta}\right]^{-1}\tilde{\Theta}(\theta_{01})'\frac{Z\varepsilon(\theta_{01})}{\sqrt{T}}, \tag{A.54}$$

where $\sigma_{\varepsilon\varepsilon}(\theta_{01}) \overset{p}{\to} \sigma^*_{\varepsilon\varepsilon}$, and $\sigma_{\varepsilon Y_i}(\theta_{01})/T \overset{p}{\to} \bar{\sigma}_{\varepsilon i}$, $i = 1, 2$. So, if $\mathrm{rank}(\Gamma_{\phi_w}) = G_2$, we have

$$\tilde{\Pi} \quad \overset{p}{\to} \quad \Pi_{\phi_w}, \ \tilde{\Gamma} \overset{p}{\to} \Gamma_{\phi_w},$$
$$\tilde{\Theta}(\theta_{01}) \quad \overset{p}{\to} \quad \Theta_{\phi_w} = \Pi_{\phi_w} - \Gamma_{\phi_w}(\Gamma'_{\phi_w}\Sigma_Z\Gamma_{\phi_w})^{-1}\Gamma'_{\phi_w}\Sigma_Z\Pi_{\phi_w}, \tag{A.55}$$

If further Θ_{ϕ_w} has full columns rank, we have

$$\mathrm{KLM}(\theta_{01}) \quad \overset{d}{\to} \quad \xi^2_{d,\phi_w} = \frac{1}{\sigma^*_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'P_{\Theta_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}). \tag{A.56}$$

where $P_{\Theta_{\phi_w}} = \Theta_{\phi_w}(\Theta'_{\phi_w}\Theta_{\phi_w})^{-1}\Theta'_{\phi_w}$. By following the same way, we get

$$\text{JKLM}(\theta_{01}) \overset{d}{\to} k_z\xi^1_{d,\phi_w} - \xi^2_{d,\phi_w} = \frac{1}{\sigma^*_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'M_{\Theta_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}), \qquad \text{(A.57)}$$

$M_{\Theta_{\phi_w}} = I - P_{\Theta_{\phi_w}}$. Using the definition of $\text{MQLR}(\theta_{01})$, we then easily have

$$\text{MQLR}(\theta_{01}) \overset{d}{\to} \frac{1}{2}[k_z\xi^1_{d,\phi_w} - \tau_m(\theta_{01})] +$$
$$\frac{1}{2}\left[\sqrt{(k_z\xi^1_{d,\phi_w} + \tau_m(\theta_{01}))^2 - 4(k_z\xi^1_{d,\phi_w} - \xi^2_{d,\phi_w})\tau_m(\theta_{01})}\right]. \qquad \text{(A.58)}$$

(B) Now, let $\Pi = \Pi_0/\sqrt{T}$ and $\Phi^1_w = \Phi^{1^0}_w/\sqrt{T}$. We can write the KLM statistic as

$$KLM(\theta_{01}) = \frac{1}{\sigma_{\varepsilon\varepsilon}(\theta_{01})}\frac{\varepsilon(\theta_{01})'Z'}{\sqrt{T}}\sqrt{T}\tilde{\Theta}(\theta_{01})\left[\sqrt{T}\tilde{\Theta}'\left(\frac{Z'Z}{T}\right)\sqrt{T}\tilde{\Theta}\right]^{-1}\sqrt{T}\tilde{\Theta}(\theta_{01})'\frac{Z\varepsilon(\theta_{01})}{\sqrt{T}},$$

where $\sqrt{T}\tilde{\Theta}(\theta_{01}) = \sqrt{T}\tilde{\Pi}(\theta_{01}) - \tilde{\Gamma}(\theta_{01})\left[\tilde{\Gamma}(\theta_{01})'\left(\frac{Z'Z}{T}\right)\tilde{\Gamma}(\theta_{01})\right]^{-1}\tilde{\Gamma}(\theta_{01})\left(\frac{Z'Z}{T}\right)\sqrt{T}\tilde{\Pi}(\theta_{01})$.
Now, we know that $\sigma_{\varepsilon\varepsilon}(\theta_{01}) \overset{p}{\to} \sigma^*_{\varepsilon\varepsilon}$, $\frac{Z\varepsilon(\theta_{01})}{\sqrt{T}} \overset{d}{\to} \mathscr{S}_d + \mathscr{S}_{\phi_w}$, $\sigma_{\varepsilon Y_i}(\theta_{01}) \overset{p}{\to} \bar{\sigma}^0_{\varepsilon i}$, $i = 1, 2$ with

$$\bar{\sigma}_{\varepsilon i} = \sigma'_{i\varepsilon} + \phi'_w\Sigma_{WY_i} - (\sigma'_{i\varepsilon} + \phi'_w\Sigma_{WY_2})\tilde{\Sigma}^{-1}_{Y_2}\Sigma_{Y_2Y_i}. \qquad \text{(A.59)}$$

From (A.59), we see that $\bar{\sigma}_{\varepsilon 2} = 0$, since $\tilde{\Sigma}_{Y_2} = \Sigma_{Y_2Y_i}$. Moreover, we have

$$\sqrt{T}\tilde{\Pi}(\theta_{01}) \overset{d}{\to} \psi_{\phi_w} = \Pi_0 + \Sigma^{-1}_Z\Sigma_{ZW}\Phi^2_{0w} + \Sigma^{-1}_Z\mathscr{S}_{z,1} - \Sigma^{-1}_Z(\mathscr{S}_d + \mathscr{S}_{\phi_w})\frac{\bar{\sigma}_{\varepsilon 1}}{\bar{\sigma}_{\varepsilon\varepsilon}}, \text{ (A.60)}$$

$$\tilde{\Gamma} \overset{p}{\to} \Gamma^0_{\phi_w} = \Gamma + \Sigma^{-1}_Z\Sigma_{ZW}\Phi^2_w, \qquad \text{(A.61)}$$

$$\sqrt{T}\tilde{\Theta}(\theta_{01}) \overset{d}{\to} \Lambda_{\phi_w} = \psi_{\phi_w} - \Gamma^0_{\phi_w}(\Gamma^{0'}_{\phi_w}\Sigma_Z\Gamma^0_{\phi_w})^{-1}\Gamma^{0'}_{\phi_w}\Sigma_Z\psi_{\phi_w}. \qquad \text{(A.62)}$$

So, we get

$$KLM(\theta_{01}) \overset{d}{\to} \xi^{*2}_{d,\phi_w} = \frac{1}{\sigma^*_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'P_{\Lambda_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}), \qquad \text{(A.63)}$$

$$JKLM(\theta_{01}) \overset{d}{\to} k_z\xi^1_{d,\phi_w} - \xi^{*2}_{d,\phi_w} = \frac{1}{\sigma^*_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mathscr{S}_{\phi_w})'M_{\Lambda_{\phi_w}}(\mathscr{S}_d + \mathscr{S}_{\phi_w}), \qquad \text{(A.64)}$$

and

$$MQLR(\theta_{01}) \overset{d}{\to} \frac{1}{2}[k_z\xi^1_{d,\phi_w} - \tau_m(\theta_{01})] +$$

33

$$\frac{1}{2}\left[\sqrt{(k_z\xi_{d,\phi_w}^1+\tau_m(\theta_{01}))^2-4(k_z\xi_{d,\phi_w}^1-\xi_{d,\phi_w}^{*2})\tau_m(\theta_{01})}\right], \quad \text{(A.65)}$$

where $P_{\Lambda_{\phi_w}}=\Lambda_{\phi_w}(\Lambda_{\phi_w}'\Lambda_{\phi_w})^{-1}\Lambda_{\phi_w}'$ and $M_{\Lambda_{\phi_w}}=I-P_{\Lambda_{\phi_w}}$. $\qquad\square$

PROOF OF COROLLARY **4.4**   Similar to the proof of Corollary **4.2**. $\qquad\square$

PROOF OF THEOREM **5.1**   Let $H_0:\theta_1=\theta_{01}$. As in Theorem **4.1**, we have

$$\varepsilon(\theta_{01})=y-Y_1\theta_{01}-Y_2\tilde{\theta}_2 \;=\; y-Y_1\theta_{01}-Y_2\theta_2-Y_2(\tilde{\theta}_2-\theta_2)$$
$$=\; u_0-Y_2(\tilde{\theta}_2-\theta_2)=(I_T-D_Y)u_0. \quad \text{(A.66)}$$

Let $\phi_w=\frac{\phi_w^0}{\sqrt{T}}$, where $\phi_w^0=\Phi_{0w}^2\theta_2$. We have

$$\frac{u_0'u_0}{T}\;=\;\frac{(\varepsilon+W\frac{\phi_w^0}{\sqrt{T}})'(\varepsilon+W\frac{\phi_w^0}{\sqrt{T}})}{T}\xrightarrow{p}\sigma_{\varepsilon\varepsilon},\quad \frac{Z'u_0}{T}=\frac{Z'(\varepsilon+W\frac{\phi_w^0}{\sqrt{T}})}{T}\xrightarrow{p}0,\quad \text{(A.67)}$$

$$\frac{u_0'Y_2}{T}\;=\;\frac{(\varepsilon+W\frac{\phi_w^0}{\sqrt{T}})'(Z\Gamma+W\frac{\Phi_{0w}^2}{\sqrt{T}}+V_2)}{T}\xrightarrow{p}\sigma_{2\varepsilon}'. \quad \text{(A.68)}$$

So, we have

$$\tilde{\theta}_2-\theta_2\xrightarrow{p}\Delta_{\theta_2}(k^*,0)=(1-k^*)\tilde{\Sigma}_{Y_2}^{-1}\sigma_{2\varepsilon}. \quad \text{(A.69)}$$

Hence, we get

$$\frac{\varepsilon(\theta_{01})'\varepsilon(\theta_{01})}{T}\xrightarrow{p}\sigma_{\varepsilon\varepsilon}^0 \;=\; \sigma_{\varepsilon\varepsilon}-2\sigma_{2\varepsilon}\Delta_{\theta_2}(k^*,0)+\Delta_{\theta_2}(k^*,0)'\Sigma_{Y_2}\Delta_{\theta_2}(k^*,0), \quad \text{(A.70)}$$

$$\frac{Z'(y-Y_1\theta_{01}-Y_2\tilde{\theta}_2)}{T}\;\xrightarrow{p}\;-\Sigma_{ZY_2}\Delta_{\theta_2}(k^*,0), \quad \text{(A.71)}$$

$$\frac{\varepsilon(\theta_{01})'P_Z\varepsilon(\theta_{01})}{T}\;\xrightarrow{p}\;\sigma_{\varepsilon,w}^0=\Delta_{\theta_2}(k^*,0)'\Sigma_{ZY_2}'\Sigma_Z^{-1}\Sigma_{ZY_2}\Delta_{\theta_2}(k^*,0), \quad \text{(A.72)}$$

$$\frac{\varepsilon(\theta_{01})'M_Z\varepsilon(\theta_{01})}{T}\;\xrightarrow{p}\;\sigma_{\varepsilon\varepsilon}^0-\sigma_{\varepsilon,w}^0\geq 0,\; k^*=\frac{\sigma_{\varepsilon\varepsilon}^0}{(\sigma_{\varepsilon\varepsilon}^0-\sigma_{\varepsilon,w}^0)}. \quad \text{(A.73)}$$

We then observe from (A.69)-(A.73) that $\Sigma_{ZY_2}\Delta_{\theta_2}(k^*,0)=0$ if and only $k^*=1$.

Suppose that $\Sigma_{ZY_2}\Delta_{\theta_2}(k^*,0) \neq 0 \Leftrightarrow k^* \neq 1$. Then, we have $\sigma^0_{\varepsilon,w} \neq 0$ and

$$\frac{\varepsilon(\theta_{01})'P_Z\varepsilon(\theta_{01})}{T} \xrightarrow{p} \sigma^0_{\varepsilon,w} \neq 0, \tag{A.74}$$

$$\varepsilon(\theta_{01})'P_Z\varepsilon(\theta_{01}) = T\frac{\varepsilon(\theta_{01})'P_Z\varepsilon(\theta_{01})}{T} \xrightarrow{p} +\infty. \tag{A.75}$$

Since the denominator of AR is $\varepsilon(\theta_{01})'M_Z\varepsilon(\theta_{01})/(T-k) \xrightarrow{p} \sigma^0_{\varepsilon\varepsilon} - \sigma^0_{\varepsilon,w} \geq 0$, we then get

$$\text{AR}(\theta_{01}) \xrightarrow{d} +\infty. \tag{A.76}$$

Suppose now that $\Sigma_{ZY_2}\Delta_{\theta_2}(k^*,0) = 0 \Leftrightarrow k^* = 1$. We can write $Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)$ as

$$\begin{aligned}
\frac{Z'(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)}{\sqrt{T}} &= \frac{Z'(I_T - D_Y)\varepsilon}{\sqrt{T}} + \frac{[Z'(I_T - D_Y)W\phi^0_w]}{T} \\
&\xrightarrow{d} \mathscr{S}_d + \mu_{\phi_w}, \tag{A.77}
\end{aligned}$$

where $\mu_{\phi_w} = \tilde{\Sigma}_{ZW}\phi^0_w$, $\tilde{\Sigma}_{ZW} = \plim_{T\to\infty} \left(\frac{Z'(I_T - D_Y)W}{T}\right) = (I_{(k_z,k_w)} - \Sigma_{ZY_2}\tilde{\Sigma}^{-1}_{Y_2}\Sigma'_{ZY_2}\Sigma^{-1}_Z)\Sigma_{ZW}$. So, we get

$$\varepsilon(\theta_{01})'P_Z\varepsilon(\theta_{01}) \xrightarrow{d} (\mathscr{S}_d + \mu_{\phi_w})'\Sigma^{-1}_Z(\mathscr{S}_d + \mu_{\phi_w}). \tag{A.78}$$

Moreover, we have

$$\frac{\varepsilon(\theta_{01})'M_Z\varepsilon(\theta_{01})}{T-k} \xrightarrow{p} \sigma^0_{\varepsilon\varepsilon} = \sigma_{\varepsilon\varepsilon}. \tag{A.79}$$

Thus

$$\text{AR}(\theta_{01}) \xrightarrow{d} \xi^1_{d,0} = \frac{1}{k_z\sigma_{\varepsilon\varepsilon}}(\mathscr{S}_d + \mu_{\phi_w})'\Sigma^{-1}_Z(\mathscr{S}_d + \mu_{\phi_w}). \tag{A.80}$$

$\square$

PROOF OF THEOREM **5.3**   The proof of Theorem **5.3** follows from those of Theorem **4.3** and Theorem **5.1**. $\square$

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.

Bean, C. , 1994. European unemployment: A survey. The Journal of Economic Literature 32(2), 573–619.

Chaudhuri, S., Zivot, E., 2010. A new method of projection based inference in GMM with weakly identified nuisance parameters. Technical report, Department of Economics, New York University N.Y.

Doko, F., Dufour, J.-M., 2006. Confidence sets for covariances between errors and endogenous regressors. Technical report, C.R.D.E., Université de Montréal.

Doko, F., Dufour, J.-M., 2008. Wald-type tests for error-regressors covariances, partial exogeneity tests and partial IV estimation. Technical report, C.R.D.E., Université de Montréal.

Doko, T. F., 2010. Inference on subset of parameters in GMM with missing instruments: Some analytical results. Technical report, School of Economics and Finance, University of Tasmania Hobart, Australia.

Donald, A., 1999. Consistent moment selection procedures for generalized method of moments estimation. Econometrica 67(3), 543–564.

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

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

Dufour, J.-M., 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., Khalaf, L., Kichian, M., 2006. Inflation dynamics and the new Keynesian Phillips curve: an identification robust econometric analysis. Journal of Economic Dynamics and Control 30(9), 1707–1727.

Dufour, J.-M., Taamouti, M., 2003. Point-optimal instruments and generalized Anderson-Rubin procedures for nonlinear models. Technical report, C.R.D.E., Université de Montréal.

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.

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

Hansen, L. P., 1985. A method for calculating bounds on the asymptotic variance-covariance matrices of the generalized method of moments estimators. Journal of Econometrics 30, 203–228.

Hansen, L. P., Heaton, J. C., Masao, O., 1988. Efficient bounds implied by multiperiod conditional moment restrictions. Journal of the American Statistical Association 83, 863–871.

Hansen, L. P., Heaton, J., Yaron, A., 1996. Finite sample properties of some alternative GMM estimators. Journal of Business and Economic Statistics 14, 262–280.

Hansen, L., Singleton, K., 1991. Computing semi-parametric efficiency bounds in linear time series models. Cambridge University Press in Barnett,W., Powell, J. and Tauchen, G (eds), Nonparametric and Semiparametric Methods in Econometrics and Statistics, 387–412.

Hayashi, F., Christopher, A. S., 1983. Nearly efficient estimation of time series models with predetermined, but not exogenous, instruments. Econometrica 51, 783–798.

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 forthcoming.

Kleibergen, F., 2008. Subset statistics in the linear IV regression model. Technical report, Department of Economics, Brown University, Providence, Rhode Island Providence, Rhode Island.

Kleibergen, F., 2009. Tests of risk premia in linear factor models. Journal of Econometrics forth-coming.

Kleibergen, F., Mavroeidis, S., 2008. Weak instrument robust tests in GMM and the new keynesian phillips curve. Technical report, Department of Economics, Brown University, Providence, Rhode Island.

Kleibergen, F., Mavroeidis, S., 2009. Inference on subsets of parameters in GMM without assuming identification. Technical report, Department of Economics, Brown University, Providence, Rhode Island.

Kocherlakota, N., 1990. On tests of representative consumer asset pricing models. Journal of Monetary Economics 26, 285–304.

Malcomson, J., Mavroeidis, S., 2006. Matching frictions, efficiency wages, and unemployment in the usa and the uk. Technical report, Department of Economics, Brown University, Providence, Rhode Island.

Mavroeidis, S., 2004. Weak identification of forward-looking models in monetary economics. Oxford Bulletin of Economics and Statistics 66, 609–635.

Mavroeidis, S., 2005. Identification issues in forward-looking models estimated by gmm with an application to the phillips curve. Journal of Money, Credit and Banking 66(3), 421–449.

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

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

Startz, S., Nelson, C., Zivot, E., 2006. Improved inference in weakly identified instrumental variables regression. *Frontiers in Analysis and Applied Research: Essays in Honor of Peter C.B. Phillips. Cambridge University Press, 2006.* .

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

West, K., 2001. On optimal instrumental variable estimation of stationary time series models. International Economic Review 42, 1043–1050.

West, K., Wilcox, D., 1996. A comparison of alternative instrumental variable estimators of a dynamic linear model. Journal of Business and Economic Statistics 14, 283–291.

School of Economics and Finance Discussion Papers

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