Econometric Theory, **31**, 2015, 1192–1228.
doi:10.1017/S0266466614000462

SUBSET HYPOTHESES TESTING AND INSTRUMENT EXCLUSION IN THE LINEAR IV REGRESSION

FIRMIN DOKO TCHATOKA
University of Tasmania

This paper explores the sensitivity of plug-in subset tests to instrument exclusion in structural models. Identification-robust statistics based on the plug-in principle have been developed for testing hypotheses specified on subsets of the structural parameters. However, their robustness to instrument exclusion has not been investigated. This paper proposes an analysis of the asymptotic distributions of the limited information maximum likelihood (LIML) estimator and plug-in statistics when potential instruments are omitted. Our results provide several new insights and extensions of earlier studies. We show that the exclusion of instruments can eliminate the first-stage, thus weakening identification and invalidating the plug-in subset inference. However, when instrument omission does not affect LIML consistency, it preserves the plug-in subset test validity, although LIML is no longer asymptotically efficient. Unlike the instrumental variable (IV) estimator, the LIML estimator of the identified linear combination of the nuisance parameter is not asymptotically a Gaussian mixture, even without instrument exclusion.

1. INTRODUCTION

Inference procedures using instrumental variable (IV) methods have received considerable attention over the last two decades. Although IV methods aim to produce consistent estimates where explanatory variables are possibly correlated with the errors, it is now well known that such methods raise identification difficulties. When the instruments are weak, IV estimators may be very imprecise, and inference procedures, such as tests and confidence sets, highly unreliable. This has led to a large body of literature aimed at producing reliable inference in the presence of weak instruments. For example, see the reviews by Stock, Wright, and Yogo (2002), Dufour (2003), and Andrews and Stock (2007).

Research on weak instruments in IV regressions has often focused on testing hypotheses specified on the full set of "structural parameters". However, testing

I am grateful to Professor Guido Kuersteiner, Co-Editor *Econometric Theory*, and three anonymous referees for constructive comments and suggestions. I would also like to thank Professors Peter C.B. Phillips, Richard J. Smith, Mardi Dungey, Jean-Marie Dufour, Lynda Khalaf, Denis Osborn, Jan Kiviet, Alistair Hall, and Jan Jacobs for several useful comments. This paper won the 2011 joint Research Award of the Reserve Bank of New Zealand and the New Zealand Econometric Study Group and I am grateful for their support. Address correspondence to Firmin Doko Tchatoka, School of Economics, The University of Adelaide, Australia; e-mail: firmin.dokotchatoka@adelaide.edu.au.

 © Cambridge University Press 2014

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

hypotheses specified on a subset of parameters may also be of interest. Subset hypotheses testing is often used to address a wide set of economic questions. This includes, but is not restricted to: (1) forward-looking models, such as the new Keynesian Phillips curve[1]; (2) stochastic discount factor models, in particular the linear factor model[2]; and (3) models of unemployment.[3]

The literature concerned with subset hypotheses testing, where instruments may be weakly correlated with the explanatory variables, falls generally into two categories. The first is the projection method based on identification-robust statistics.[4] 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. In addition to being robust to weak identification, the projection method based on the Anderson and Rubin (1949, AR) statistic also enjoys robustness to instrument exclusion. However, this method has often been criticized for being overly conservative and 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. However, there are no studies that explore the behavior of EPK when instruments are omitted.

The second category includes the robust subset procedures originally suggested by Stock and Wright (2000) and recently developed by Kleibergen (2004, 2008) and Startz, Nelson, and Zivot (2006). These procedures, known as conventional plug-in based tests, consist of replacing the nuisance parameters that are not specified by the hypothesis of interest by estimators. It is well known that plug-in based tests never over-reject the true parameter values when the nuisance parameters are identified. Guggenberger, Kleibergen, Mavroeidis, and Chen (2012) and Guggenberger and Chen (2011) show that the plug-in subset AR-test has a *correct asymptotic size* even when identification is weak, i.e., this statistic is robust to identifying assumptions, while the subset KLM-test is sensitive to such assumptions.

However, it is not clear how sensitive the plug-in principle is to instrument exclusion. 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 may be quite uncertain or "left out" of the analysis. Secondly, in many 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. A typical example of instrument exclusion problem is the conflicting results in Habib and Ljungqvist (2001) and Loughran and Ritter (2004). Both papers address the issue of endogeneity of underwriter quality when the reputation of the lead underwriter is used as an explanatory variable in an initial public offering (IPO) underpricing regression. The first-step specifications

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

differ between the two papers, which means that the instruments used in Habib and Ljungqvist (2001) are omitted in Loughran and Ritter (2004) and vice versa. As a result, in Habib and Ljungqvist's (2001) paper, the (predicted) instrument has a negative sign, whereas in Loughran and Ritter (2004), the (predicted) instrument has a positive sign. This underscores how instrumental variable regression is not always robust to alternative first-stage specification.

In this paper, we focus on the linear IV model and explore the sensitivity of the plug-in based procedures to instrument exclusion. Specifically, four plug-in based procedures are considered: the AR-test (Anderson and Rubin, 1949), the KLM-test (Kleibergen, 2002), the MQLR-test (Moreira, 2003), and the J-statistic (JKLM) that tests the misspecification of the model under the subset null hypothesis.

After formulating a general asymptotic framework which allows us to address this issue in a convenient way, we consider two main setups. In the first setup, the parameter matrix that controls the quality of the omitted instruments in the first-stage regression is fixed. In the second setup, this parameter matrix converges to zero at rate $[T^{-\frac{1}{2}}]$ when the sample size T increases (similar to Staiger and Stock, 1997). We stress the fact that analyzing the latter setup where the omitted instruments are weak is useful in capturing situations where instrument omission biases the first-stage estimates toward zero (thus weakening identification) but does not completely eliminate the first-stage. In both setups, the parameter matrix that controls the quality of the included instruments can be fixed with an arbitrary rank or local to zero. By allowing this parameter matrix to have an arbitrary rank, we extend earlier results in the partial identified model by Phillips (1989) and Choi and Phillips (1992) to limited information maximum likelihood (LIML) estimators.

We propose an analysis of the asymptotic distributions of the restricted LIML estimator, as well as the plug-in subset statistics. Our results provide several new insights and extensions of earlier studies. We show that exclusion of instruments, even if it is not coupled with weak instruments, can eliminate the first stage, thus weakening identification and invalidating the plug-in subset inference. However, the exclusion of instruments alone, if it does not affect the consistency of LIML, preserves the plug-in subset inference even though LIML is no longer asymptotically efficient. In addition, our results indicate that the limiting distribution of the LIML estimator of the identified linear combination of the nuisance parameters is not necessarily a Gaussian mixture, with or without instrument exclusion. This contrasts with the identified linear combination of the usual IV estimator, which is still asymptotically Gaussian mixture, despite instrument exclusion (similar to Choi and Phillips, 1992).[5] We present a Monte Carlo experiment which confirms our theoretical findings.

In the remainder of the paper, Section 2 formulates the model, while Section 3 studies the asymptotic distribution of the restricted LIML estimator, as well as the subset statistics. Conclusions are drawn in Section 4 and proofs are presented in the Appendix.

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

Throughout this paper, I_q stands for the identity matrix of order q. For any full-column rank $n \times m$ matrix A, $P_A = A(A'A)^{-1}A'$ is the projection matrix on the space spanned by A, and $M_A = I_n - P_A$. The notation $vec(A)$ is the $nm \times 1$ dimensional column vectorization of A. $B > 0$ for a squared matrix B means that B is positive definite (p.d.) and $|B|$ is the determinant of B. The convergence in probability is symbolized by "$\xrightarrow{p}$" while "$\xrightarrow{d}$" stands for convergence in distribution. The usual (stochastic) orders of magnitude are denoted by $O_p(.)$ and $o_p(.)$. $\|U\| = [tr(U'U)]^{\frac{1}{2}}$ denotes the usual Euclidean or Frobenius norm for a matrix U, while $rank(U)$ is the rank of U. The symbol "$\overset{d}{=}$" signifies equivalence in distribution, and, finally, $\mathcal{O}(n) = \{H \in \mathbb{M}(n,n) : H'H = I_n\}$ denotes the orthogonal group of $n \times n$ matrices, where $\mathbb{M}(n,n)$ is the set of all squared matrices of order n.

2. FRAMEWORK

We consider the simplified structural equation of the form:

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

where y is a $T \times 1$ vector of observations on a dependent variable, Y_1 and Y_2 are $T \times m_1$ and $T \times m_2$ matrices of (supposedly) endogenous explanatory variables ($m = m_1 + m_2 \geq 1$), $\varepsilon = [\varepsilon_1, \ldots, \varepsilon_T]'$ is a vector of structural disturbances, θ_1 and θ_2 are $m_1 \times 1$ and $m_2 \times 1$ unknown coefficient vectors. Further, we assume that Y_1 and Y_2 satisfy the model:

$$Y_1 = Z\Pi_{21} + V_1^*,\ Y_2 = Z\Pi_{22} + V_2^*, \tag{2.2}$$

where Z is a $T \times l$ matrix of observations on exogenous variables (instruments), Π_{21} and Π_{22} are $l \times m_1$ and $l \times m_2$ coefficient matrices, and $V^* = [V_1^*, V_2^*] = [v_1^*, \ldots, v_T^*]'$ is a $T \times m$ matrix of reduced form disturbances with $V_1^* : T \times m_1$ and $V_2^* : T \times m_2$. It is assumed that the instrument matrix Z has full-column rank l with probability one.[6]

Without loss of generality, let θ_1 be the parameter vector of interest. We are concerned with the sensitivity to instrument exclusion of plug-in subset AR, KLM, JKLM, and MQLR tests[7] for the null hypothesis

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

where θ_{01} is a $m_1 \times 1$ fixed vector. More precisely, we investigate the asymptotic behavior of the plug-in subset tests when the "true" data generating processes for Y_1 and Y_2 are given by:

$$Y_1 = Z\Pi_{21} + W\Phi_1 + V_1,\ Y_2 = Z\Pi_{22} + W\Phi_2 + V_2, \tag{2.4}$$

but model (2.2) is used for inference; where W is a $T \times l_1$ matrix of instruments that is not accounted for, $\Phi_1 \in \mathbb{R}^{l_1 \times m_1}$ and $\Phi_2 \in \mathbb{R}^{l_1 \times m_2}$ are unknown coefficient

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

matrices, and $V = [V_1, V_2] = [v_1, \ldots, v_T]'$ is a $T \times m$ matrix of first-stage errors. Let $Y = [Y_1 : Y_2]$ denote the matrix of endogenous regressors and define $\Pi_2 = [\Pi_{21} : \Pi_{22}]$, $\Phi = [\Phi_1 : \Phi_2]$. The first-stage regression (2.4) can then be expressed as:

$$Y = Z\Pi_2 + W\Phi + V. \tag{2.5}$$

The restricted LIML estimator of θ_2 under H_0 in model (2.1) and (2.2) where W is excluded can be expressed (Startz et al., 2006) as:

$$\tilde{\theta}_2 = [Y_2'(P_Z - \tilde{\kappa}M_Z)Y_2]^{-1}Y_2'(P_Z - \tilde{\kappa}M_Z)(y - Y_1\theta_{01}), \tag{2.6}$$

where $\tilde{\kappa} = \tilde{\kappa}(\tilde{\theta}_2)$ is the smallest root of the determinantal equation $|\bar{Y}_0'P_Z\bar{Y}_0 - \tilde{k}\bar{Y}_0'M_Z\bar{Y}_0| = 0$ and $\bar{Y}_0 = [y - Y_1\theta_{01} : Y_2]$. All subset tests considered in this study use $\tilde{\theta}_2$ as the plug-in estimator.

The usual assumptions on the joint asymptotic behavior of the errors and the instruments are made:

Assumption 2.1. When the sample size T converges to infinity, the following convergence results hold jointly:

$AS1$. $T^{-1}[\varepsilon : V]'[\varepsilon : V] \xrightarrow{p} \Sigma$, $T^{-1}[Z : W]'[\varepsilon : V] \xrightarrow{p} 0$; $T^{-1}[Z : W]'[Z : W] \xrightarrow{p} Q$,

where $\Sigma = \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{V_1\varepsilon}' & \sigma_{V_2\varepsilon}' \\ \sigma_{V_1\varepsilon} & \Sigma_{V_1} & \sigma_{V_2V_1}' \\ \sigma_{V_2\varepsilon} & \sigma_{V_2V_1} & \Sigma_{V_2} \end{pmatrix} > 0$, $Q = \begin{pmatrix} Q_Z & Q_{ZW} \\ Q_{ZW}' & Q_W \end{pmatrix}$, $Q_Z > 0$, $Q_W \geq 0$; $\sigma_\varepsilon^2 : 1 \times 1$, $\sigma_{V_1\varepsilon} : m_1 \times 1$, $\sigma_{V_2\varepsilon} : m_2 \times 1$, $\Sigma_{V_1} : m_1 \times m_1$, $\Sigma_{V_2V_1} : m_2 \times m_1$, $\Sigma_{V_2} : m_2 \times m_2$;

$AS2$. $T^{-1/2}Z'[\varepsilon : V] \xrightarrow{d} [\psi_{Z\varepsilon} : \psi_{ZV_1} : \psi_{ZV_2}]$, where $\psi_{Z\varepsilon} : l \times 1$, $\psi_{ZV_1} : l \times m_1$, $\psi_{ZV_2} : l \times m_2$, and $vec([\psi_{Z\varepsilon} : \psi_{ZV_1} : \psi_{ZV_2}]) \sim N(0, \Sigma \otimes Q_Z)$.

The convergence in assumptions $AS1$ is the weak law of large numbers (WLLN) property. It is assumed that the IVs and errors are asymptotically uncorrelated. So, instrument endogeneity (such as in Kiviet and Niemczyk, 2007, 2012; Doko Tchatoka and Dufour, 2008; Guggenberger, 2011; Berkowitz, Caner, and Fang, 2008, 2012) is not an issue. Assumption $AS2$ holds by the central limit theorem (CLT) property (see Staiger and Stock, 1997; Kleibergen, 2002, 2004; Kleibergen and Mavroeidis, 2011 for similar assumptions). It is required that $rank(Z) = l$ with probability one, but no such restriction is imposed on W, i.e., some linear combinations of the columns of W may be dependent or close to being so. We now wish to discuss the link between instrument exclusion and model identification.

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

2.1. Instrument Omission and Parameter Identification

The weak IV literature[8] documents that the high-level assumption for the identification[9] of the nuisance parameters θ_2 under the subset null hypothesis H_0 in model (2.1) and (2.2) is:

$$rank(\Pi_{22}) = m_2. \tag{2.7}$$

However, this identifying high-level assumption may not hold if potential candidate instruments are omitted from the first-stage specification (2.2). In this paper, we stress the fact the exclusion of instruments can eliminate the first-stage, thus weakening identification and invalidating the plug-in subset inference. Indeed, the consistency of the restricted LIML estimator given in (2.6) depends on the rank of the matrix $Q_{ZY_2} = p\lim_{T\to\infty}\left(T^{-1}Z'Y_2\right)$. Under (2.1)–(2.5) and Assumption 2.1, we have

$$Q_{ZY_2} = p\lim_{T\to\infty}\left(T^{-1}Z'Y_2\right) = Q_Z\Pi_{22} + Q_{ZW}\Phi_2, \tag{2.8}$$

where the last term $Q_{ZW}\Phi_2$ is a measure of the bias induced by the omission of W from the first-stage specification of Y_2. If no relevant instrument is omitted in the specification of Y_2, i.e., if $\Phi_2 = 0$, then it is obvious from (2.8) that $rank(Q_{ZY_2}) = m_2$ if, and only if, $rank(\Pi_{22}) = m_2$; and (2.7) is a necessary and sufficient condition for the identification of θ_2 under H_0. However, if at least one relevant IV is omitted ($\Phi_2 \neq 0$) in the first-stage specification of Y_2, it is possible that $rank(Q_{ZY_2}) < m_2$ even if $rank(\Pi_{22}) = m_2$. Clearly, omitting instruments can lead to biased parameter estimates in the first-stage regression. It is then possible that this bias leads to estimated $\widehat{\Pi}_{22}$ that has a deficient rank, thus affecting the consistency of LIML. Although it is not clear how relevant this scenario is in practice, analyzing cases where the bias of these estimates is close to zero may be useful to capture situations where instrument exclusion biases the first-stage estimates toward zero (thus weakening identification) but does not completely eliminate the first-stage.

Equation (2.8) shows how inference procedures which exploit the first-stage specification of Y_2—such as the restricted LIML estimator in (2.6) and the plug-in subset tests in Section 2.2—can be affected by relevant instrument omission. Dufour Taamouti (2007)[10] show that if inference is performed on the entire vector $\theta = (\theta_1', \theta_2')'$, the AR-test is robust to instrument exclusion, i.e., it never overrejects the null hypothesis even when strong instruments are omitted. Meanwhile, Kleibergen's (2002) score and Moreira's (2003) likelihood ratio tests do not share this property. One can then assess any subset hypothesis such as H_0 in (2.3) by resorting to the method of projection[11] based on inverting the AR-statistic which is robust to instrument exclusion. In spite of its attractiveness in terms of level control and robustness to IV exclusion, the method of projection-type inference can often be too conservative; see Moreira (2003) and Startz et al. (2006). Kleibergen and Mavroeidis (2011) show that the plug-in method yields more powerful tests than their projection-based counterpart. Recently, Guggenberger et al. (2012) and

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

Guggenberger and Chen (2011) show that the subset AR-test has a *correct asymptotic size* even when identification is weak, i.e., this statistic is robust to identifying assumptions, while the subset KLM-test is sensitive to such assumptions. However, the question of when and to what extent instrument exclusion is detrimental to the plug-in subset inference remains open, and this paper attempts to make progress in this direction.

2.2. The Plug-in Subset Inference

The plug-in principle consists of two steps. The first step is to take an identification-robust statistic [usually the Anderson and Rubin (1949, AR), Kleibergen (2002, K), and Moreira (2003, CLR) type statistics] which results from the test of the joint hypothesis $\mathrm{H}(\theta_{01}, \theta_{02}) : \theta_1 = \theta_{01}, \theta_2 = \theta_{02}$. The second step consists of replacing the nuisance parameters θ_2 which are not specified by the subset null hypothesis of interest (2.3) by an estimator in the expression of the above statistics. In this paper, we consider four plug-in statistics, namely AR, KLM, JKLM, and MQLR that use the restricted LIML estimator $\tilde{\theta}_2$ given in (2.6). These statistics are computed from model (2.1) and (2.2), where W is omitted from the variables that determine Y, as follows:

(a) the AR subset statistic to test $\mathrm{H}_0 : \theta_1 = \theta_{01}$ reads

$$\mathrm{AR}(\theta_{01}) = \frac{1}{l\tilde{\sigma}_\varepsilon^2(\theta_{01})}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)' P_Z (y - Y_1\theta_{01} - Y_2\tilde{\theta}_2), \tag{2.9}$$

where $\tilde{\sigma}_\varepsilon^2(\theta_{01}) = \frac{1}{T-l}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)' M_Z (y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)$;

(b) Kleibergen's (2002) Lagrange multiplier (KLM) statistic to test $\mathrm{H}_0 : \theta_1 = \theta_{01}$ reads (see Kleibergen, 2004)

$$\mathrm{KLM}(\theta_{01}) = \frac{1}{\tilde{\sigma}_\varepsilon^2(\theta_{01})}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)' P_{\tilde{Z}} (y - Y_1\theta_{01} - Y_2\tilde{\theta}_2), \tag{2.10}$$

where

$$\tilde{Z} = M_{Z\tilde{\Pi}_{22}} \times Z\tilde{\Pi}_{21}, \tag{2.11}$$

$$\tilde{\Pi}_{21} = \tilde{\Pi}_{21}(\theta_{01}) = (Z'Z)^{-1}Z'\left[Y_1 - (y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)\frac{\tilde{\sigma}_{\varepsilon Y_1}(\theta_{01})}{\tilde{\sigma}_\varepsilon^2(\theta_{01})}\right], \tag{2.12}$$

$$\tilde{\Pi}_{22} = \tilde{\Pi}_{22}(\theta_{01}) = (Z'Z)^{-1}Z'\left[Y_2 - (y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)\frac{\tilde{\sigma}_{\varepsilon Y_2}(\theta_{01})}{\tilde{\sigma}_\varepsilon^2(\theta_{01})}\right], \tag{2.13}$$

and $\tilde{\sigma}_{\varepsilon Y_i}(\theta_{01}) = \frac{1}{T-l}(y - Y_1\theta_{01} - Y_2\tilde{\theta}_2)' M_Z Y_i,\ i = 1, 2$;

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

$$\mathrm{JKLM}(\theta_{01}) = l\,\mathrm{AR}(\theta_{01}) - \mathrm{KLM}(\theta_{01}); \tag{2.14}$$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

(d) and finally the subset extension of the conditional likelihood ratio statistic to test $H_0 : \theta_1 = \theta_{01}$ reads (see Moreira, 2003)

$$\begin{aligned} MQLR(\theta_{01}) = &\frac{1}{2}(lAR(\theta_{01}) - \tau_m) \\ &+ \frac{1}{2}\sqrt{[lAR(\theta_{01}) + \tau_m]^2 - 4[lAR(\theta_{01}) - KLM(\theta_{01})]\tau_m}], \end{aligned} \tag{2.15}$$

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

$$\hat{\Sigma}(\theta_{01}) = \hat{\Sigma} = \hat{\mathbb{T}}'\hat{\mathbb{T}}, \tag{2.16}$$

$$\begin{aligned} \hat{\mathbb{T}} = \hat{\mathbb{T}}(\theta_{01}) &= (Z'Z)^{\frac{1}{2}}[\tilde{\Pi}_{21} \vdots \tilde{\Pi}_{22}]\hat{\Sigma}_{Y\varepsilon}^{-1/2}, \\ \hat{\Sigma}_{Y\varepsilon}^{-1/2} &= \begin{bmatrix} \hat{\Sigma}_{1,\varepsilon}^{-1/2} & 0 \\ -\hat{\Sigma}_{2\varepsilon}^{-1}\hat{\Sigma}_{21,\varepsilon}\hat{\Sigma}_{1,\varepsilon}^{-1/2} & \hat{\Sigma}_{2\varepsilon}^{-1/2} \end{bmatrix}, \\ \hat{\Sigma}_{1,\varepsilon} &= \frac{1}{T-l}Y_1'M_{(Z:Y_2:\tilde{\varepsilon})}Y_1,\ \hat{\Sigma}_{21,\varepsilon} = \frac{1}{T-l}Y_2'M_{(Z:\tilde{\varepsilon})}Y_1, \\ \hat{\Sigma}_{2\varepsilon} &= \frac{1}{T-l}Y_2'M_{(Z:\tilde{\varepsilon})}Y_2,\ \tilde{\varepsilon} = y - Y_1\theta_{01} - Y_2\tilde{\theta}_2. \end{aligned} \tag{2.17}$$

Lemma 2.2 summarizes the asymptotic distributions of the plug-in statistics from the previous literature under the assumption that all potential candidate instruments have been used within inference.

LEMMA 2.2. *Suppose Assumption* 2.1 *and* H_0 *are satisfied. Suppose also that no instrument is omitted from* (2.2). *Then*:

(a) *if* Π_{22} *is fixed and has full-column rank* m_2, *we have*

$$\begin{aligned} &l \times AR(\theta_{01}) \xrightarrow{d} \chi_{m_1} + \chi_{l-m},\ KLM(\theta_{01}) \xrightarrow{d} \chi_{m_1},\ JKLM(\theta_{01}) \xrightarrow{d} \chi_{l-m}, \\ &MQLR(\theta_{01})|_{\tilde{\Pi}_{21},\tilde{\Pi}_{22}} \xrightarrow{d} \frac{1}{2}(\chi_{m_1} + \chi_{l-m} - \tau_m) \\ &\qquad + \frac{1}{2}\sqrt{(\chi_{m_1} + \chi_{l-m} + \tau_m)^2 - 4\chi_{l-m}\tau_m}, \end{aligned}$$

where χ_{m_1} *and* χ_{l-m} *are independent* $\chi^2(m_1)$ *and* $\chi^2(l-m)$ *distributed random variables*;

(b) *if* Π_{22} *has a weak value, i.e., for example, if* Π_{22} *is fixed with a deficient rank or if* $\Pi_{22} = C/\sqrt{T}$ *where* C *is a* $l \times m_2$ *constant matrix* (*possibly zero*), *the chi-square limiting distributions of the subset AR-statistic in* (*a*) *provide an upper bound on its actual limiting distribution. But this result does not directly generalize to the other subset statistics.*

The proof of (a) is given in Stock and Wright (2000) and Startz et al. (2006) for the subset AR statistic; Kleibergen (2004) and Mikusheva (2010) for all other

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

statistics. The proof of of (b) is given in Guggenberger et al. (2012) and Guggenberger and Chen (2011). Lemma 2.2(a) states that with no instrument exclusion, all subset statistics are asymptotically pivotal if the high level assumption of the identification of θ_2 is satisfied. This means that all plug-in subset tests have an asymptotically correct level under H_0. Lemma 2.2(b) extends the validity of the plug-in subset AR-statistic to weakly identified setups. Guggenberger et al. (2012, eqn. 59) also provide the analytical characterization of the asymptotic distribution of the subset KLM-statistic. However, there are no such characterizations of the limiting distributions of the subset JKLM and MQLR statistics even when $\Phi = 0$ (no instrument exclusion), especially when partial identification holds (i.e., when Π_{22} is fixed with $rank(\Pi_{22}) < m_2$). One of our contributions to the literature is the provision of an explicit characterization of the limiting distributions of all subset statistics under all possible scenarios of identification, with or without instrument omission.

3. TEST SENSITIVITY TO INSTRUMENT OMISSION

To analyze the impact of instrument exclusion on the subsets tests, we find it useful to distinguish two main setups. In the first setup, $\Phi = [\Phi_1, \Phi_2]$ is fixed such that $rank(\Pi_{2W}) = p_2 \leq m_2$, where $\Pi_{2W} = \Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2$. In the second one, we consider sequences $\Phi = [D_1/\sqrt{T}, \Phi_2^* + D_2/\sqrt{T}]$ such that the rank of Π_{2W} is local to rank deficient, i.e., $\lim_{T\to+\infty} rank(\Pi_{2W}) = rank(\Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2^*) = p_2 < m_2$ for some fixed matrices $\Phi_2^* \in \mathbb{R}^{l_1 \times m_2}$, $D_1 \in \mathbb{R}^{l_1 \times m_1}$, and $D_2 \in \mathbb{R}^{l_1 \times m_2}$ (possibly zeros). Because the parameter matrix Φ that controls the quality of the excluded instruments W in the first-stage regression (2.5) is fixed (does not depend on the sample size T), we shall call this setup "*fixed instrument exclusion*". Note that the parameter $\Pi_2 = \left[\Pi_{21}, \Pi_{22}\right]$ that controls the quality of the included instruments Z may have any arbitrary rank. In particular, Π_2 can be zero or close to being so. Clearly, this setup allows for cases where the included and excluded instruments imply different behaviors. The second setup expands the results when Π_{2W} is fixed with a deficient rank p_2 to cases where Π_{2W} is local to rank deficient, i.e., the rank of Π_{2W} is not exactly p_2 but approaches p_2 as T increases. We stress the fact that analyzing this latter setup is important to establish that the results of the first setup, when Π_{2W} is fixed with $rank(\Pi_{2W}) < m_2$, do not depend on a specific value of the rank of Π_{2W} in the neighborhood of rank deficiency. So, our results remain the same whether Π_{2W} is fixed with $rank(\Pi_{2W}) < m_2$ or Π_{2W} is local to rank deficient such that $\Phi = [D_1/\sqrt{T}, \Phi_2^* + D_2/\sqrt{T}]$. Because the sequences $\Phi = [D_1/\sqrt{T}, \Phi_2^* + D_2/\sqrt{T}]$ include the local to zero weak omitted instruments (similar to Staiger and Stock, 1997), we shall call this setup "*possibly local to zero weak instrument exclusion*". It is worthwhile mentioning that, in spite of Π_{2W} being local to rank deficient, the included and excluded instruments can imply different behaviors as in the first setup.

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

It is worthwhile noting that the results will not change qualitatively when considering sequences of the form $\Phi = [\Phi_1^* + D_1/\sqrt{T},\ \Phi_2^* + D_2/\sqrt{T}]$ for some fixed matrices $\Phi_1^* \in \mathbb{R}^{l_1 \times m_1}$ and $\Phi_2^* \in \mathbb{R}^{l_1 \times m_2}$. This is because the key findings are driven by the rank condition on the matrix Π_{2W} which does not involve Φ_1 (for example, see the discussion below Theorems 3.2 and 3.3). So, without loss of generality, we can limit the analysis to sequences $\Phi = [D_1/\sqrt{T},\ \Phi_2^* + D_2/\sqrt{T}]$. However, note that except in the case of the subset AR-statistic, the expressions of the limiting distributions of the other statistics also depend on Φ_1 through the score factor. Nevertheless, these expressions can be easily modified in accordance with $\Phi = [\Phi_1^* + D_1/\sqrt{T},\ \Phi_2^* + D_2/\sqrt{T}]$ when necessary.

Section 3.1 presents the results in the *fixed instrument exclusion* setup, while Section 3.2 deals with *possibly local to zero weak instrument exclusion*. Finally, Section 3.3 illustrates the behavior of the tests through a Monte Carlo experiment.

3.1. Fixed Instrument Exclusion

In this section, we study the situations where Φ is fixed and may have an arbitrary rank. Since all plug-in subset statistics considered use the restricted LIML estimator, $\tilde{\theta}_2$, we find it useful to investigate the asymptotic behavior of $\tilde{\theta}_2$ first. Let $S = [S_1 : S_2] \in \mathcal{O}(m_2)$[12] such that S_2 spans the null space of $[\Pi_{22}' : \Phi_2']'$ (see Section B.1 of the Supplemental Appendix for further details). Also, let κ denote the smallest root of the determinantal equation

$$\left| \Xi_{\Phi\Pi}^0 - \kappa \Lambda \right| = 0, \tag{3.1}$$

where $\Xi_{\Phi\Pi}^0$ and Λ are given by

$$\Xi_{\Phi\Pi}^0 = \begin{bmatrix} \psi_{Z\varepsilon}' Q_Z^{-1} \psi_{Z\varepsilon} & \psi_{Z\varepsilon}' \Pi_{2W} S_1 & \psi_{Z\varepsilon}' Q_Z^{-1} \psi_{ZV_2} S_2 \\ S_1' \Pi_{2W}' \psi_{Z\varepsilon} & S_1' \Pi_{2W}' Q_Z \Pi_{2W} S_1 & S_1' \Pi_{2W}' \psi_{ZV_2} S_2 \\ S_2' \psi_{ZV_2}' Q_Z^{-1} \psi_{Z\varepsilon} & S_2' \psi_{ZV_2}' \Pi_{2W} S_1 & S_2' \psi_{ZV_2}' Q_Z^{-1} \psi_{ZV_2} S_2 \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \sigma_{V_2\varepsilon}' S_2 \\ 0 & 0 & 0 \\ S_2' \sigma_{V_2\varepsilon} & 0 & S_2' \Sigma_{V_2} S_2 \end{bmatrix}$$

with $\Pi_{2W} = \Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2$; the other variables are defined in Assumption 2.1. Define $\Upsilon_T = diag(\sqrt{T} I_{p_2}, I_{m_2 - p_2})$ where I_{p_2} and $I_{m_2 - p_2}$ are identity matrices of dimensions $p_2 = rank(\Pi_{2W})$ and $m_2 - p_2$, respectively. We can now state the following theorem on the asymptotic distribution of $\tilde{\theta}_2$.

THEOREM 3.1. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [\Phi_1, \Phi_2]$ *is fixed. If further* H_0 *holds, then we have*:

(a) $\sqrt{T}(\tilde{\theta}_2 - \theta_2) \xrightarrow{d} \Delta_2 \sim N(0, \sigma_\varepsilon^2 \Omega_2^{-1})$ *when* $p_2 = m_2$, *where* $\Omega_2 = \Pi_{2W}' Q_Z \Pi_{2W}$; *and*

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

(b) $\Upsilon_T S'(\tilde{\theta}_2 - \theta_2) = \begin{pmatrix} \sqrt{T}(\tilde{\theta}_{21} - \theta_{21}) \\ \tilde{\theta}_{22} - \theta_{22} \end{pmatrix} \xrightarrow{d} \theta_2^*(\kappa) = \Xi_2^{-1}(\kappa)\zeta_2^*(\kappa)$ *for all values of* Π_2, *when* $p_2 < m_2$, *where*

$$\Xi_2(\kappa) = \begin{bmatrix} S_1'\Pi_{2W}' Q_Z \Pi_{2W} S_1 & S_1'\Pi_{2W}'\psi_{ZV_2} S_2 \\ S_2'\psi_{ZV_2}'\Pi_{2W} S_1 & S_2'(\psi_{ZV_2}' Q_Z^{-1}\psi_{ZV_2} - \kappa\Sigma_{V_2})S_2 \end{bmatrix},$$

$$\zeta_2^*(\kappa) = \begin{pmatrix} S_1'\Pi_{2W}'\psi_{Z\varepsilon} \\ S_2'(\psi_{ZV_2}' Q_Z^{-1}\psi_{Z\varepsilon} - \kappa\sigma_{V_2\varepsilon}) \end{pmatrix},$$

κ *and* S_j $(j = 1, 2)$ *are defined in* (3.1).

Theorem 3.1(a) shows that exclusion of instruments alone, if it does not weaken identification, preserves LIML consistency. However, the limiting distribution of LIML is affected through the covariance matrix Ω_2. Indeed, we can decompose Ω_2 as:

$$\Omega_2 = \Pi_{22}' Q_Z \Pi_{22} + \Omega_{\Phi\Pi} + \Omega_\Phi \quad \text{where}$$

$$\Omega_{\Phi\Pi} = \Phi_2' Q_{ZW}' \Pi_{22} + \Pi_{22}' Q_{ZW} \Phi_2, \quad \Omega_\Phi = \Phi_2' Q_{ZW}' Q_Z^{-1} Q_{ZW} \Phi_2. \tag{3.2}$$

Note that $\sigma_\varepsilon^2 \Pi_{22}' Q_Z \Pi_{22}$ is the usual asymptotic covariance of LIML under strong identification and no instrument is exclusion. The two other terms, $\Omega_{\Phi\Pi}$ and Ω_Φ, are present because of the exclusion of W. This decomposition shows clearly that exclusion of instruments can bias the covariance matrix estimate of LIML, thus affecting its asymptotic efficiency, although consistency still holds.

Theorem 3.1(b) can be viewed as a generalization of the usual *fixed instrument asymptotics*.[13] The parameter matrix Π_2 in the first-stage regression (2.5) may be fixed with an arbitrary rank. By allowing Π_2 and Φ to have any arbitrary rank, Theorem 3.1 extends the partial identification setup in Phillips (1989) and Choi and Phillips (1992) to LIML. Most studies of weak instrument effects on the asymptotic behavior of LIML usually consider a Staiger and Stock (1997) setup, which basically excludes partial identification. Moreover, even with no instrument omission, Theorem 3.1(b) confirms[14] that lack of identifiability is detrimental to the asymptotic normality of LIML. Even the asymptotic distribution of the identified linear combination of the LIML estimator, $\sqrt{T}(\tilde{\theta}_{21} - \theta_{21})$, is not necessarily a Gaussian mixture as indicated by the expression of $\theta_2^*(\kappa)$. This contrasts with the usual IV estimator of the identified linear combination of θ_2, which converges to a Gaussian mixture process (see Choi and Phillips, 1992), despite the exclusion of potential relevant IVs.

As argued in Section 2.1, it is possible that instrument exclusion alone, even if it not coupled with weak identification, leads to the result of Theorem 3.1(b). Although it is difficult to explicitly quantify the effects of instrument exclusion on LIML from Theorem 3.1(b), one can analyze the impact of instrument exclusion by considering the following decomposition of $\Xi_2(\kappa)$ and $\zeta_2^*(\kappa)$ [in the expression of $\theta_2^*(\kappa)$]:

$$\Xi_2(\kappa) = \Xi_\Pi(\kappa) + \Xi_{\Phi\Pi}(\kappa), \zeta_2^*(\kappa) = \zeta_{2,\Pi}^*(\kappa) + \zeta_{2,\Phi\Pi}^*(\kappa) \quad \text{where} \tag{3.3}$$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

$$\Xi_{\Pi}(\kappa)=\begin{bmatrix} S_1'\Pi_{22}'Q_Z\Pi_{22}S_1 & S_1'\Pi_{22}'\psi_{ZV_2}S_2 \\ S_2'\psi_{ZV_2}'\Pi_{22}S_1 & S_2'(\psi_{ZV_2}'Q_Z^{-1}\psi_{ZV_2}-\kappa\Sigma_{V_2})S_2 \end{bmatrix},$$

$$\zeta_{2,\Pi}^{*}(\kappa)=\begin{pmatrix} S_1'\Pi_{22}'\psi_{Z\varepsilon} \\ S_2'(\psi_{ZV_2}'Q_Z^{-1}\psi_{Z\varepsilon}-\kappa\sigma_{V_2\varepsilon}) \end{pmatrix}$$

$$\Xi_{\Phi\Pi}(\kappa)=\begin{bmatrix} S_1'(\Omega_{\Phi\Pi}+\Omega_{\Phi})S_1 & S_1'\Phi_2'Q_{ZW}'Q_Z^{-1}\psi_{ZV_2}S_2 \\ S_2'\psi_{ZV_2}'Q_Z^{-1}Q_{ZW}\Phi_2S_1 & 0 \end{bmatrix},$$

$$\zeta_{2,\Phi\Pi}^{*}(\kappa)=\begin{pmatrix} S_1'\Phi_2'Q_{ZW}'Q_Z^{-1}\psi_{Z\varepsilon} \\ 0 \end{pmatrix}.$$

The presence in (3.3) of the two terms, $\Xi_{\Phi\Pi}(\kappa)$ and $\zeta_{2,\Phi\Pi}^{*}(\kappa)$, is due to instrument omission. When $\Phi_2\neq 0$, we have $\Xi_2(\kappa)\neq 0$ and $\zeta_2^{*}(\kappa)\neq 0$ with probability one, hence omitting relevant instruments can dramatically affect the asymptotic distribution of LIML. There is however no clear direction of this effect.

We can now state the following theorem on the asymptotic behavior of the subset statistics when $p_2=m_2$.

THEOREM 3.2. *Suppose Assumption* 2.1 *is satisfied and* $\Phi=[\Phi_1,\Phi_2]$ *is fixed. If further* H_0 *holds and* $p_2=m_2$, *then we have*:

$$AR(\theta_{01})\xrightarrow{d}\xi_1=\frac{1}{l}\parallel z_{w\varepsilon}\parallel^2,\quad KLM(\theta_{01})\xrightarrow{d}\xi_2=\parallel P_{\Omega(r_1)}.z_{w\varepsilon}\parallel^2,$$

$$JKLM(\theta_{01})\xrightarrow{d}l\xi_1-\xi_2,$$

$$MQLR(\theta_{01})|_{\tau_m}\xrightarrow{d}\frac{1}{2}(l\xi_1-\tau_m)+\frac{1}{2}\sqrt{(l\xi_1+\tau_m)^2-4[l\xi_1-\xi_2]\tau_m},$$

where ξ_2 *and* $l\xi_1-\xi_2$ *are independent* $\chi^2(m_1)$ *and* $\chi^2(l-m)$ *distributed random variables,* $z_{w\varepsilon}=M_{Q_Z^{1/2}\Pi_{2W}}.z_{\varepsilon}$, $z_{\varepsilon}=\sigma_{\varepsilon}^{-1}Q_Z^{-1/2}\psi_{Z\varepsilon}$, $\Omega(r_1)=M_{Q_Z^{1/2}\Pi_{2W}}.Q_Z^{1/2}\Pi_{1W}$ *if* $r_1=rank(\Pi_{1W})=m_1$ *and* $\Omega(r_1)=M_{Q_Z^{1/2}\Pi_{2W}}.Q_Z^{1/2}\Psi_{1R}$ *if* $r_1<m_1$, $\Pi_{1W}=\Pi_{21}+Q_Z^{-1}Q_{ZW}\Phi_1$, $\Psi_{1R}=[\Pi_{1W}R_1,\ Q_Z^{-1}\psi_{ZV_1}R_2-\sigma_{\varepsilon}^{-1}z_{w\varepsilon}\sigma_{V_1\varepsilon}'R_2]$, *and* $R=[R_1:R_2]\in\mathcal{O}(m_1)$ *is such that* R_2 *spans the null space of* $[\Pi_{21}':\Phi_1']'$.

Although the exclusion of instruments affects the asymptotic efficiency of LIML, Theorem 3.2 shows that the plug-in subset inference validity is preserved if LIML consistency holds ($p_2=m_2$). Unlike Lemma 2.2(a), Theorem 3.2 holds for all values of Π_2 and Φ, provided $rank(\Pi_{2W})=m_2$ where $\Pi_{2W}=\Pi_{22}+Q_Z^{-1}Q_{ZW}\Phi_2$. Clearly, Theorem 3.2 captures situations where instrument exclusion biases the first stage estimates but does not weaken identification, thus preserving the plug-in subset inference. If $\Phi_2=0$ (no excluded IV in the first-stage regression of Y_2), $\Pi_{2W}=\Pi_{22}$ and the rank condition of Theorem 3.2 is exactly the same as Lemma 2.2(a), although we may still have instrument

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

exclusion in the first-stage regression of Y_1 ($\Phi_1 \neq 0$). So, if $p_2 = m_2$, the exclusion of instrument matters only in the first-stage specification of Y_2. This result is intuitive for AR(θ_{01}). Indeed, AR(θ_{01}) depends only on $y - Y_1\theta_1 - Y_2\tilde{\theta}_2$ and the included IVs Z, but under H_0, $y - Y_1\theta_1 - Y_2\tilde{\theta}_2 = Y_2(\theta_2 - \tilde{\theta}_2) + \varepsilon$ does not involve Y_1 or the first-stage specification of Y_1. However, the results are less intuitive for the three other statistics. This is because all three statistics, unlike the AR-statistic, exploit the first-stage regression of Y_1 through the score vectors in (2.11)–(2.13). They are asymptotically pivotal because the score factor $\Omega(r_1)$ is always independent of $z_{w\varepsilon}$ for all values of $r_1 = rank(\Pi_{1W})$ when $p_2 = m_2$ [see discussion below (A.13) in Appendix A]. So, the driving force of the asymptotic pivotality in Theorem 3.2, in spite of relevant instrument exclusion, is the consistency of LIML ($p_2 = m_2$). Otherwise, the validity of plug-in inference is not necessarily preserved, as shown in Theorem 3.3.

THEOREM 3.3. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [\Phi_1, \Phi_2]$ *is fixed. If further* H_0 *holds and* $p_2 < m_2$, *then we have*:

$$AR(\theta_{01}) \xrightarrow{d} \tilde{\xi}_1 = \frac{1}{l} \parallel z_\varepsilon - \zeta_{\pi S}\theta_2^*(\kappa) \parallel^2,$$

$$KLM(\theta_{01}) \xrightarrow{d} \tilde{\xi}_2(r_1) = \parallel Q_Z^{1/2} P_{\tilde{\Omega}(r_1)}.[z_\varepsilon - \zeta_{\pi S}\theta_2^*(\kappa)] \parallel^2,$$

$$JKLM(\theta_{01}) \xrightarrow{d} l\tilde{\xi}_1 - \tilde{\xi}_2(r_1),$$

$$MQLR(\theta_{01})|_{\tau_m} \xrightarrow{d} \frac{1}{2}(l\tilde{\xi}_1 - \tau_m) + \frac{1}{2}\sqrt{(l\tilde{\xi}_1 + \tau_m)^2 - 4[l\tilde{\xi}_1 - \tilde{\xi}_2(r_1)]\tau_m},$$

where $\theta_2^*(\kappa)$ *is given in Theorem* 3.1, $\zeta_{\pi S} = [Q_Z^{1/2}\Pi_{2W}S_1,\ Q_Z^{-1/2}\psi_{ZV_2}S_2]$ *and*: (i) $\tilde{\Omega}(r_1) = M_{Q_Z^{1/2}\Psi_{2S}}.Q_Z^{1/2}\Pi_{1W}$ *if* $r_1 = m_1$; (ii) $\tilde{\Omega}(r_1) = M_{Q_Z^{1/2}\Psi_{2S}}.Q_Z^{1/2}\tilde{\Psi}_{1R}$ *if* $r_1 < m_1$, $\Psi_{2S} = [\Pi_{2W}S_1,\ Q_Z^{-1}\psi_{ZV_1}S_2 - \sigma_\varepsilon^{-1}Q_Z^{1/2}(z_\varepsilon - \zeta_{\pi S}\theta_2^*(\kappa))\sigma'_{V_1\varepsilon}S_2]$, *and* $\tilde{\Psi}_{1R} = [\Pi_{1W}R_1,\ Q_Z^{-1}\psi_{ZV_1}R_2 - \sigma_\varepsilon^{-1}Q_Z^{1/2}(z_\varepsilon - \zeta_{\pi S}\theta_2^*(\kappa))\sigma'_{V_1\varepsilon}R_2]$.

Theorem 3.3 has two main contributions.

First, we provide an explicit analytical characterization of the asymptotic distributions of the plug-in subset tests, including when partial identification holds, i.e., the first-stage regression parameter matrix Π_2 can be fixed with an arbitrary rank. Guggenberger et al. (2012, eqn. 14) provide an upper bound on the asymptotic distribution of the subset AR-statistic, and an analytical characterization of the asymptotic distribution of the subset KLM-statistic (see Guggenberger et al., 2012, eqn. 59). However, both results assume no IV exclusion ($\Phi = 0$) and the asymptotic distribution of the subset KLM-statistic is proved under sequences of Π_2 local to zero (see Guggenberger et al., 2012, eqn. 20). Here, in addition to IV exclusion ($\Phi \neq 0$), we also allow both Π_2 and Φ to have any arbitrary rank (partial identification). Furthermore, to the best of our knowledge, the analytical characterization of the limiting distributions of the subset AR, JKLM, and MQLR statistics in Theorem 3.3, allowing for both Π_2 and Φ to have any arbitrary rank,

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

is new. In addition, the characterization in Theorem 3.3 also provides some new insights. For example, if at least one element of θ_2 is not identifiable ($p_2 < m_2$), $z_\varepsilon - \zeta_{\pi S}\theta_2^*(\kappa)$ has a nonstandard distribution and is always correlated with the score factor $\tilde{\Omega}(r_1)$. As a result, both $\tilde{\xi}_1$ and $\tilde{\xi}_2$ have nonstandard distributions (even when $\Phi = 0$). This contrasts with Theorem 3.2 where $z_{w\varepsilon}$ is a Gaussian process and independent of the score factor[15] $\Omega(r_1)$.

Second, because it is possible that $p_2 < m_2$ even if $rank(\Pi_{22}) = m_2$, Theorem 3.3 shows that lack of identification, if it is due to instrument omission, can be detrimental to the plug-in subset inference. Indeed, under the condition of Theorem 3.3, the upper bound on the limiting distributions of the plug-in subset AR-statistic in Lemma 2.2(b) is not necessarily preserved. There is an intuitive justification with regard to why weak instruments alone, when not coupled with instrument omission, do not invalidate the plug-in subset inference, while exclusion of instruments, even alone, does. Indeed, weak identification alone, although it biases LIML, does not eliminate the first-stage. However, instrument exclusion can eliminate the first-stage, thus weakening identification and invalidating the plug-in inference. However, quantifying the effects of instrument exclusion analytically from Theorem 3.3 is complex, because the effects appear in the asymptotic distribution of the statistics in a nonlinear way, as shown in (3.1), (3.3), and the expression of the score factor $\tilde{\Omega}(r_1)$ in the theorem. Nevertheless, we can see that the subset AR-test is affected by instrument omission only through the term $\zeta_{\pi S}\theta_2^*(\kappa)$. Meanwhile, the effect on the subset KLM-test (and the others) is twofold: a direct effect through $\zeta_{\pi S}\theta_2^*(\kappa)$ and an indirect effect through the score factor $\tilde{\Omega}(r_1)$. It is therefore likely that the subset AR-test, which does not use the score vector, may be less sensitive to instrument omission than the others.

We now analyze the sensitivity of plug-in subset statistics to *possibly local to zero weak instrument exclusion*.

3.2. Possibly Local to Zero Weak Instrument Exclusion

We now focus on cases where the omitted instruments imply a *local to rank deficient* of Π_{2W}, i.e., the situations where the rank of Π_{2W} is not exactly p_2, but approaches p_2 as the sample T increases. As discussed in Section 3, we consider sequences $\Phi = [D_1/\sqrt{T},\ \Phi_2^* + D_2/\sqrt{T}]$ such that $rank(\Pi_{22} + Q_Z^{-1}Q_{ZW}\Phi_2^*) = p_2 < m_2$, for some fixed matrices $D_1 \in \mathbb{R}^{l_1 \times m_1}$, $\Phi_2^* \in \mathbb{R}^{l_1 \times m_2}$, and $D_2 \in \mathbb{R}^{l_1 \times m_2}$ (where all matrices may take zero value). In the above setup, $\Phi_2^* = 0$ implies *local to zero weak omitted instruments* (similar to Staiger and Stock, 1997), while $\Phi_2^* \neq 0$ means that the omitted instruments are only relevant for the first-stage specification of Y_2, but not for those of Y_1. To ease the exposition, we find it useful to separate the case $\Phi_2^* \neq 0$ from $\Phi_2^* = 0$. Although both cases design a *local to rank deficient* of Π_{2W}, the above distinction facilitates the interpretation of the results. When $\Phi_2^* \neq 0$, there is no additional restriction on $\Pi_2 = [\Pi_{21}, \Pi_{22}]$ except the maintained condition $rank(\Pi_{22} + Q_Z^{-1}Q_{ZW}\Phi_2^*) = p_2 < m_2$. In particular, Π_{22} may be local to zero such as in Staiger and Stock (1997). When $\Phi_2^* = 0$,

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

it is straightforward to see that the local to rank deficient assumption of Π_{2W} implies a local to rank deficient of Π_{22}. So, the results of this section hold for any arbitrary rank of Π_{22}, provided $rank(\Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2^*) = p_2 < m_2$. It is important to note that although this section maintains the rank deficiency of Π_{2W} (i.e., $p_2 < m_2$), there is no impediment to expand it to $p_2 = m_2$, which will give similar results as in Theorems 3.1(a) and 3.2. So, without loss of this generality, we shall focus only on cases where $p_2 < m_2$.

We can now state the following two theorems on the asymptotic behavior of the restricted LIML estimator $\tilde{\theta}_2$ (Theorem 3.4) and the plug-in subset statistic (Theorem 3.5), when $\Phi_2^* \neq 0$.

THEOREM 3.4. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [D_1/\sqrt{T},\ \Phi_2^* + D_2/\sqrt{T}]$ *such that* $rank(\Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2^*) = p_2 < m_2$ *for some fixed matrices* $D_1 \in \mathbb{R}^{l_1 \times m_1}$ *and* $D_2 \in \mathbb{R}^{l_1 \times m_2}$ (*possibly zeros*). *If further* $\Phi_2^* \neq 0$ *is fixed, then Theorem* 3.1(*b*) *holds with* Π_{2W} *replaced by* $\Pi_{2W}^* = \Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2^*$.

THEOREM 3.5. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [D_1/\sqrt{T},\ \Phi_2^* + D_2/\sqrt{T}]$ *such that* $rank(\Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2^*) = p_2 < m_2$ *for some fixed matrices* $D_1 \in \mathbb{R}^{l_1 \times m_1}$ *and* $D_2 \in \mathbb{R}^{l_1 \times m_2}$ (*possibly zeros*). *If further* $\Phi_2^* \neq 0$ *is fixed, then Theorem* 3.3 *holds with* Π_{2W} *and* Π_{1W} *replaced by* $\Pi_{2W}^* = \Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2^*$ *and* $\Pi_{1W}^* = \Pi_{21}$, *respectively.*

Theorems 3.4 and 3.5 reinforce the results of Theorems 3.1(b) and 3.3, respectively. It is shown that the results of Theorems 3.1(b) and 3.3 do not depend on any specific value of the rank of Π_{2W} in the neighborhood of rank deficiency. So, as soon as $rank(\Pi_{2W})$ approaches p_2, as T increases, the interpretations and conclusions of Theorems 3.1(b) and 3.3 remain valid for any arbitrary rank of $\Pi_2 = [\Pi_{21} : \Pi_{22}]$ and $\Phi_2^* \neq 0$. In particular, lack of identification, even if it is due to instrument omission in the first-stage regression of Y_2 only ($D_1 = 0$), can be detrimental to LIML and plug-in subset inference.

We will now analyze the case where the omitted instruments are weak in both the first-stage specifications of Y_1 and Y_2, i.e., $\Phi_2^* = 0$. To ease the interpretation of our results, we find it more illuminating to examine the case where $p_2 = \lim_{T \to +\infty} rank(\Pi_{2W}) = 0$ and those where $0 < p_2 = \lim_{T \to +\infty} rank(\Pi_{2W}) < m_2$, separately. This is because when $\Phi_2^* = 0$, both cases are conceptually different. Indeed, the former implies that Π_{2W} is local to zero (similar to Staiger and Stock, 1997), so that the nuisance parameters are completely unidentified. The latter entails that at least one linear combination of the nuisance parameter is identifiable (partial identification). The asymptotic distribution of the subset KLM-statistic in Guggenberger et al. (2012) considers the first case (but with somewhat different parametrization and notations) with no IV exclusion ($D_1 = 0$ and $D_2 = 0$). Furthermore, note that the first case (Π_{2W} is local to zero) is automatically satisfied if Π_2 is local to zero, i.e., if $\Pi_2 = [\Pi_{21} : \Pi_{22}] = [C_1/\sqrt{T} : C_2/\sqrt{T}]$, where $C_1 \in \mathbb{R}^{l \times m_1}$ and $C_2 \in \mathbb{R}^{l \times m_2}$ are fixed matrices (possibly zeros).

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

Theorem 3.6 characterizes the asymptotic behavior of the restricted LIML estimator, $\tilde{\theta}_2$, while Theorem 3.7 presents the limiting distributions of plug-in subset statistics, in both cases. Before we proceed, let us first introduce the following additional definitions and notations:

$$z_{V_2} = Q_Z^{-1/2}\psi_{ZV_2}\Sigma_{V_2}^{-1/2}, \quad \lambda_2 = \left(Q_Z^{1/2}C_2 + \Sigma_Z^{-1/2}Q_{ZW}D_2\right)\Sigma_{V_2}^{-1/2}$$

$$\upsilon_1 = (\lambda_2 + z_{V_2})'(\lambda_2 + z_{V_2}), \quad \upsilon_2 = (\lambda_2 + z_{V_2})'z_\varepsilon$$

$$z_\kappa = (\lambda_2 + z_{V_2})(\upsilon_1 - \kappa I_{m_2})^{-1}(\upsilon_2 - \kappa\rho_2), \quad \rho_2 = \sigma_\varepsilon^{-1}\Sigma_{V_2}^{-1/2}\sigma_{V_2\varepsilon} \tag{3.4}$$

$$\Xi = \begin{pmatrix} z_\varepsilon z_\varepsilon' & (\lambda_2 + z_{V_2}) \\ (\lambda_2 + z_{V_2})' & \upsilon_1 \end{pmatrix}, \quad \bar{\Sigma}_2 = \begin{pmatrix} 1 & \rho_2' \\ \rho_2 & I_{m_2} \end{pmatrix} \tag{3.5}$$

$$\Omega_{D_2\Pi} = D_2'Q_{ZW}'\Pi_{22} + \Pi_{22}'Q_{ZW}D_2, \quad \Omega_{D_2} = D_2'Q_{ZW}'Q_Z^{-1}Q_{ZW}D_2, \tag{3.6}$$

where $z_\varepsilon = \sigma_\varepsilon^{-1}Q_Z^{-1/2}\psi_{Z\varepsilon} \sim \mathbf{N}(0, I_l)$ by Assumption 2.1-*AS*2. We can now characterize the asymptotic distribution of $\tilde{\theta}_2$.

THEOREM 3.6. *Suppose Assumption* 2.1 *and* H_0 *are satisfied. Let* $\Phi = [D_1/\sqrt{T}, D_2/\sqrt{T}]$ *for some fixed matrices* $D_1 \in \mathbb{R}^{l_1 \times m_1}$ *and* $D_2 \in \mathbb{R}^{l_1 \times m_2}$ (*possibly zeros*). *Then the following convergence holds*:

(a) *if* $\Pi_2 = [\Pi_{21} : \Pi_{22}] \neq 0$ *is fixed with* $rank(\Pi_{22}) = p_2 < m_2$, *then we have* $\Upsilon_T S'(\tilde{\theta}_2 - \theta_2) = \begin{pmatrix} \sqrt{T}(\tilde{\theta}_{21} - \theta_{21}) \\ \tilde{\theta}_{22} - \theta_{22} \end{pmatrix} \xrightarrow{d} \theta_2^*(\kappa^*) \overset{d}{\equiv} \Xi_2^{-1}(\kappa^*)\zeta_2^*(\kappa^*)$, *where* $\Xi_2(\kappa^*)$ *and* $\zeta_2^*(\kappa^*)$ *are defined as in* (3.1) *and Theorem* 3.1(*b*) *with* $\Pi_{2W} = \Pi_{22} + T^{-1/2}Q_Z^{-1}Q_{ZW}D_2$;

(b) *if* $\Pi_2 = [\Pi_{21} : \Pi_{22}] = [C_1/\sqrt{T} : C_2/\sqrt{T}]$ *where* $C_1 \in \mathbb{R}^{l \times m_1}$ *and* $C_2 \in \mathbb{R}^{l \times m_2}$ *are fixed matrices* (*possibly zeros*), *then we have*

$$\tilde{\theta}_2 - \theta_2 \xrightarrow{d} \bar{\theta}_2(\bar{\kappa}) = \sigma_\varepsilon\Sigma_{V_2}^{-1/2}(\upsilon_1 - \bar{\kappa}I_{m_2})^{-1}(\upsilon_2 - \bar{\kappa}\rho_2),$$

where $\bar{\kappa}$ *is the smallest eigenvalue of* $|\Xi - \kappa\bar{\Sigma}_2| = 0$, υ_1, υ_2, ρ_2, Ξ, *and* $\bar{\Sigma}_2$ *are defined in* (3.4) *and* (3.5).

Theorem 3.6(a) reinforces our previous analysis in Theorem 3.1(b). It shows that the exclusion of weak instruments can affect the behavior of LIML. Indeed, following (3.3), we can decompose $\Xi_2(\kappa^*)$ and $\zeta_2^*(\kappa^*)$ [in the expression of $\theta_2^*(\kappa^*)$] as:

$$\Xi_2(\kappa^*) = \Xi_\Pi(\kappa^*) + \Xi_{D_2\Pi}(\kappa^*), \ \zeta_2^*(\kappa^*) = \zeta_{2,\Pi}^*(\kappa^*) + \zeta_{2,D_2\Pi}^*(\kappa^*) \quad \text{where} \tag{3.7}$$

$$\Xi_\Pi(\kappa^*) = \begin{bmatrix} S_1'\Pi_{22}'Q_Z\Pi_{22}S_1 & S_1'\Pi_{22}'\psi_{ZV_2}S_2 \\ S_2'\psi_{ZV_2}'\Pi_{22}S_1 & S_2'(\psi_{ZV_2}'Q_Z^{-1}\psi_{ZV_2} - \kappa^*\Sigma_{V_2})S_2 \end{bmatrix},$$

$$\zeta_{2,\Pi}^*(\kappa^*) = \begin{pmatrix} S_1'\Pi_{22}'\psi_{Z\varepsilon} \\ S_2'(\psi_{ZV_2}'Q_Z^{-1}\psi_{Z\varepsilon} - \kappa^*\sigma_{V_2\varepsilon}) \end{pmatrix}$$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

$$\Xi_{D_2\Pi}(\kappa^*) = \begin{bmatrix} S_1'(T^{-1/2}\Omega_{D_2\Pi} + T^{-1}\Omega_{D_2})S_1 & T^{-1/2}S_1'D_2'Q_{ZW}'Q_Z^{-1}\psi_{ZV_2}S_2 \\ T^{-1/2}S_2'\psi_{ZV_2}'Q_Z^{-1}Q_{ZW}D_2S_1 & 0 \end{bmatrix},$$

$$\zeta^*_{2,D_2\Pi}(\kappa^*) = \begin{pmatrix} T^{-1/2}S_1'D_2'Q_{ZW}'Q_Z^{-1}\psi_{Z\varepsilon} \\ 0 \end{pmatrix}.$$

Again, the presence of the terms $\Xi_{D_2\Pi}(\kappa^*)$ and $\zeta^*_{2,D_2\Pi}(\kappa^*)$ in (3.7) is due to instrument omission. Although both terms converge to zero almost surely, they never vanish completely in the sample unless $D_2 = 0$ (no IV exclusion). So, even the exclusion of weak instruments can bias LIML estimates and their standard errors in the sample.

Part (b) shows the exclusion of instruments, if it is coupled with weak identification, can substantially affect the consistency and asymptotic normality of LIML. Indeed, since any linear combination of the nuisance parameters is identifiable in Theorem 3.6(b), the effects of instrument exclusion can be more worse than there are in Theorem 3.1(b). Note, however, that it is difficult to quantify the impact of instrument exclusion alone in Theorem 3.6. This is because both $\theta_2^*(\kappa^*)$ and $\bar{\theta}_2(\bar{\kappa})$ depend on D_2 in a complex way through the expressions of $\Xi_2(\kappa^*)$, $\zeta_2^*(\kappa^*)$, v_1, v_2, and $\bar{\kappa}$. Moreover, it is worthwhile noting that although the condition of Theorem 3.6(a) excludes zero value for Π_2, there is no impediment to expand it to $\Pi_2 = 0$. Indeed, we can easily show from the definition of the rotation $S = [S_1, S_2]$ in Supplemental Appendix A, along with Theorem 3.1(b), that setting $\Pi_2 = 0$ in Theorem 3.6(a) gives the result of Theorem 3.6(b) when $C_1 = 0$, and $C_2 = 0$.

We can now state the following theorem on the asymptotic distributions of plug-in subset statistics.

THEOREM 3.7. *Suppose Assumption* 2.1 *and* H_0 *are satisfied. Let* $\Phi = [D_1/\sqrt{T}, D_2/\sqrt{T}]$ *for some fixed matrices* $D_1 \in \mathbb{R}^{l_1 \times m_1}$ *and* $D_2 \in \mathbb{R}^{l_1 \times m_2}$ (*possibly zeros*). *Then the following convergence holds*:

(a) *if* $\Pi_2 = [\Pi_{21} : \Pi_{22}] \neq 0$ *is fixed with* $rank(\Pi_{22}) = p_2 < m_2$, *then Theorem* 3.3 *holds with* $\Pi_{2W} = \Pi_{22} + T^{-1/2}Q_Z^{-1}Q_{ZW}D_2$ *and* $\Pi_{1W} = \Pi_{22} + T^{-1/2}Q_Z^{-1}Q_{ZW}D_1$;

(b) *if* $\Pi_2 = [\Pi_{21} : \Pi_{22}] = [C_1/\sqrt{T} : C_2/\sqrt{T}]$ *where* $C_1 \in \mathbb{R}^{l \times m_1}$ *and* $C_2 \in \mathbb{R}^{l \times m_2}$ *are fixed matrices* (*possibly zeros*), *then*

$$AR(\theta_{01}) \xrightarrow{d} \xi_1^L(\bar{\kappa}) = \frac{1}{l} \parallel z_\varepsilon - z_{\bar{\kappa}} \parallel^2,$$

$$KLM(\theta_{01}) \xrightarrow{d} \xi_2^L(\bar{\kappa}) = \parallel P_{\Omega(\bar{\kappa})}.(z_\varepsilon - z_{\bar{\kappa}}) \parallel^2,$$

$$JKLM(\theta_{01}) \xrightarrow{d} l\xi_1^L(\bar{\kappa}) - \xi_2^L(\bar{\kappa}), \quad \textit{and}$$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

$$MQLR(\theta_{01})|_{\tau_m} \xrightarrow{d} \frac{1}{2}(l\xi_1^L(\bar{\kappa}) - \tau_m) + \frac{1}{2}\sqrt{(l\xi_1^L(\bar{\kappa}) + \tau_m)^2 - 4[l\xi_1^L(\bar{\kappa}) - \xi_2^L(\bar{\kappa})]\tau_m},$$

where $\Omega(\bar{\kappa}) = M_{Q_Z^{1/2}\Pi_{22}(\bar{\kappa})} \times Q_Z^{1/2}\Pi_{21}(\bar{\kappa})$, $\Pi_{21}(\bar{\kappa}) = C_1 + Q_Z^{-1}Q_{ZW}D_1 + \sigma_\varepsilon^{-1}Q_Z^{-1/2}(z_\varepsilon - z_{\bar{\kappa}})\sigma'_{V_1\varepsilon}$, $\Pi_{22}(\bar{\kappa}) = C_2 + Q_Z^{-1}Q_{ZW}D_2 + \sigma_\varepsilon^{-1}Q_Z^{-1/2}(z_\varepsilon - z_{\bar{\kappa}})\sigma'_{V_2\varepsilon}$, z_ε *and* $z_{\bar{\kappa}}$ *are given in* (3.4)–(3.5) *and* $\bar{\kappa}$ *is defined in Theorem* 3.6(*b*).

Theorem 3.7 shows that instrument exclusion, if coupled with partial or weak identification, can worsen the size property of the plug-in subset tests. This is confirmed by our Monte Carlo study in the next section. It is worth noting that except when $D = [D_1, D_2] = 0$, the bound on the limiting distribution of the subset AR-statistic in Lemma 2.2(b) and Guggenberger et al. (2012, eqn. 14)[16] does not necessarily apply. The characterization of the asymptotic distribution of the subset KLM-statistic in Guggenberger et al. (2012, eqn. 59) is similar to those of Theorem 3.7(b) when $D = 0$ (no IV exclusion), but with somewhat different parametrization and notations. So, the result of Theorem 3.7(b) for this statistic can be viewed as a generalization of Guggenberger et al. (2012, eqn. 59) to instrument exclusion setting. The extension to weak values of Π_2 in Theorem 3.7(a) for the subset KLM-statistic as well as the characterization of the asymptotic distributions of the other subset statistics (including the subset AR-statistic) is new here, in all cases even when $D = 0$ (no IV exclusion).

It is worth mentioning that it is difficult to quantify analytically the effects of instrument exclusion alone because of the complexity of the forms of $\tilde{\xi}_1$, $\tilde{\xi}_2$ (Theorem 3.3), and $\xi_1^L(\bar{\kappa})$, and $\xi_2^L(\bar{\kappa})$ (Theorem 3.7(b)). Thus, we will complete our sensitivity analysis through a Monte Carlo experiment.

3.3. Monte Carlo Experiment

We use simulation to examine the size performance of the subset tests when subject to instrument omission. The data generating process is described by (2.1) and (2.5) where y, Y_1, and Y_2 are $T \times 1$ vectors.[17] The T rows of $[\varepsilon : V]$ are i.i.d. (independent and identically distributed) with zero mean, unit variance, and the correlation between u_t and V_{jt} $(j = 1, 2)$ is set at $\rho = .52$. The correlation coefficient, ρ, represents the degree of endogeneity of Y_j, $j = 1, 2$, and the distribution for $[\varepsilon : V]$ will vary as described below in (3.8). The l columns of the included instruments Z, $l \in \{5, 20, 40\}$, are $N(0, I_l)$, independently from $[\varepsilon : V]$. The omitted instrument W is a $T \times 1$ vector[18] such $W = M_Z\tilde{W}$ and the $\tilde{W}_t$ are generated as i.i.d $\mathbf{N}(0, 1)$ variables.[19] Both Z and $\tilde{W}$ are kept fixed over the simulation experiment. The true values of θ_1 and θ_2 are set at 1 and .5, respectively. The reduced-form coefficient matrix Π_2 is chosen as $\Pi_2 = (\frac{\mu^2}{n\|ZC_0\|})^{1/2}\Pi_0$, where Π_0

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

is obtained by taking the first 2 columns[20] of an identity matrix of order l, μ^2 is the concentration parameter which describes the strength of the included IVs Z. In this experiment, we vary μ^2 in $\{0, 13, 413, 1,000\}$, where $\mu^2 = 0$ is a complete nonidentification or irrelevant IVs setup, $\mu^2 \in \{13, 413\}$ represents weak instruments, and $\mu^2 = 1,000$ implies strong instruments.[21] The coefficients in the first-stage specification of the omitted instrument W is set at $\Phi = \varphi(1, 1)$. The parameter φ characterizes the relevance of the omitted instruments. We vary φ in $\{0, 0.01, 0.1, 0.5, 1\}$, where $\varphi = 0$ is a design of no relevant instrument exclusion, $\varphi = .01$ symbolizes weak omitted instruments, $\varphi = .1$ designs moderately weak omitted instruments, and finally $\varphi \in \{.5, 1\}$ are designs of strong omitted instruments. The design $\varphi = 0$ satisfies the conditions of Lemma 2.2 (which is the benchmark in our analysis).

We consider the following specification of the distribution of $[\varepsilon : V]$ which allows both normal and nonnormal distributions (see Kotz et al., 2000):

$$\begin{aligned} \varepsilon_t &= a + bu_{0t} + cu_{0t}^2 + du_{0t}^3 \\ V_{jt} &= a + bu_{jt} + cu_{jt}^2 + du_{jt}^3,\ j = 1, 2 \end{aligned} \tag{3.8}$$

where $(u_{0t}, u_{1t}, u_{2t})' \overset{i.i.d.}{\sim} N\left[0, \Sigma = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & 0 \\ \rho & 0 & 1 \end{pmatrix}\right]$ for all $t = 1, \ldots, T$. Fleishman (1978) describes the simultaneous equations system that a, b, c, d should satisfy to yield the desired skewness (κ_3) and kurtosis (κ_4), for (ε_t, V_{jt}). We use two parametrization corresponding to: (**1**) normal errors[22] and (**2**) nonnormal errors.[23] In each setup, the empirical rejection frequencies of the subset tests are evaluated for sample sizes $T = 100$ and $T = 300$ using $10,000$ replications. Except for the subset MQLR statistic, whose critical values are computed within experiment (see Moreira, 2003), the nominal level of the other tests is set at 5%.

Table 1 presents the results. The first column of the table shows the statistics. The second column indicates the number of instruments included in the computation of the statistics. The other columns show, for each value of φ and instrument quality μ^2, the rejection frequencies. First, we note that the results are similar for normal and nonnormal errors even when the sample size is relatively moderate ($T = 100$). Second, when $\varphi = 0$ (which is the reference case of no IV exclusion), all plug-in subset tests are conservative with weak instruments ($\mu^2 \in \{0, 13, 413\}$), but their rejection frequencies is close to the nominal level with strong instruments ($\mu^2 = 1,000$). Third, we see that even the exclusion of weak instruments affects all subset tests negatively (including the AR one) and the impact increases with the quality of the omitted instruments. Indeed, the subset KLM-test is highly sensitive to instrument omission and more often rejects the null hypothesis than the other tests. This test has a maximal rejection frequency as great as 92%, even when the included instruments are strong; see, for example, columns $\varphi \in \{.5, 1\}$ and $\mu^2 \in \{0, 13, 413\}$ in Table 1. The subset MQLR-test is more affected by

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

TABLE 1. Rejection frequencies (in %) at nominal level 5%

Normal errors, $T = 100$																				
		$\varphi = 0$				$\varphi = .01$				$\varphi = .1$				$\varphi = .5$				$\varphi = 1$			
Statistics	$l \downarrow \mu^2 \rightarrow$	0	13	413	1,000	0	13	413	1,000	0	13	413	1,000	0	13	413	1,000	0	13	413	1,000
AR	5	0.3	0.5	4.4	5.1	0.2	0.4	5.0	5.4	0.2	0.5	4.5	5.4	0.1	0.1	4.3	4.9	0.0	0.0	2.7	4.8
KLM	-	1.2	1.2	4.5	5.1	0.9	1.5	4.8	5.1	1.0	1.5	4.6	5.3	0.9	1.1	4.4	4.7	0.7	0.8	3.0	4.3
JKLM	-	0.4	0.6	4.3	4.8	0.3	0.6	5.1	5.3	0.3	0.6	4.9	5.7	0.1	0.2	4.5	4.8	0.0	0.0	3.2	4.9
MQLR	-	0.2	0.2	1.4	3.1	0.1	0.3	1.3	1.1	0.1	0.3	1.0	0.8	0.0	0.0	1.4	1.0	0.0	0.0	1.2	2.5
AR	20	0.4	0.4	3.0	6.0	0.4	0.4	2.6	4.9	0.3	0.4	2.9	6.1	0.0	0.0	0.7	3.8	0.0	0.0	0.0	0.4
KLM	-	3.4	3.8	6.3	6.6	3.9	3.9	7.6	7.8	3.6	3.8	5.7	6.0	24.4	22.3	5.5	4.9	39.9	39.5	29.9	16.5
JKLM	-	0.4	0.4	3.0	5.6	0.4	0.3	2.2	4.4	0.3	0.4	2.9	5.9	0.0	0.0	0.6	3.7	0.0	0.0	0.0	0.1
MQLR	-	1.5	1.7	6.2	5.3	1.5	1.6	6.1	6.0	1.4	1.6	5.8	5.2	0.2	0.3	1.5	6.4	0.0	0.0	0.1	0.9
AR	40	0.8	0.7	2.6	5.4	0.8	0.9	2.1	4.7	0.7	0.7	2.0	5.1	0.0	0.0	0.2	0.8	0.0	0.0	0.0	0.0
KLM	-	5.3	5.6	5.8	5.9	7.6	7.4	8.3	9.8	7.2	7.3	8.1	9.6	55.5	55.1	35.7	20.6	90.0	89.3	88.4	86.0
JKLM	-	0.6	0.6	2.1	4.7	0.6	0.8	1.8	4.4	0.6	0.7	1.9	4.5	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0
MQLR	-	5.2	5.7	5.9	5.8	5.7	5.9	11.0	14.3	4.6	4.7	10.4	15.8	0.8	0.9	2.1	5.2	0.0	0.0	0.0	0.1
Normal errors, $T = 300$																					
AR	5	0.3	0.3	3.4	4.9	0.2	0.2	3.9	4.8	0.2	0.3	4.0	4.7	0.1	0.1	3.1	4.6	0.0	0.0	0.9	3.5
KLM	-	0.8	1.2	4.1	4.7	1.0	1.1	4.8	5.0	1.0	1.1	4.5	5.1	0.9	0.9	3.5	4.8	0.7	0.6	2.1	4.0
JKLM	-	0.4	0.3	3.5	5.0	0.3	0.4	4.0	4.7	0.3	0.5	4.1	4.6	0.0	0.1	3.5	4.5	0.0	0.0	1.3	3.9
MQLR	-	0.1	0.2	1.4	3.0	0.1	0.1	1.3	1.6	0.1	0.1	1.5	0.9	0.0	0.0	1.2	1.1	0.0	0.0	0.2	1.8
AR	20	0.2	0.3	1.0	3.8	0.2	0.2	1.3	3.5	0.1	0.2	1.3	2.8	0.0	0.0	0.1	0.8	0.0	0.0	0.0	0.0
KLM	-	3.2	3.0	3.9	4.8	2.7	2.7	5.2	5.9	2.8	2.8	4.1	5.2	18.6	17.7	9.9	5.9	40.4	39.9	38.7	27.4
JKLM	-	0.2	0.3	1.1	3.9	0.3	0.2	1.3	3.2	0.1	0.3	1.3	2.7	0.0	0.0	0.1	0.7	0.0	0.0	0.0	0.0
MQLR	-	1.5	1.2	3.6	5.5	1.1	1.1	4.4	5.3	1.0	1.0	4.0	5.8	0.2	0.2	0.7	2.5	0.0	0.0	0.0	0.3
AR	40	0.2	0.3	0.7	2.9	0.3	0.3	0.7	1.6	0.2	0.2	0.7	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
KLM	-	3.9	4.0	5.1	5.9	3.9	4.3	5.0	6.6	4.2	3.8	4.6	5.9	58.4	57.7	43.2	28.1	91.3	92.0	91.0	89.6
JKLM	-	0.3	0.2	0.7	3.9	0.3	0.3	0.7	1.5	0.2	0.2	0.7	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MQLR	-	4.1	3.9	5.8	5.4	4.1	4.2	7.6	12.2	3.7	3.4	7.2	12.0	0.8	0.7	1.1	2.5	0.0	0.0	0.0	0.1

Table continues on overleaf

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

TABLE 1. *continued*

		$\varphi = 0$				$\varphi = .01$				$\varphi = .1$				$\varphi = .5$				$\varphi = 1$			
Nonnormal errors, $T = 100$																					
AR	5	0.2	0.6	5.4	5.7	0.2	1.0	5.5	5.7	0.3	0.8	5.4	5.4	0.0	0.1	5.1	5.6	0.0	0.0	4.1	5.3
KLM	-	1.2	1.8	5.2	5.3	1.4	2.1	5.5	5.5	1.3	1.7	5.4	4.7	1.0	1.3	5.1	5.1	0.1	0.3	4.5	5.0
JKLM	-	0.3	0.8	5.4	5.4	0.4	1.1	5.4	5.4	0.3	1.0	5.3	5.8	0.0	0.1	5.0	5.6	0.0	0.0	4.3	5.2
MQLR	-	0.3	0.6	1.0	3.2	0.4	0.8	0.6	0.4	0.3	0.7	1.3	0.9	0.0	0.1	1.6	1.1	0.0	0.0	3.2	3.6
AR	20	0.3	0.4	5.4	5.6	0.3	0.4	6.0	7.2	0.4	0.5	5.2	6.7	0.0	0.0	2.0	6.0	0.0	0.0	0.0	0.1
KLM	-	3.7	3.9	6.3	5.8	3.7	3.7	6.8	5.7	3.8	3.9	5.5	5.7	37.5	35.8	10.5	6.4	15.6	16.7	38.7	31.7
JKLM	-	0.2	0.4	5.1	5.9	0.3	0.4	5.6	7.0	0.4	0.4	5.0	6.4	0.0	0.0	1.2	5.7	0.0	0.0	0.0	0.0
MQLR	-	5.1	5.8	5.8	3.4	5.5	6.1	4.9	1.8	4.8	5.1	6.9	3.4	0.2	0.2	11.4	20.5	0.0	0.0	0.1	1.1
AR	40	0.8	0.7	5.1	5.8	0.7	0.8	4.2	8.4	0.6	0.6	4.6	8.7	0.0	0.0	0.1	2.2	0.0	0.0	0.0	0.0
KLM	-	5.6	5.3	5.4	6.0	7.8	7.5	10.2	9.4	7.7	8.1	10.5	9.8	76.9	74.9	41.9	18.3	87.4	87.6	85.4	77.7
JKLM	-	0.6	0.6	4.8	6.4	0.6	0.7	3.6	7.4	0.6	0.5	3.9	7.8	0.0	0.0	0.0	1.3	0.0	0.0	0.0	0.0
MQLR	-	6.0	5.4	5.7	6.1	15.6	17.0	18.4	12.2	13.9	14.3	21.2	14.1	0.9	0.8	6.2	24.4	0.0	0.0	0.1	1.1
Nonnormal errors, $T = 300$																					
AR	5	0.3	0.3	4.5	5.1	0.3	0.5	4.8	4.7	0.2	0.4	4.9	4.9	0.0	0.1	3.9	4.5	0.0	0.0	1.1	4.0
KLM	-	1.1	1.3	4.6	5.0	1.1	1.3	4.7	5.1	0.9	1.3	4.7	5.2	1.1	1.2	4.3	4.8	0.1	0.1	3.4	4.7
JKLM	-	0.4	0.6	4.6	5.3	0.4	0.7	4.7	4.6	0.3	0.7	5.0	5.0	0.0	0.1	4.1	4.7	0.0	0.0	1.0	4.2
MQLR	-	0.2	0.4	1.3	3.3	0.3	0.4	1.0	0.8	0.2	0.5	1.0	0.9	0.0	0.1	2.0	1.3	0.0	0.0	0.5	3.2
AR	20	0.2	0.2	2.2	4.5	0.2	0.3	2.5	4.7	0.1	0.1	2.5	4.7	0.0	0.0	0.2	2.6	0.0	0.0	0.0	0.0
KLM	-	2.8	2.9	4.6	5.4	2.7	2.8	4.6	5.3	3.2	2.9	4.5	5.0	34.5	32.7	14.4	7.2	11.1	12.5	36.2	45.2
JKLM	-	0.2	0.2	2.2	4.5	0.2	0.2	2.4	4.6	0.1	0.1	2.7	4.6	0.0	0.0	0.1	2.1	0.0	0.0	0.0	0.0
MQLR	-	4.6	5.2	7.4	3.8	4.9	5.4	7.2	3.7	3.9	4.1	7.7	4.2	0.2	0.2	4.0	15.6	0.0	0.0	0.0	0.1
AR	40	0.3	0.3	1.2	3.6	0.2	0.2	1.5	3.9	0.1	0.2	1.0	3.6	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0
KLM	-	4.2	3.9	6.0	5.4	4.0	4.1	6.1	6.5	4.2	4.4	5.1	6.7	85.4	85.8	70.6	42.9	78.4	78.8	85.7	91.6
JKLM	-	0.3	0.3	1.2	3.2	0.2	0.3	1.4	3.6	0.1	0.1	1.0	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MQLR	-	5.6	5.9	5.7	5.6	15.5	16.2	19.8	13.4	14.3	14.7	21.1	15.6	0.5	0.4	2.8	12.9	0.0	0.0	0.0	0.0

instrument omission than AR and JKLM, but its maximal rejection is only around 16% with normal errors and 25% with nonnormal errors. However, AR and JKLM become overly conservative and sometimes never reject the null hypothesis despite strong instruments exclusion; see columns $\varphi \in \{.5, 1\}$ and $\mu^2 = 1,000$. The latter observation also holds for the subset MQLR-test. This behavior of the tests is fundamentally different from the reference case of no excluded instrument and strong included IVs: column $\varphi = 0$ and $\mu^2 = 1,000$. Clearly, the exclusion of instruments alone, even if it not coupled with weak identification, can be detrimental to the plug-in subset inference. This confirms our theoretical findings in Section 3. Furthermore, we also observe that there are cases where the exclusion of relevant instruments does not invalidate the plug-in subset inference [see column $\varphi = .1$, $\mu^2 = 1,000$ for example], thus confirming the results of Theorem 3.2.

4. CONCLUSIONS

In this paper, we analyze the sensitivity of the LIML estimator and the plug-in subset tests—namely AR, KLM, JKLM, and MQLR tests—to instrument exclusion. Our analysis of the asymptotic behavior of the statistics considered two main setups: *possibly strong instrument omission* and *local-to-zero weak instrument omission*. In both setups, we allowed the first-stage regression parameter matrix to have an arbitrary rank, thus extending earlier results in partial identified models by Phillips (1989) and Choi and Phillips (1992) to LIML estimators.

In both setups, our characterization of the limiting distributions of the LIML estimator and the plug-in subset statistics provides several new insights and extensions of earlier studies. We show that exclusion of instruments can eliminate the first-stage, thus weakening identification and invalidating the plug-in subset inference. However, when instrument omission does not affect the consistency of LIML, it preserves the plug-in subset test validity, although LIML is no longer asymptotically efficient. Moreover, we also show that the limiting distribution of LIML estimator of the identified linear combination of the nuisance parameters is not necessarily a Gaussian mixture. This contrasts with the identified linear combination of the usual IV estimator, which is still asymptotically a Gaussian mixture, in spite of instrument exclusion.

We present a Monte Carlo experiment which indicates that the subset KLM and MQLR tests are highly sensitive to instrument omission and can reject the subset null hypothesis in up to 92% (KLM) and 25% (MQLR) of the cases. AR and JKLM, however, do not show dramatic rejections of the subset null hypothesis even when relevant instruments are omitted. They are, however, overly conservative.

NOTES

1. Mavroeidis (2004, 2005), Dufour, Khalaf, and Kichian (2006), and Kleibergen and Mavroeidis (2008).

2. Kocherlakota (1990), Hansen, Heaton, and Yaron (1996), and Kleibergen (2005, 2009).

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

3. Bean (1994) and Malcomson and Mavroeidis (2006).

4. See Dufour (1997), Dufour and Jasiak (2001), Dufour and Taamouti (2005, 2007), Dufour, Khalaf, and Beaulieu (2010), Doko Tchatoka, and Dufour (2013), and Dufour, Khalaf, and Kichian (2013).

5. See also Stock and Wright (2000), Guggenberger and Smith (2005), and Antoine and Renault (2012).

6. This is a normalization that requires excluding redundant columns from Z so that $Z'Z$ is nonsingular with probability one. If Z_t is generated by power series or splines through an underlying scalar instrument, say x_t, *i.e.*, if $Z_t = p(x_t) = (1, x_t, \ldots, x_t^{l-1})'$, then Z has full column rank with probability one if $T > l$; see Hansen, Hausman, and Newey (2008, Assump. 1).

7. See Kleibergen (2004), Kleibergen (2008), and Kleibergen and Mavroeidis (2011).

8. See Stock and Wright (2000), Kleibergen (2004), and Startz et al. (2006).

9. The reduced-form equation for the dependent variable $y - Y_1\theta_{01}$ can be expressed under H_0 as:

$$y - Y\theta_{01} = Z\delta^* + \varepsilon^*,$$

where $\delta^* = \Pi_{22}\theta_2$ and $\varepsilon^* = V_2^*\theta_2 + \varepsilon$. If (ε, V_2^*) has a zero mean and uncorrelated with Z, then ε^* also has a zero mean and is uncorrelated with Z under H_0. Therefore, as soon as Z has full-column rank with probability one, the parameter δ^* in the above equation is identifiable under H_0, and the relationship $\delta^* = \Pi_{22}\theta_2$ implies that θ_2 is identifiable under H_0 if, and only if, $rank(\Pi_{22}) = m_2$.

10. See also Dufour (2009).

11. See Dufour (1997), Dufour and Jasiak (2001), and Dufour and Taamouti (2005, 2007).

12. $\mathcal{O}(m_2) = \{H \in \mathbb{M}(m_2, m_2) : H'H = I_{m_2}\}$ denotes the orthogonal group of $m_2 \times m_2$ matrices, where $\mathbb{M}(m_2, m_2)$ is the set of all squared matrices of order m_2.

13. *Fixed instrument asymptotics* usually imposes a full-column rank on the first stage regression parameter matrix that controls the strength of the instruments, see for example, Lemma 2.2(a). To the best of our knowledge, there is no explicit characterization in the previous literature of the asymptotic distribution of LIML when Π_2 is fixed with a deficient rank, even if $\Phi = 0$ (no instrument exclusion). Choi and Phillips (1992) provide such a characterization for the 2SLS estimator.

14. See for example, Staiger and Stock (1997).

15. Note that $z_\varepsilon - \zeta_{\pi S}\theta_2^*(\kappa)$ and $\tilde{\Omega}(r_1)$ are analogous to the terms $z_{w\varepsilon}$ and $\Omega(r_1)$ of Theorem 3.2, respectively.

16. See also Guggenberger and Chen (2011)

17. We ran the simulations with two endogenous regressors in both Y_1 and Y_2 and found similar results.

18. It is worth mentioning that we reach similar conclusions when W contains 5 and 10 columns.

19. Dufour and Taamouti (2007) use a similar data generating process for the omitted instrument.

20. The simulations where Π_0 is a $l \times 2$ matrix of ones yield the same conclusions.

21. Hansen et al. (2008) show that the values of μ^2 below 613 correspond to weak instruments, while those above 613 indicate strong instruments. Guggenberger (2010) uses a similar parametrization.

22. $a = c = d = 0$ and $b = 1$.

23. $a = c = 0$, $b = -.3 + \sqrt{.94}$, and $d = 1/10$, which yields $\kappa_3 = 0$ and $\kappa_4 \approx 5.8$ as skewness and kurtosis. The results are qualitatively the same with alternative values of (κ_3, κ_4).

REFERENCES

Anderson, T.W. & H. Rubin (1949) Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* 20, 46–63.

Andrews, D.W.K. & J.H. Stock (2007) Inference with weak instruments. In R. Blundell, W. Newey, & T. Pearson (eds.), *Advances in Economics and Econometrics, Theory and Applications, 9th Congress of the Econometric Society*, Vol. 3, Chap. 6. Cambridge University Press.

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

Antoine, B. & E. Renault (2012) Efficient minimum distance estimation with multiple rates of convergence. *Journal of Econometrics* 170(2), 350–367.

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

Berkowitz, D., M. Caner, & Y. Fang (2008) Are nearly exogenous instruments reliable? *Economics Letters* 101, 20–23.

Berkowitz, D., M. Caner, & Y. Fang (2012) The validity of instruments revisited. *Journal of Econometrics* 166, 255–266.

Chaudhuri, S. & E. Zivot (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.

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

Doko Tchatoka, F. & J.-M. Dufour (2008) Instrument endogeneity and identification-robust tests: Some analytical results. *Journal of Statistical Planning and Inference* 138(9), 2649–2661.

Doko Tchatoka, F. & J.-M. Dufour (2014) Identification-robust inference for endogeneity parameters in linear structural models. *The Econometrics Journal* 17(1), 165–187.

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

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

Dufour, J.-M. (2009) Comments on "Weak instrument robust tests in GMM and the new Keynesian Phillips curve" by F. Kleibergen and S. Mavroeidis. *Journal of Business and Economic Statistics* 27, 318–321.

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

Dufour, J.-M., L. Khalaf, & M.-C. Beaulieu (2010) Multivariate residual-based finite-sample tests for serial dependence and GARCH with applications to asset pricing models. *Journal of Applied Econometrics* 25, 263–285.

Dufour, J.-M., L. Khalaf, & M. Kichian (2006) Inflation dynamics and the new Keynesian Phillips curve: An identification robust econometric analysis. *Journal of Economic Dynamics and Control* 30(9–10), 1707–1727.

Dufour, J.-M., L. Khalaf, & M. Kichian (2013) Identification-robust analysis of DSGE and structural macroeconomic models. *Journal of Monetary Economics* 60, 340–350.

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

Dufour, J.-M. & M. Taamouti (2007) Further results on projection-based inference in IV regressions with weak, collinear or missing instruments. *Journal of Econometrics* 139(1), 133–153.

Fleishman, A. (1978) A method for simulating nonnormal distributions. *Psychometrika* 43, 521–532.

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

Guggenberger, P. (2012) On the asymptotic size distortion of tests when instruments locally violate the exogeneity assumption. *Econometric Theory* 18, 387–421 .

Guggenberger, P. & L. Chen (2011) On the Asymptotic Size of Subvector Tests in the Linear Instrumental Variables Model. Technical report, Department of Economics, UCSD.

Guggenberger, P., F. Kleibergen, S. Mavroeidis, & L. Chen (2012) On the asymptotic sizes of subset Anderson-Rubin and Lagrange multiplier tests in linear instrumental variables regression. *Econometrica* 80(6), 2649–2666.

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

Habib, A.M. & A.P. Ljungqvist (2001) Interpricing and entrepreneurial wealth losses in IPOs: Theory and evidence. *Review of Financial Studies* 14(2), 433–458.

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

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

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

Kiviet, J.F. & J. Niemczyk (2007) The asymptotic and finite-sample distributions of OLS and simple IV in simultaneous equations. *Computational Statistics and Data Analysis* 51, 3296–3318.

Kiviet, J.F. & J. Niemczyk (2012) Comparing the asymptotic and empirical (un)conditional distributions of OLS and IV in a linear static simultaneous equation. *Computational Statistics and Data Analysis* 56, 3567–3586.

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

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

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

Kleibergen, F. (2008) Subset Statistics in the Linear IV Regression Model. Technical report, Department of Economics, Brown University, Providence, Rhode Island.

Kleibergen, F. (2009) Tests of risk premia in linear factor models. *Journal of Econometrics* 149, 149–173.

Kleibergen, F. & S. Mavroeidis (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. & S. Mavroeidis (2011) 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.

Kotz, S., N. Balakrishnan, & N. Johnson (2000) *Continuous Multivariate Distributions*, 2nd ed. Wiley.

Loughran, T. & J. Ritter (2004) Why has IPO underpricing changed over time? *Financial Management* 33(3), 5–37.

Malcomson, J. & S. Mavroeidis (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.

Mikusheva, A. (2010) Robust confidence sets in the presence of weak instruments. *Journal of Econometrics* 157, 236–247.

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

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

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

Startz, R., C.R.N. Nelson, & E. Zivot (2006) Improved inference in weakly identified instrumental variables regression. In D. Corbae, S.N. Durlauf, & B.E. Hansen (eds.), *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, Chap. 5. Cambridge University Press.

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

Stock, J.H., J.H. Wright, & M. Yogo (2002) A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics* 20(4), 518–529.

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

APPENDIX

A. Proofs of the main results

Proof of Theorem 3.1. To simplify, we prove the results when Π_{22} fixed. The case where Π_{22} is local to zero (i.e., $\Pi_{22} = C_2/\sqrt{T}$ for some fixed matrix C_2, possibly zero) follows immediately.

Suppose first that $p_2 = rank(\Pi_{2W}) = m_2$ and let us write $\sqrt{T}(\tilde{\theta}_2 - \theta_2)$ (under H_0) as:

$$\sqrt{T}(\tilde{\theta}_2 - \theta_2) = \mathbf{L}_T(\kappa_T)^{-1}\mathbf{N}_T(\kappa_T), \tag{A.1}$$

where $\mathbf{L}_T(\kappa_T) = Y_2'P_ZY_2/T - \frac{1}{T}\kappa_T Y_2'M_ZY_2/T$, $\mathbf{N}_T(\kappa_T) = Y_2'P_Z\varepsilon/\sqrt{T} - \frac{1}{T}\kappa_T Y_2'M_Z\varepsilon/T$, and $\kappa_T = T\tilde{\kappa}$. Under Assumption 2.1, we have $Y_2'P_ZY_2/T \xrightarrow{p} Q'_{ZY_2}Q_Z^{-1}Q_{ZY_2} = \Pi'_{2W}Q_Z\Pi_{2W}$, $Y_2'M_ZY_2/T \xrightarrow{p} Q_{Y_2} - \Pi'_{2W}Q_Z\Pi_{2W} = \Sigma_{V_2} + \Phi_2'(Q_W - Q'_{ZW}Q_Z^{-1}Q_{ZW})\Phi_2$, $Y_2'M_Z\varepsilon/T \xrightarrow{p} \sigma_{V_2\varepsilon}$, and $Y_2'P_Z\varepsilon/\sqrt{T} \xrightarrow{d} \Pi'_{2W}\psi_{Z\varepsilon}$, where $\Pi_{2W} = \Pi_{22} + Q_Z^{-1}Q_{ZW}\Phi_2$ (note that $\Pi_{2W} = Q_Z^{-1}Q_{ZW}\Phi_2$ if $\Pi_{22} = C_2/\sqrt{T}$). Since $\kappa_T \xrightarrow{d} \kappa = O_p(1)$ (by Lemma B.1 of the Supplemental Appendix), we have $\mathbf{L}_T(\kappa_T) \xrightarrow{p} \Pi'_{2W}Q_Z\Pi_{2W}$ and $\mathbf{N}_T(\kappa_T) \xrightarrow{d} \Pi'_{2W}\psi_{Z\varepsilon}$ jointly with $\kappa_T \xrightarrow{d} \kappa$. Because $p_2 = m_2$, we have $rank(\Pi_{2W}) = m_2$ so that $\Pi'_{2W}Q_Z\Pi_{2W} > 0$. Therefore, we get

$$\sqrt{T}(\tilde{\theta}_2 - \theta_2) \xrightarrow{d} \zeta_2(\kappa) = \left(\Pi'_{2W}Q_Z\Pi_{2W}\right)^{-1}\Pi'_{2W}\psi_{Z\varepsilon} \sim \mathbf{N}(0, \Omega_2),$$
$$\text{where } \Omega_2 = \sigma_\varepsilon^2\left(\Pi'_{2W}Q_Z\Pi_{2W}\right)^{-1}. \tag{A.2}$$

Suppose now that $p_2 = rank(\Pi_{2W}) < m_2$. Define $\Upsilon_T = diag(\sqrt{T}I_{p_2}, I_{m_2-p_2})$. We can write $\Upsilon_T S'(\tilde{\theta}_2 - \theta_2)$ under H_0 as:

$$\Upsilon_T S'(\tilde{\theta}_2 - \theta_2) = \left[\Upsilon_T^{-1}A_T\Upsilon_T^{-1} - \kappa_T\Upsilon_T^{-1}(B_T/T)\Upsilon_T^{-1}\right]^{-1}\left(\Upsilon_T^{-1}D_T - \kappa_T\Upsilon_T^{-1}E_T/T\right), \tag{A.3}$$

where we have

$$A_T = S'Y_2'P_ZY_2S = \begin{bmatrix} Y_{21}'P_ZY_{21} & Y_{21}'P_ZY_{22} \\ Y_{22}'P_ZY_{21} & Y_{22}'P_ZY_{22} \end{bmatrix},$$
$$B_T = S'Y_2'M_ZY_2S = \begin{bmatrix} Y_{21}'M_ZY_{21} & Y_{21}'M_ZY_{22} \\ Y_{22}'M_ZY_{21} & Y_{22}'M_ZY_{22} \end{bmatrix},$$
$$D_T = S'Y_2'P_Z\varepsilon = \begin{pmatrix} Y_{21}'P_Z\varepsilon \\ Y_{22}'P_Z\varepsilon \end{pmatrix},\ E_T = S'Y_2'M_Z\varepsilon = \begin{pmatrix} Y_{21}'M_Z\varepsilon \\ Y_{22}'M_Z\varepsilon \end{pmatrix};$$

$Y_{2j} = Y_2S_j$, $j = 1, 2$, and $S = [S_1 : S_2] \in \mathcal{O}(m_2)$, are defined in Section B.1 of the Supplemental Appendix. If Assumption 2.1 holds, we have $Y_2'M_ZY_2/T \xrightarrow{p} Q_{Y_2} - \Pi'_{2W}Q_Z\Pi_{2W} > 0$ so that: $Y_{21}'M_ZY_{21}/T = S_1'Y_2'M_ZY_2S_1/T \xrightarrow{p} S_1'(Q_{Y_2} - \Pi'_{2W}Q_Z\Pi_{2W})S_1$, $Y_{21}'M_ZY_{22}/T \xrightarrow{p} S_1'(Q_{Y_2} - \Pi'_{2W}Q_Z\Pi_{2W})S_2$, $Y_{22}'M_Z\varepsilon/T \xrightarrow{p} S_2'\sigma_{V_2\varepsilon}$, and $Y_{22}'M_ZY_{22}/T \xrightarrow{p} S_2(Q_{Y_2} - \Pi'_{2W}Q_Z\Pi_{2W})S_2 = S_2\Sigma_{V_2}S_2 > 0$. Therefore, we get

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

$$\Upsilon_T^{-1}(B_T/T)\Upsilon_T^{-1} = \begin{bmatrix} Y_{21}'M_Z Y_{21}/T^2 & Y_{21}'M_Z Y_{22}/T^{3/2} \\ Y_{22}'M_Z Y_{21}/T^{3/2} & Y_{22}'M_Z Y_{22}/T \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 0 & 0 \\ 0 & S_2'\Sigma_{V_2} S_2 \end{bmatrix}$$

$$\text{and} \quad \Upsilon_T^{-1}E_T/T = \begin{pmatrix} Y_{21}'M_Z\varepsilon/T^2 \\ Y_{22}'M_Z\varepsilon/T \end{pmatrix} \xrightarrow{p} \begin{pmatrix} 0 \\ S_2'\sigma_{V_2\varepsilon} \end{pmatrix}. \tag{A.4}$$

By the same way, we have $Y_{21}'P_Z Y_{21}/T \xrightarrow{p} S_1'\Pi_{2W}'Q_Z\Pi_{2W}S_1$, $Y_{21}'P_Z Y_{22}/\sqrt{T} \xrightarrow{d} S_1'\Pi_{2W}'\psi_{ZV_2}S_2$, and $Y_{22}'P_Z Y_{22} \xrightarrow{d} S_2'\psi_{ZV_2}'Q_Z^{-1}\psi_{ZV_2}S_2$, so that

$$\begin{aligned} \Upsilon_T^{-1}A_T\Upsilon_T^{-1} &= \begin{bmatrix} Y_{21}'P_Z Y_{21}/T & Y_{21}'P_Z Y_{22}/\sqrt{T} \\ Y_{22}'P_Z Y_{21}/\sqrt{T} & Y_{22}'P_Z Y_{22} \end{bmatrix} \\ &\xrightarrow{p} \begin{bmatrix} S_1'\Pi_{2W}'Q_Z\Pi_{2W}S_1 & S_1'\Pi_{2W}'\psi_{ZV_2}S_2 \\ S_2'\psi_{ZV_2}'\Pi_{2W}S_1 & S_2'\psi_{ZV_2}'Q_Z^{-1}\psi_{ZV_2}S_2 \end{bmatrix} \\ \text{and} \quad \Upsilon_T^{-1}D_T &\xrightarrow{p} \begin{bmatrix} S_1'\Pi_{2W}'\psi_{Z\varepsilon} \\ S_2'\psi_{ZV_2}'Q_Z^{-1}\psi_{Z\varepsilon} \end{bmatrix}. \end{aligned} \tag{A.5}$$

From (A.4) and (A.5), along with the fact that $\kappa_T \xrightarrow{d} \kappa$, it is clear that

$$\Upsilon_T S'(\tilde{\theta}_2 - \theta_2) \xrightarrow{d} \theta_2^*(\kappa) = \Xi_1(\kappa)\zeta^*(\kappa) \tag{A.6}$$

and the convergence is also uniform in κ over compact sets, where

$$\Xi_1(\kappa) = \begin{bmatrix} S_1'\Pi_{2W}'Q_Z\Pi_{2W}S_1 & S_1'\Pi_{2W}'\psi_{ZV_2}S_2 \\ S_2'\psi_{ZV_2}'\Pi_{2W}S_1 & S_2'\left(\psi_{ZV_2}'Q_Z^{-1}\psi_{ZV_2} - \kappa\Sigma_{V_2}\right)S_2 \end{bmatrix}^{-1},$$

$$\zeta^*(\kappa) = \begin{bmatrix} S_1'\Pi_{2W}'\psi_{Z\varepsilon} \\ S_2'\left(\psi_{ZV_2}'Q_Z^{-1}\psi_{Z\varepsilon} - \kappa\sigma_{V_2\varepsilon}\right) \end{bmatrix}.$$

■

Proof of Theorem 3.2. To simplify, we prove the results when $\Pi_2 = [\Pi_{21} : \Pi_{22}]$ fixed. The case where Π_2 is local to zero is deduced straightforwardly.

First, let us focus on $\mathrm{AR}(\theta_{01})$. Under H_0, we have:

$$\mathrm{AR}(\theta_{01}) = \frac{1}{l\tilde{\sigma}_\varepsilon^2}\frac{\tilde{\varepsilon}'Z}{\sqrt{T}}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'\tilde{\varepsilon}}{\sqrt{T}}, \tag{A.7}$$

where $\tilde{\sigma}_\varepsilon^2 = \tilde{\varepsilon}'M_Z\tilde{\varepsilon}/(T-l) \xrightarrow{p} \sigma_\varepsilon^2 > 0$ and from Lemma B.2(1) of the Supplemental Appendix, we have $(Z'Z/T)^{-1/2}Z'\tilde{\varepsilon}/\sqrt{T} \xrightarrow{d} M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon}$ for all values of $r_1 = rank(\Pi_{1W})$. From (A.7), it is clear that

$$\mathrm{AR}(\theta_{01}) \xrightarrow{d} \xi_1 = \frac{1}{l\sigma_\varepsilon^2}\psi_{Z\varepsilon}'Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon} = \frac{1}{l}\parallel z_{w\varepsilon}\parallel^2, \tag{A.8}$$

where $z_{w\varepsilon} = M_{Q_Z^{1/2}\Pi_{2W}}(\sigma_\varepsilon^{-1}Q_Z^{-1/2}\psi_{Z\varepsilon}) = M_{Q_Z^{1/2}\Pi_{2W}} \times z_\varepsilon$. By Assumption 2.1, we have $z_\varepsilon = \sigma_\varepsilon^{-1}Q_Z^{-1/2}\psi_{Z\varepsilon} \sim N(0, I_l)$ so that

$$\mathrm{AR}(\theta_{01}) \xrightarrow{d} \xi_1 \sim \frac{1}{l}\chi^2(r_2), \quad \text{where} \quad r_2 = rank\left(M_{Q_Z^{1/2}\Pi_{2W}}\right) \tag{A.9}$$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

for all values of $r_1 = rank(\Pi_{1W})$. Since $p_2 = rank(\Pi_{2W}) = m_2$ by assumption, we have $r_2 = rank(M_{Q_Z^{1/2}\Pi_{2W}}) = l - m_2$, so that $\mathrm{AR}(\theta_{01}) \xrightarrow{d} \frac{1}{l}\chi^2(l-m_2)$ for all values of r_1 as required.

Now, we can write $\mathrm{KLM}(\theta_{01})$ under H_0 as:

$$\mathrm{KLM}(\theta_{01}) = \frac{1}{\tilde{\sigma}_\varepsilon^2}\left(\tilde{\varepsilon}'Z/\sqrt{T}\right)\bar{\Pi}_{21}\left[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}_{21}'\left(Z'\tilde{\varepsilon}/\sqrt{T}\right), \tag{A.10}$$

where $\bar{\Pi}_{21} = \tilde{\Pi}_{21} - \tilde{\Pi}_{22}(\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{22})^{-1}\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{21}$.

Suppose first that $r_1 = rank(\Pi_{1W}) = m_1$. From Lemma B.2(1) of the Supplemental Appendix, we have $[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} (\Pi_{21}^{*'}Q_Z\Pi_{21}^*)^{-1}\Pi_{21}^{*'}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon}$ where $\Pi_{21}^* = Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Pi_{1W}$. So, we get

$$\begin{aligned}\mathrm{KLM}(\theta_{01}) \xrightarrow{d} \xi_2 &= \frac{1}{\sigma_\varepsilon^2}\psi_{Z\varepsilon}'Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}P_{M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Pi_{1W}}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon} \\ &= \| P_{\Omega(r_1)}.z_{w\varepsilon} \|^2, \quad \text{where} \quad \Omega(r_1) = M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Pi_{1W}. \end{aligned} \tag{A.11}$$

Since Π_{1W} has a full-column rank m_1, we can show from (A.11) that

$$\mathrm{KLM}(\theta_{01}) \xrightarrow{d} \xi_2 \sim \chi^2(m_1). \tag{A.12}$$

In addition, it is straightforward to see that ξ_1 and $l\xi_1 - \xi_2$ are independent each of another.

Suppose now that $r_1 = rank(\Pi_{1W}) < m_1$. From Lemma B.2(2) of the Supplemental Appendix, we have $\bar{\Upsilon}_T[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} (\Psi_{21}s'Q_Z\Psi_{21})^{-1}\Psi_{21}'M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon}$, where $\Psi_{21} = Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Psi_{1R}$ and $\Psi_{1R} = [\Pi_{1W}R_1, Q_Z^{-1}\psi_{ZV_1}R_2 - \sigma_\varepsilon^{-2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon}\sigma_{V_1\varepsilon}'R_2]$, so that we get

$$\begin{aligned}\mathrm{KLM}(\theta_{01}) \xrightarrow{d} \xi_2 &= \psi_{Z\varepsilon}'Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}P_{M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Psi_{1R}}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon} \\ &= \| P_{\Omega(r_1)}.z_{w\varepsilon} \|^2, \quad \text{where} \quad \Omega(r_1) = M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Pi_{1R}. \end{aligned} \tag{A.13}$$

First, we prove that $\Omega(r_1)$ is independent of $z_{w\varepsilon}$. For this, it is sufficient to show that any column $\psi_{V_1,j}$, $j = 1, \ldots, m_1$, of $\Psi_{V_1} = Q_Z^{-1}\psi_{ZV_1} - \frac{1}{\sigma_\varepsilon}z_{w\varepsilon}\sigma_{V_1\varepsilon}'$ is independent of $z_{w\varepsilon}$. However, it is straightforward to see that $cov(z_{w\varepsilon}, \psi_{V_1,j}) = \sigma_{V_1\varepsilon,j}(M_{Q_Z^{1/2}\Pi_{2W}} - M_{Q_Z^{1/2}\Pi_{2W}}) = 0$ and $z_{w\varepsilon}$ is independent of $\psi_{V_1,j}$ for all $j = 1, \ldots, m_1$, by the Gaussian Assumption 2.1- AS2. Second, because $\Omega(r_1)$ has a full-column rank m_1 with probability one, and further, is independent of $z_\varepsilon \sim N(0, I_l)$, it follows that the conditional distribution of $\| P_{\Omega(r_1)}.z_{w\varepsilon} \|^2$, given $\Omega(r_1)$, is $\chi^2(m_1)$. Since the conditional distribution, given $\Omega(r_1)$, does not depend on $\Omega(r_1)$, it is equal to the unconditional distribution, so, $\xi_2 \sim \chi^2(m_1)$ as stated.

The results of $\mathrm{JKLM}(\theta_{01})$ and $\mathrm{MQLR}(\theta_{01})$ follow by definition. ■

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

Proof of Theorem 3.3. The proof uses the results of Lemma B.3 of the Supplemental Appendix and follows similar steps as in the above proof of Theorem 3.2, therefore is omitted. ■

Proof of Theorem 3.4. The proof of Theorem 3.4 follows immediately from those of Theorem 3.1, therefore is omitted. ■

Proof of Theorem 3.5. Theorem 3.5 follows immediately from Theorem 3.3, therefore its proof is omitted. ■

Proof of Theorem 3.6. The proof of Theorem 3.6(a) is similar to those of Theorem 3.1(b) and is therefore omitted. We now prove Theorem 3.6(b). First, recall that we can write $\tilde{\theta}_2 - \theta_2$ under H_0 as:

$$\tilde{\theta}_2 - \theta_2 = (Y_2' P_Z Y_2 - \kappa_T Y_2' M_Z Y_2/T)^{-1}(Y_2' P_Z \varepsilon - \kappa_T Y_2' M_Z \varepsilon/T). \tag{A.14}$$

From Lemma B.5 in the Supplemental Appendix, we have $Y_2' P_Z Y_2 - \kappa_T Y_2' M_Z Y_2/T \xrightarrow{d} \Sigma_{V_2}^{1/2} \upsilon_1 \Sigma_{V_2}^{1/2} - \bar{\kappa}\Sigma_{V_2}$ and $Y_2' P_Z \varepsilon - \kappa_T Y_2' M_Z \varepsilon/T \xrightarrow{d} \sigma_\varepsilon \Sigma_{V_2}^{1/2} \upsilon_2 - \bar{\kappa}\sigma_{V_2\varepsilon}$, so that

$$\begin{aligned}\tilde{\theta}_2 - \theta_2 \xrightarrow{d} \bar{\theta}_2(\bar{\kappa}) &= \left(\Sigma_{V_2}^{1/2} \upsilon_1 \Sigma_{V_2}^{1/2} - \kappa \Sigma_{V_2}\right)^{-1} \left(\sigma_\varepsilon \Sigma_{V_2}^{1/2} \upsilon_2 - \bar{\kappa}\sigma_{V_2\varepsilon}\right) \\ &= \Sigma_{V_2}^{-1/2}(\upsilon_1 - \bar{\kappa} I_{m_2})^{-1} \sigma_\varepsilon (\upsilon_2 - \bar{\kappa}\rho_2)\end{aligned} \tag{A.15}$$

as required. ■

Proof of Theorem 3.7. The proof follows immediately from those of Theorems 3.2 and 3.3 and is omitted. ■

Proof of Theorem 3.7. Again, part (a) of Theorem 3.7 follows immediately from Theorem 3.3, so we omit the proof. We now prove part (b).

First, let us focus on $\mathrm{AR}(\theta_{01})$. Under H_0, we can write $\mathrm{AR}(\theta_{01})$ as

$$\begin{aligned}\mathrm{AR}(\theta_{01}) = \tilde{\varepsilon}' P_Z \tilde{\varepsilon}/l\tilde{\sigma}_\varepsilon^2 = \frac{1}{l\tilde{\sigma}_\varepsilon^2}\Big[&\varepsilon' P_Z \varepsilon - 2\varepsilon' P_Z Y_2(\tilde{\theta}_2 - \theta_2) \\ &+ (\tilde{\theta}_2 - \theta_2)' Y_2' P_Z Y_2 (\tilde{\theta}_2 - \theta_2)\Big],\end{aligned} \tag{A.16}$$

where $\tilde{\sigma}_\varepsilon^2 = \tilde{\varepsilon}' M_Z \tilde{\varepsilon}/(T-l) = \varepsilon'\varepsilon/(T-l) - \frac{T}{T-l}(\tilde{\varepsilon}' Z/T)\left(Z'Z/T\right)^{-1}(Z'\tilde{\varepsilon}/T) \xrightarrow{p} \sigma_\varepsilon^2$, because $Z'\varepsilon/T \xrightarrow{p} 0$, $Z'Y_2/T \xrightarrow{p} 0$, and $\tilde{\theta}_2 - \theta_2 \xrightarrow{d} \bar{\theta}_2(\bar{\kappa})$ by Theorem 3.6(b). So, it follows from the proof of Theorem 3.6(b) that

$$\begin{aligned}\mathrm{AR}(\theta_{01}) \xrightarrow{d} \xi_1^L(\bar{\kappa}) &= \frac{1}{l\sigma_\varepsilon^2}\left[\sigma_\varepsilon^2 z_\varepsilon' z_\varepsilon - 2\sigma_\varepsilon \upsilon_2' \Sigma_{V_2}^{1/2} \bar{\theta}_2(\bar{\kappa}) + \bar{\theta}_2(\bar{\kappa})' \Sigma_{V_2}^{1/2} \upsilon_1 \Sigma_{V_2}^{1/2} \bar{\theta}_2(\bar{\kappa})\right] \\ &= \frac{1}{l} \parallel z_\varepsilon - z_\kappa \parallel^2,\end{aligned} \tag{A.17}$$

where $z_\kappa = (\lambda_2 + z_{V_2})(\upsilon_1 - \kappa I_{m_2})^{-1}(\upsilon_2 - \kappa\rho_2)$.

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

We shall now prove the result of KLM(θ_{01}). First, note that the same denominator of KLM(θ_{01}) is $\tilde{\sigma}_\varepsilon^2$ and from the proof of the result of AR(θ_{01}), we have $\tilde{\sigma}_\varepsilon^2 \xrightarrow{p} \sigma_\varepsilon^2 > 0$. Based on this, we can write KLM(θ_{01}) as:

$$\begin{aligned}\text{KLM}(\theta_{01}) &= \frac{1}{\sigma_\varepsilon^2}\tilde{\varepsilon}' P_{Z\bar{\Pi}_{21}}\tilde{\varepsilon} + o_p(1)\\ &= \frac{1}{\sigma_\varepsilon^2}(\tilde{\varepsilon}'Z/\sqrt{T})\sqrt{T}\bar{\Pi}_{21}(\sqrt{T}\bar{\Pi}_{21}'(Z'Z/T)\sqrt{T}\bar{\Pi}_{21})^{-1}\sqrt{T}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T})\\ &\quad + o_p(1), \end{aligned} \tag{A.18}$$

where $\bar{\Pi}_{21} = \tilde{\Pi}_{21} - \tilde{\Pi}_{22}(\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{22})^{-1}\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{21}$. Now under H_0, we have

$$\begin{aligned} &Z'\tilde{\varepsilon}/\sqrt{T} \xrightarrow{d} \sigma_\varepsilon Q_Z^{1/2}(z_\varepsilon - z_\kappa),\ \tilde{\sigma}_{\varepsilon Y_i}(\theta_{01}) \xrightarrow{d} \sigma_{\varepsilon V_i},\ i = 1, 2,\\ &\sqrt{T}\tilde{\Pi}_{21} \xrightarrow{d} \Pi_{21}(\kappa) = C_1 + Q_Z^{-1}Q_{ZW}D_1 + \sigma_\varepsilon^{-1}Q_Z^{-1/2}(z_\varepsilon - z_\kappa)\sigma_{V_1\varepsilon}'\\ &\sqrt{T}\tilde{\Pi}_{22} \xrightarrow{d} \Pi_{22}(\kappa) = C_2 + Q_Z^{-1}Q_{ZW}D_2 + \sigma_\varepsilon^{-1}Q_Z^{-1/2}(z_\varepsilon - z_\kappa)\sigma_{V_2\varepsilon}'. \end{aligned}$$

So, we find $\sqrt{T}\bar{\Pi}_{21} \xrightarrow{d} \bar{\Pi}_{21}(\kappa) = \Pi_{21}(\kappa) - \Pi_{22}(\kappa)[\Pi_{22}'(\kappa)Q_Z\Pi_{22}(\kappa)]^{-1}\Pi_{22}'(\kappa)$ $Q_Z\Pi_{21}(\kappa) = Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{22}(\kappa)}Q_Z^{1/2}\Pi_{21}(\kappa)$. It follows immediately that

$$\begin{aligned}\text{KLM}(\theta_{01}) \xrightarrow{d} \xi_2^L(\kappa) &= (z_\varepsilon - z_\kappa)' P_{M_{Q_Z^{1/2}\Pi_{22}(\kappa)}Q_Z^{1/2}\Pi_{21}(\kappa)}(z_\varepsilon - z_\kappa)\\ &= \| P_{\Omega(\bar{\kappa})}.(z_\varepsilon - z_\kappa) \|^2, \end{aligned} \tag{A.19}$$

where $\Omega(\bar{\kappa}) = M_{Q_Z^{1/2}\Pi_{22}(\kappa)}Q_Z^{1/2}\Pi_{21}(\kappa)$. The results of JKLM($\theta_{01}$) and MQLR($\theta_{01}$) follow by definition. ■

B. Supplemental appendix

This supplemental appendix contains two main sections. Section B.1 presents the coordinate rotation introduced in Section 3.1 of the main text. The auxiliary lemmas used in the proof of the main theorems, as well as their proofs, are presented in Section B.2. The notations and definitions are the same as in the main text.

B.1. Rotation of the coordinates

To allow for partial identification of model parameters where the reduced-form matrices $[\Pi_{21}' : \Phi_1']'$ and $[\Pi_{22}' : \Phi_2']'$ [in Eqs. (2.4) and (2.5) of the main text] may be fixed with arbitrary ranks, we consider the coordinates rotations in the spaces of Y_2 and Y_1 as in Phillips (1989) and Choi and Phillips (1992).

Let $\Pi_{2W} = \Pi_{22} + Q_Z^{-1}Q_{ZW}\Phi_2$ and $\Pi_{1W} = \Pi_{21} + Q_Z^{-1}Q_{ZW}\Phi_1$, and define $rank(\Pi_{2W}) = p_2$, $rank(\Pi_{1W}) = r_1$; where Q_Z and Q_{ZW} are given in Assumption 2.1. Let $\mathcal{O}(m_2)$ and $\mathcal{O}(m_1)$ denote the orthogonal groups of $m_2 \times m_2$ and $m_1 \times m_1$ matrices, and consider the following two elements of $\mathcal{O}(m_2)$ and $\mathcal{O}(m_1)$:

$$S = [S_1, S_2] \in \mathcal{O}(m_2), \qquad R = [R_1, R_2] \in \mathcal{O}(m_1) \tag{B.1}$$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

such that: $S_1 : m_2 \times p_2,\ S_2 : m_2 \times (m_2 - p_2),\ R_1 : m_1 \times r_1,\ R_2 : m_1 \times (m_1 - r_1)$, S_2 and R_2 span the null space of $[\Pi'_{22} : \Phi'_2]'$ and $[\Pi'_{21} : \Phi'_1]'$ respectively, i.e., $[\Pi'_{22} : \Phi'_2]' S_2 = 0$ and $[\Pi'_{21} : \Phi'_1]' R_2 = 0$. We shall assume in (B.1) that the matrix S_j or R_j, $j = 1, 2$, is simply not present if its number of columns is equal to zero. For example, if $p_2 = m_2$, S_2 is not present in (B.1) and θ_2 is completely identified and we set $S = S_1 = I_{m_2}$. By the same way, if $p_2 = 0$, θ_2 is completely unidentified and S_1 vanishes so that $S \equiv S_2$. The same applies to the R_j matrices.

Now, since from (B.1) we have $SS' = I_{m_2}$ and $RR' = I_{m_1}$, we can write models (2.1) and (2.5) in the main text as:

$$\begin{aligned} y &= Y_1 RR'\theta_1 + Y_2 SS'\theta_2 + u, \\ &= Y_{11}\theta_{11} + Y_{12}\theta_{12} + Y_{21}\theta_{21} + Y_{22}\theta_{22} + u, \end{aligned} \tag{B.2}$$

$$Y_{11} = Z\tilde{\Pi}_{21} + W\tilde{\Phi}_1 + V_{11}, \quad Y_{12} = V_{12}, \tag{B.3}$$

$$Y_{21} = Z\tilde{\Pi}_{22} + W\tilde{\Phi}_2 + V_{21}, \quad Y_{22} = V_{22}, \tag{B.4}$$

where $\tilde{\Pi}_{22} = \Pi_{22} S_1,\ \tilde{\Phi}_2 = \Phi_2 S_1,\ \tilde{\Pi}_{21} = \Pi_{21} R_1,\ \tilde{\Phi}_1 = \Phi_1 R_1,\ \theta_{21} = S'_1\theta_2,\ \theta_{22} = S'_2\theta_2,\ \theta_{11} = R'_1\theta_1,\ \theta_{12} = R'_2\theta_1,\ V_{21} = V_2 S_1,\ V_{22} = V_2 S_2,\ V_{11} = V_1 R_1,\ V_{12} = V_1 R_2,\ Y_{21} = Y_2 S_1,\ Y_{22} = Y_2 S_2,\ Y_{11} = Y_1 R_1$, and $Y_{12} = Y_1 R_2$.

The parameter vectors θ_{21} and θ_{11} in (B.2) and (B.3) represent the linear combinations of θ_2 and θ_1 that are identified, while θ_{22} and θ_{12} are those that are not. The original coefficients θ_1 and θ_2 can then be recovered from (B.2) and (B.3) by the identities:

$$\theta_1 = R_1\theta_{11} + R_2\theta_{12}, \qquad \theta_2 = S_1\theta_{21} + S_2\theta_{22}. \tag{B.5}$$

B.2. Auxiliary lemmata and their proofs

This subsection contains the auxiliary lemmas and their proofs.

B.2.1. Auxiliary lemmata

LEMMA B.1. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [\Phi_1, \Phi_2]$ *is fixed. If further* $\theta_1 = \theta_{01}$, *then we have* $T\tilde{\kappa} \xrightarrow{d} \kappa = O_p(1)$ *for all values of* Π_2, *where* κ *is the smallest solution of* $|\Xi - \kappa\Lambda| = 0$ *with*

$$\Xi^0_{\Phi\Pi} = \begin{bmatrix} \psi'_{Z\varepsilon} Q_Z^{-1} \psi_{Z\varepsilon} & \psi'_{Z\varepsilon} \Pi_{2W} S_1 & \psi'_{Z\varepsilon} Q_Z^{-1} \psi_{ZV_2} S_2 \\ S'_1 \Pi'_{2W} \psi_{Z\varepsilon} & S'_1 \Pi'_{2W} Q_Z \Pi_{2W} S_1 & S'_1 \Pi'_{2W} \psi_{ZV_2} S_2 \\ S'_2 \psi'_{ZV_2} Q_Z^{-1} \psi_{Z\varepsilon} & S'_2 \psi'_{ZV_2} \Pi_{2W} S_1 & S'_2 \psi'_{ZV_2} Q_Z^{-1} \psi_{ZV_2} S_2 \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} \sigma^2_\varepsilon & 0 & \sigma'_{V_2\varepsilon} S_2 \\ 0 & 0 & 0 \\ S'_2 \sigma_{V_2\varepsilon} & 0 & S'_2 \Sigma_{V_2} S_2 \end{bmatrix}$$

and S_j, $j = 1, 2$, *are defined in* (B.1).

LEMMA B.2. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [\Phi_1, \Phi_2]$ *is fixed. Assume further that* $p_2 = rank(\Pi_{2W}) = m_2$ *and define* $\bar{\Upsilon}_T = diag(I_{r_1}, \sqrt{T} I_{m_1 - r_1})$, *where* $\Pi_{2W} = \Pi_{22} + Q_Z^{-1} Q_{ZW} \Phi_2$, I_{r_1} *and* $I_{m_1 - r_1}$ *are identity matrices of order* r_1 *and* $m_1 - r_1$, *respectively. Then under* H_0, *we have*:

(1) $(Z'Z/T)^{-1/2} Z'\tilde{\varepsilon}/\sqrt{T} \xrightarrow{d} M_{Q_Z^{1/2}\Pi_{2W}} Q_Z^{-1/2} \psi_{Z\varepsilon}$ *for all values of* $r_1 = rank(\Pi_{1W})$, *where* $\Pi_{1W} = \Pi_{21} + Q_Z^{-1} Q_{ZW} \Phi_1$;

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

(2) $\left[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} Q_\pi^* M_{Q_Z^{1/2}\Pi_{2W}} Q_Z^{-1/2}\psi_{Z\varepsilon}$ *when* $r_1 = m_1$ *and* $\bar{\Upsilon}_T\left[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} \Psi_\pi M_{Q_Z^{1/2}\Pi_{2W}} Q_Z^{-1/2}\psi_{Z\varepsilon}$ *when* $r_1 < m_1$, *where* $\bar{\Pi}_{21} = (Z'Z)^{-1/2}M_{(Z'Z)^{1/2}\tilde{\Pi}_{22}}(Z'Z)^{1/2}\tilde{\Pi}_{21}$, $Q_\pi^* = (\Pi_{21}^{*'}Q_Z\Pi_{21}^*)^{-1}\Pi_{21}^{*'}Q_Z^{1/2}$, $\Psi_\pi = (\Psi_{21}'Q_Z\Psi_{21})^{-1}\Psi_{21}'Q_Z^{1/2}$, $\Pi_{21}^* = Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Pi_{1W}$, $\Psi_{1R} = \left[\Pi_{1W}R_1, Q_Z^{-1}\psi_{ZV_1}R_2 - \sigma_\varepsilon^{-2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon}\sigma_{V_1\varepsilon}'R_2\right]$, *and* $\Psi_{21} = Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Psi_{1R}$.

LEMMA B.3. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [\Phi_1, \Phi_2]$ *is fixed. Assume further that* $p_2 = rank(\Pi_{2W}) < m_2$ *and* H_0 *holds. Then, we have*:

(1) $(Z'Z/T)^{-1/2}Z'\tilde{\varepsilon}/\sqrt{T} \xrightarrow{d} Q_Z^{-1/2}\left(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa)\right)$ *for all values of* r_1;

(2) $\left[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} (\Psi_{21}^{*'}Q_Z\Psi_{21}^*)^{-1}\Psi_{21}^{*'}\left(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa)\right)$ *when* $r_1 = m_1$ *and* $\bar{\Upsilon}_T\left[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} (\tilde{\Psi}_{21}'Q_Z\tilde{\Psi}_{21})^{-1}\tilde{\Psi}_{21}'\left(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa)\right)$ *when* $r_1 < m_1$, *where* $\bar{\Upsilon}_T$ *is defined in Lemma B.2,* $\Psi_{21}^* = Q_Z^{-1/2}M_{Q_Z^{1/2}\Psi_{2S}}Q_Z^{1/2}\Pi_{1W}$, $\Psi_{2S} = \left[\Pi_{2W}S_1, Q_Z^{-1}\psi_{ZV_2}S_2\right] - \sigma_\varepsilon^{-2}(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa))\left[0, \sigma_{V_2\varepsilon}'S_2\right]$, $\tilde{\Psi}_{21} = Q_Z^{-1/2}M_{Q_Z^{1/2}\Psi_{2S}}Q_Z^{1/2}\tilde{\Psi}_{1R}$, $\tilde{\Psi}_{1R} = \left[\Pi_{1W}R_1, Q_Z^{-1}\psi_{ZV_1}R_2\right] - \sigma_\varepsilon^{-2}(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa))\left[0, \sigma_{V_1\varepsilon}'R_2\right]$, *and finally* $\zeta_{\pi S} = [Q_Z\Pi_{2W}S_1, \psi_{ZV_2}S_2]$.

LEMMA B.4. *Suppose Assumption* 2.1 *is satisfied and* $\Phi = [D_1/\sqrt{T}, \Phi_2^* + D_2/\sqrt{T}]$ *such that* $rank(\Pi_{22} + Q_Z^{-1}Q_{ZW}\Phi_2^*) = p_2 < m_2$ *for some fixed matrices* $D_1 \in \mathbb{R}^{l_1\times m_1}$ *and* $D_2 \in \mathbb{R}^{l_1\times m_2}$ (*possibly zeros*). *If further* $\Phi_2^* \neq 0$ *is fixed, then Lemmas B.1 and B.3 hold with* Π_{2W} *replaced by* $\Pi_{2W}^* = \Pi_{22} + Q_Z^{-1}Q_{ZW}\Phi_2^*$ *and* Π_{1W} *by* $\Pi_{1W}^* = \Pi_{21}$.

LEMMA B.5. *Suppose Assumption* 2.1 *and* H_0 *are satisfied. Let* $\Phi = [D_1/\sqrt{T}, D_2/\sqrt{T}]$ *for some fixed matrices* $D_1 \in \mathbb{R}^{l_1\times m_1}$ *and* $D_2 \in \mathbb{R}^{l_1\times m_2}$ (*possibly zeros*). *Then the following convergence holds*:

(a) *if* $\Pi_2 = [\Pi_{21} : \Pi_{22}] \neq 0$ *is fixed with* $rank(\Pi_{22}) = p_2 < m_2$, *then Lemmas B.1 and B.3 hold with* $\Pi_{2W} = \Pi_{22} + T^{-1/2}Q_Z^{-1}Q_{ZW}D_2$ *and* $\Pi_{1W} = \Pi_{21} + T^{-1/2}Q_Z^{-1}Q_{ZW}D_1$;

(b) *if* $\Pi_2 = [\Pi_{21} : \Pi_{22}] = [C_1/\sqrt{T} : C_2/\sqrt{T}]$ *where* $C_1 \in \mathbb{R}^{l\times m_1}$ *and* $C_2 \in \mathbb{R}^{l\times m_2}$ *are fixed matrices* (*possibly zeros*), *then*

$T\tilde{\kappa} \xrightarrow{d} \bar{\kappa}$ *where* $\bar{\kappa}$ *is the smallest root of* $|\Xi - \kappa\bar{\Sigma}_2| = 0$, $\Xi = \begin{pmatrix} z_\varepsilon z_\varepsilon' & (\lambda_2 + z_{V_2}) \\ (\lambda_2 + z_{V_2})' & v_1 \end{pmatrix}$, *and* $\bar{\Sigma}_2 = \begin{pmatrix} 1 & \rho_2' \\ \rho_2 & I_{m_2} \end{pmatrix}$, v_1, v_2, ρ_2, z_ε, λ_2, *and* z_{V_2} *are defined in* (3.4)–(3.6).

B.2.2. Proofs of auxiliary lemmata

Proof of Lemma B.1. Suppose Assumption 2.1 is satisfied and Φ is fixed. Let $\tilde{\kappa}$ be the smallest root of the determinantal equation

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

$$\left|\bar{Y}_0' P_Z \bar{Y}_0 - \tilde{\kappa} \bar{Y}_0' M_Z \bar{Y}_0\right| = \left|\bar{Y}_0' P_Z \bar{Y}_0 - \kappa_T (\bar{Y}_0' M_Z \bar{Y}_0 / T)\right| = 0, \tag{B.6}$$

where $\bar{Y}_0 = [y - Y_1\theta_{01} : Y_2]$ and $\kappa_T = T\tilde{\kappa}$. For any nonsingular $(m_2+1)\times(m_2+1)$ matrix J, the roots of (B.6) are also those of $|J'\bar{Y}_0' P_Z \bar{Y}_0 J - \kappa_T J'(\bar{Y}_0' M_Z \bar{Y}_0/T)J| = 0$, hence by choosing J as

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}, \quad J_{11} = 1, \quad J_{12} = 0 : 1 \times m_2, J_{21} = -\theta_2, J_{22} = I_{m_2} \tag{B.7}$$

we have $\bar{Y}_0 J = [\varepsilon, Y_2]$ under H_0 and it is clear that κ_T is the smallest root of

$$\left|(\varepsilon, Y_2)' P_Z (\varepsilon, Y_2) - \kappa_T (\varepsilon, Y_2)' M_Z (\varepsilon, Y_2)/T\right| = 0. \tag{B.8}$$

Now, let $\Upsilon_T^* = diag(1, \sqrt{T} I_{p_2}, I_{m_2-p_2}) : (m_2+1)\times(m_2+1)$ and $\tilde{S} = diag(1, S) : (m_2+1)\times(m_2+1)$, where S is the orthogonal matrix defined in (B.1). Since Υ_T^* and $\tilde{S}$ are nonsingular, pre- and post-multiply (B.8) by $\Upsilon_T^{*-1}\tilde{S}'$ and $\tilde{S}\Upsilon_T^{*-1}$, respectively, do not change the roots. Which means that κ_T is the smallest root of

$$\left|\Upsilon_T^{*-1}\tilde{S}'(\varepsilon, Y_2)' P_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^{*-1} - \kappa_T \Upsilon_T^{*-1}\tilde{S}'(\varepsilon, Y_2)' M_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^{*-1}/T\right| = 0, \tag{B.9}$$

where both terms in the determinantal equation (B.9) can be written explicitly as:

$$\begin{aligned}
\Upsilon_T^{*-1}\tilde{S}'(\varepsilon, Y_2)' P_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^{*-1} &= \begin{bmatrix} \varepsilon' P_Z \varepsilon & \frac{\varepsilon' P_Z Y_{21}}{\sqrt{T}} & \varepsilon' P_Z Y_{22} \\ \frac{Y_{21}' P_Z \varepsilon}{\sqrt{T}} & \frac{Y_{21}' P_Z Y_{21}}{T} & \frac{Y_{21}' P_Z Y_{22}}{\sqrt{T}} \\ Y_{22}' P_Z \varepsilon & \frac{Y_{22}' P_Z Y_{21}}{\sqrt{T}} & Y_{22}' P_Z Y_{22} \end{bmatrix} \\
\frac{1}{T}\Upsilon_T^{*-1}\tilde{S}'(\varepsilon, Y_2)' M_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^{*-1} &= \begin{bmatrix} \frac{\varepsilon' M_Z \varepsilon}{T} & \frac{\varepsilon' M_Z Y_{21}}{T^{3/2}} & \frac{\varepsilon' M_Z Y_{22}}{T} \\ \frac{Y_{21}' M_Z \varepsilon}{T^{3/2}} & \frac{Y_{21}' M_Z Y_{21}}{T^2} & \frac{Y_{21}' M_Z Y_{22}}{T^{3/2}} \\ \frac{Y_{22}' M_Z \varepsilon}{T} & \frac{Y_{22}' M_Z Y_{21}}{T^{3/2}} & \frac{Y_{22}' M_Z Y_{22}}{T} \end{bmatrix}.
\end{aligned} \tag{B.10}$$

From (B.1) and if further Assumption 2.1 and H_0 hold, (B.10) gives $\Upsilon_T^{*-1}\tilde{S}'(\varepsilon, Y_2)' P_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^{*-1} \xrightarrow{p} \Xi_{\Phi\Pi}^0$ and $\frac{1}{T}\Upsilon_T^{*-1}\tilde{S}'(\varepsilon, Y_2)' M_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^{*-1} \xrightarrow{p} \Lambda$ where

$$\Xi_{\Phi\Pi}^0 = \begin{bmatrix} \psi_{Z\varepsilon}' Q_Z^{-1}\psi_{Z\varepsilon} & \psi_{Z\varepsilon}'\Pi_{2W} S_1 & \psi_{Z\varepsilon}' Q_Z^{-1}\psi_{ZV_2} S_2 \\ S_1'\Pi_{2W}'\psi_{Z\varepsilon} & S_1'\Pi_{2W}' Q_Z \Pi_{2W} S_1 & S_1'\Pi_{2W}'\psi_{ZV_2} S_2 \\ S_2'\psi_{ZV_2}' Q_Z^{-1}\psi_{Z\varepsilon} & S_2'\psi_{ZV_2}'\Pi_{2W} S_1 & S_2'\psi_{ZV_2}' Q_Z^{-1}\psi_{ZV_2} S_2 \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \sigma_{V_2\varepsilon}' S_2 \\ 0 & 0 & 0 \\ S_2'\sigma_{V_2\varepsilon} & 0 & S_2'\Sigma_{V_2} S_2 \end{bmatrix},$$

$\Pi_{2W} = \Pi_{22} + Q_Z^{-1} Q_{ZW}\Phi_2$. It follows immediately that $\kappa_T \xrightarrow{d} \kappa$ jointly with $\Upsilon_T^*\tilde{S}'(\varepsilon, Y_2)' P_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^* - \kappa_T \Upsilon_T^*\tilde{S}'(\varepsilon, Y_2)' M_Z (\varepsilon, Y_2)\tilde{S}\Upsilon_T^*/T \xrightarrow{d} \Xi_0 - \kappa\Lambda$, and the

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

latter convergence is uniform in κ over compact sets. Thus the roots of (B.9) converge to the roots of

$$\left|\Xi^0_{\Phi\Pi} - \kappa\Lambda\right| = 0. \tag{B.11}$$

Therefore, κ is the smallest solution of (B.11). Furthermore, it is clear that $\kappa = O_p(1)$ because $\Xi^0_{\Phi\Pi} = O_p(1)$ and $\Lambda = O_p(1)$. ■

Proof of Lemma B.2. Without loss of generality, we prove the results for Π_2 fixed.

Under H_0, we have $\tilde{\varepsilon} = y - Y_1\theta_{01} - Y_2\tilde{\theta}_2 = \varepsilon - Y_2(\tilde{\theta}_2 - \theta_2)$ so that we can write $(Z'Z/T)^{-1/2}Z'\tilde{\varepsilon}/\sqrt{T}$ as:

$$\frac{(Z'Z/T)^{-1/2}Z'\tilde{\varepsilon}}{\sqrt{T}} = (Z'Z/T)^{-1/2}\left\{Z'\varepsilon/\sqrt{T} - (Z'Y_2/T)L_T(\kappa_T)^{-1}\mathbf{N}_T(\kappa_T)\right\}, \tag{B.12}$$

where $\kappa_T \xrightarrow{d} \kappa$ by Lemma B.1, and because $\mathbf{L}_T(\kappa_T) = Y_2'P_ZY_2/T - \frac{1}{T}\kappa_T Y_2'M_ZY_2/T \xrightarrow{p} \Pi_{2W}'Q_Z\Pi_{2W} > 0$ and $\mathbf{N}_T(\kappa_T) = Y_2'P_Z\varepsilon/\sqrt{T} - \frac{1}{T}\kappa_T Y_2'M_Z\varepsilon/T \xrightarrow{d} \Pi_{2W}'\psi_{Z\varepsilon}$. Thus we get

$$(Z'Z/T)^{-1/2}Z'\tilde{\varepsilon}/\sqrt{T} \xrightarrow{d} M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon} \tag{B.13}$$

for all values of $r_1 = rank(\Pi_{1W})$, so that part (1) holds. To complete the proof of part (2), we now derive the asymptotic distribution of $[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}]^{-1/2}\bar{\Pi}_{21}'$, where $\bar{\Pi}_{21} = \tilde{\Pi}_{21} - \tilde{\Pi}_{22}(\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{22})^{-1}\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{21}$. Suppose first that $r_1 = rank(\Pi_{1W}) = m_1$. Then we have $\tilde{\Pi}_{21} \xrightarrow{p} \Pi_{1W}$ and $\tilde{\Pi}_{22} \xrightarrow{p} \Pi_{2W}$, where Π_{1W} and Π_{2W} have full-column ranks. It is then easy to see that $\bar{\Pi}_{21} \xrightarrow{p} \Pi^*_{21} = Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Pi_{1W}$, where Π^*_{21} also has full-column rank. It follows that

$$\left[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}_{21}'\left(Z'\tilde{\varepsilon}/\sqrt{T}\right) \xrightarrow{d} Q^*_\pi M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon},$$
$$Q^*_\pi = \left(\Pi^{*'}_{21}Q_Z\Pi^*_{21}\right)^{-1}\Pi^{*'}_{21}Q_Z^{1/2}.$$

Suppose now that $r_1 < m_1$. From Section B.1, we can write $\bar{\Upsilon}_T[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T})$ as:

$$\bar{\Upsilon}_T\left[\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T})$$
$$= R\left[\bar{\Upsilon}_T R'\bar{\Pi}_{21}'(Z'Z/T)\bar{\Pi}_{21}R\bar{\Upsilon}_T\right]^{-1}\bar{\Upsilon}_T R'\bar{\Pi}_{21}'(Z'\tilde{\varepsilon}/\sqrt{T}). \tag{B.14}$$

Because we have

$$\bar{\Pi}_{21}R\bar{\Upsilon}_T = \tilde{\Pi}_{21}R\bar{\Upsilon}_T - \tilde{\Pi}_{22}\left(\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{22}\right)^{-1}\tilde{\Pi}_{22}'(Z'Z)\tilde{\Pi}_{21}R\bar{\Upsilon}_T$$

where $\tilde{\Pi}_{21}R\bar{\Upsilon}_T = \left[(Z'Z/T)^{-1}Z'Y_1R_1/T, (Z'Z/T)^{-1}Z'Y_1R_2/\sqrt{T}\right]$

$$- \frac{1}{\tilde{\sigma}^2_{\tilde{\varepsilon}}}(Z'Z/T)^{-1}\frac{Z'\tilde{\varepsilon}}{\sqrt{T}}\left[\tilde{\varepsilon}'M_ZY_1R_1/T^{3/2}, \tilde{\varepsilon}'M_ZY_1R_2/T\right],$$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

and $\tilde{\Pi}_{22} \xrightarrow{p} \Pi_{2W}$ $(Z'Z/T)^{-1}Z'Y_1R_1/T \xrightarrow{p} \Pi_{1W}R_1$, $(Z'Z/T)^{-1}Z'Y_1R_2/\sqrt{T} \xrightarrow{d} Q_Z^{-1}\psi_{ZV_1}R_2$, $\tilde{\varepsilon}'M_ZY_1R_1/T^{3/2} \xrightarrow{p} 0$, $\tilde{\varepsilon}'M_ZY_1R_2/T \xrightarrow{p} \sigma'_{V_1\varepsilon}R_2$, and $\tilde{\sigma}_\varepsilon^2 \xrightarrow{p} \sigma_\varepsilon^2$, hence we have $\tilde{\Pi}_{21}R\bar{\Upsilon}_T \xrightarrow{d} \Psi_{1R} \equiv [\Pi_{1W}R_1, Q_Z^{-1}\psi_{ZV_1}R_2] - \sigma_\varepsilon^{-2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon}[0, \sigma'_{V_1\varepsilon}R_2]$, where Ψ_{1R} has full-column rank with probability one. Thus $\bar{\Pi}_{21}R\bar{\Upsilon}_T \xrightarrow{d} \Psi_{21} = Q_Z^{-1/2}M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{1/2}\Psi_{1R}$, where Ψ_{21} also has full-column rank with probability one. From (B.13), it follows that

$$\bar{\Upsilon}_T\left[\bar{\Pi}'_{21}(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}'_{21}(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} \Psi_\pi M_{Q_Z^{1/2}\Pi_{2W}}Q_Z^{-1/2}\psi_{Z\varepsilon}, \tag{B.15}$$

where $\Psi_\pi = (\Psi'_{21}Q_Z\Psi_{21})^{-1}\Psi'_{21}Q_Z^{1/2}$. ■

Proof of Lemma B.3. We also prove the results for Π_2 fixed. First, we can write $(Z'Z/T)^{-1/2}Z'\tilde{\varepsilon}/\sqrt{T}$ under H_0 as:

$$\begin{aligned}&\left(Z'Z/T\right)^{-1/2}Z'\tilde{\varepsilon}/\sqrt{T}\\ &\quad= \left(Z'Z/T\right)^{-1/2}\left[Z'\varepsilon/\sqrt{T} - \left(Z'Y_2S\Upsilon_T^{-1}/\sqrt{T}\right)\Upsilon_TS'(\tilde{\theta}_2-\theta_2)\right],\end{aligned} \tag{B.16}$$

where $(Z'Z/T)^{-1/2} \xrightarrow{p} Q_Z^{-1/2}$, $Z'\varepsilon/\sqrt{T} \xrightarrow{d} \psi_{Z\varepsilon}$, and $Z'Y_2S\Upsilon_T^{-1}/\sqrt{T} = \left[Z'Y_{21}/T, Z'Y_{22}/\sqrt{T}\right] \xrightarrow{d} \zeta_{\pi S} = \left[Q_Z\Pi_{2W}S_1, \psi_{ZV_2}S_2\right]$ from (B.1) to (B.4). Moreover, we have $\Upsilon_TS'(\tilde{\theta}_2-\theta_2) \xrightarrow{d} \theta_2^*(\kappa)$ by Theorem 3.1(b). Thus we get

$$\left(Z'Z/T\right)^{-1/2}Z'\tilde{\varepsilon}/\sqrt{T} \xrightarrow{d} Q_Z^{-1/2}\left(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa)\right) \tag{B.17}$$

for all values of $r_1 = rank(\Pi_{1W})$ and Lemma B.3(1) holds. From (B.17), we need the asymptotic distribution of $[\bar{\Pi}'_{21}(Z'Z/T)\bar{\Pi}_{21}]^{-1}\bar{\Pi}'_{21}$ to prove Lemma B.3(2), where $\tilde{\Pi}_{21} = \tilde{\Pi}_{21} - \tilde{\Pi}_{22}(\tilde{\Pi}'_{22}Z'Z\tilde{\Pi}_{22})^{-1}\tilde{\Pi}'_{22}(Z'Z)\tilde{\Pi}_{21}$.

Suppose first that $r_1 = m_1$, then we have $\tilde{\Pi}_{21} \xrightarrow{p} \Pi_{1W}$ with $rank(\Pi_{1W}) = m_1$. Moreover, we can write $\tilde{\Pi}_{22}(\tilde{\Pi}'_{22}Z'Z\tilde{\Pi}_{22})^{-1}\tilde{\Pi}'_{22}(Z'Z)$ as:

$$\begin{aligned}\tilde{\Pi}_{22}\left[\tilde{\Pi}'_{22}(Z'Z)\tilde{\Pi}_{22}\right]^{-1}\tilde{\Pi}'_{22}(Z'Z) &= \tilde{\Pi}_{22}S\Upsilon_T\left[\Upsilon_TS'\tilde{\Pi}'_{22}(Z'Z/T)\tilde{\Pi}_{22}S\Upsilon_T\right]^{-1}\\ &\quad\times\Upsilon_TS'\tilde{\Pi}'_{22}(Z'Z/T)\\ \text{where } \tilde{\Pi}_{22}S\Upsilon_T &= \left[(Z'Z/T)^{-1}Z'Y_2S_1/T, (Z'Z/T)^{-1}Z'Y_2S_2/\sqrt{T}\right]\\ &\quad-\frac{1}{\tilde{\sigma}_\varepsilon^2}(Z'Z/T)^{-1}\\ &\quad\times\frac{Z'\tilde{\varepsilon}}{\sqrt{T}}\left[\tilde{\varepsilon}'M_ZY_2S_1/T^{3/2}, \tilde{\varepsilon}'M_ZY_2S_2/T\right],\end{aligned} \tag{B.18}$$

where $\Upsilon_T = diag(I_{p_2}, \sqrt{T}I_{m_2-p_2})$ and I_{p_2} and $I_{m_2-p_2}$ are identity matrices of order p_2 and m_2-p_2, respectively. However, we can see from (B.1) to (B.4) that $\tilde{\Pi}_{22}S\Upsilon_T \xrightarrow{d}$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

$\Psi_{2S} \equiv [\Pi_{2W} S_1, Q_Z^{-1}\psi_{ZV_2} S_2] - \frac{1}{\sigma_\varepsilon^2}(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa))[0, \sigma'_{V_2\varepsilon} S_2]$ and Ψ_{2S} has full-column rank with probability one. So, we get $\tilde{\Pi}_{22}[\tilde{\Pi}'_{22}(Z'Z)\tilde{\Pi}_{22}]^{-1}\tilde{\Pi}'_{22}(Z'Z) \xrightarrow{d} \Psi_{2S}(\Psi'_{2S} Q_Z \Psi_{2S})^{-1}\Psi'_{2S} Q_Z$ and $\bar{\Pi}_{21} \xrightarrow{d} \Psi^*_{21} = Q_Z^{-1/2} M_{Q_Z^{1/2}\Psi_{2S}} Q_Z^{1/2}\Pi_{1W}$, where Ψ^*_{21} also has full-column rank with probability one. It follows immediately that

$$\left[\bar{\Pi}'_{21}(Z'Z/T)\bar{\Pi}_{21}\right)^{-1}\bar{\Pi}'_{21}(Z'\tilde{\varepsilon}/\sqrt{T}) \xrightarrow{d} (\Psi^{*'}_{21} Q_Z \Psi^*_{21}\Big]^{-1}\Psi^{*'}_{21}\left(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa)\right). \quad \textbf{(B.19)}$$

Suppose now that $r_1 < m_1$. From (B.14), we can write $\bar{\Upsilon}_T[\bar{\Pi}'_{21}(Z'Z/T)\bar{\Pi}_{21}]^{-1}\bar{\Pi}'_{21}(Z'\tilde{\varepsilon}/\sqrt{T})$ as:

$$\begin{aligned}&\bar{\Upsilon}_T\left[\bar{\Pi}'_{21}(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}'_{21}(Z'\tilde{\varepsilon}/\sqrt{T})\\ &\quad = R\left[\bar{\Upsilon}_T R'\bar{\Pi}'_{21}(Z'Z/T)\bar{\Pi}_{21} R\bar{\Upsilon}_T\right]^{-1}\bar{\Upsilon}_T R'\bar{\Pi}'_{21}\left(Z'\tilde{\varepsilon}/\sqrt{T}\right). \quad \textbf{(B.20)}\end{aligned}$$

Meanwhile, $\bar{\Pi}_{21} R\bar{\Upsilon}_T$ can also be written as:

$$\bar{\Pi}_{21} R\bar{\Upsilon}_T = \tilde{\Pi}_{21} R\bar{\Upsilon}_T - \tilde{\Pi}_{22}(\tilde{\Pi}'_{22}(Z'Z)\tilde{\Pi}_{22})^{-1}\tilde{\Pi}'_{22}(Z'Z)\tilde{\Pi}_{21} R\bar{\Upsilon}_T$$

and $\tilde{\Pi}_{22}[\tilde{\Pi}'_{22}(Z'Z)\tilde{\Pi}_{22}]^{-1}\tilde{\Pi}'_{22}(Z'Z) \xrightarrow{d} \Psi_{2S}(\Psi'_{2S} Q_Z \Psi_{2S})^{-1}\Psi'_{2S} Q_Z$ from (B.18) and (B.19). Since we also have $\frac{Z'\tilde{\varepsilon}}{\sqrt{T}} \xrightarrow{d} \psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa)$, it is easy to see [like in the proof of Lemma B.2(2)] that $\tilde{\Pi}_{21} R\bar{\Upsilon}_T \xrightarrow{d} \tilde{\Psi}_{1R} = [\Pi_{1W} R_1, Q_Z^{-1}\psi_{ZV_1} R_2] - \sigma_\varepsilon^{-2}(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa))[0, \sigma'_{V_1\varepsilon} R_2]$ and $\bar{\Pi}_{21} R\bar{\Upsilon}_T \xrightarrow{d} \tilde{\Psi}_{21} = Q_Z^{-1/2} M_{Q_Z^{1/2}\Psi_{2S}} Q_Z^{1/2}\tilde{\Psi}_{1R}$, where $\tilde{\Psi}_{1R}$ and $\tilde{\Psi}_{21}$ have full-column ranks with probability one, respectively. Therefore, we find

$$\begin{aligned}\bar{\Upsilon}_T\left[\bar{\Pi}'_{21}(Z'Z/T)\bar{\Pi}_{21}\right]^{-1}\bar{\Pi}'_{21}(Z'\tilde{\varepsilon}/\sqrt{T}) &\xrightarrow{d} (\tilde{\Psi}'_{21} Q_Z \tilde{\Psi}_{21})^{-1}\\ &\quad \times \tilde{\Psi}'_{21}\left(\psi_{Z\varepsilon} - \zeta_{\pi S}\theta_2^*(\kappa)\right). \quad \textbf{(B.21)}\end{aligned}$$

■

Proof of Lemma B.4. The proof of Lemma B.4 is similar to those of Lemmas B.1 and B.3 and is therefore omitted. ■

Proof of Lemma B.5. The proof of Lemma B.5(a) is similar to those of Lemmas B.1 and B.3 and is therefore omitted. We now prove part (b). We know from (B.8) that $\kappa_T = T\tilde{\kappa}$ is the smallest root of the determinantal equation

$$|B_T(\kappa_T)| = 0, \quad \textbf{(B.22)}$$

where $B_T(\kappa_T) = (\varepsilon, Y_2)' P_Z(\varepsilon, Y_2) - \kappa_T(\varepsilon, Y_2)' M_Z(\varepsilon, Y_2)/T$. Since we have $\Pi_2 = [C_1/\sqrt{T}, C_2/\sqrt{T}]$ and $\Phi = [D_1/\sqrt{T}, D_2/\sqrt{T}]$, then we have $Y_2' P_Z Y_2 \xrightarrow{d} \Sigma_{V_2}^{1/2} \upsilon_1 \Sigma_{V_2}^{1/2}$, $Y_2' M_Z Y_2/T \xrightarrow{p} \Sigma_{V_2}$, $Y_2' P_Z \varepsilon \xrightarrow{d} \sigma_\varepsilon \Sigma_{V_2}^{1/2}\upsilon_2$, $Y_2' M_Z \varepsilon/T \xrightarrow{p} \sigma'_{\varepsilon V_2}$, $\varepsilon' P_Z \varepsilon \xrightarrow{d} \sigma_\varepsilon^2 z'_\varepsilon z_\varepsilon$, $\varepsilon' M_Z \varepsilon/T \xrightarrow{p} \sigma_\varepsilon^2$ where $\upsilon_1 = (\lambda_2 + z_{V_2})'(\lambda_2 + z_{V_2})$, $\upsilon_2 = (\lambda_2 + z_{V_2})' z_\varepsilon$, $\lambda_2 = (Q_Z^{1/2} C_2 +$

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press

$Q_Z^{-1/2} Q_{ZW} D_2) \Sigma_{V_2}^{-1/2}$, $z_{V_2} = Q_Z^{-1/2} \psi_{ZV_2} \Sigma_{V_2}^{-1/2}$, and $z_\varepsilon = \sigma_\varepsilon^{-1} Q_Z^{-1/2} \psi_{Z\varepsilon}$. Therefore, $\kappa_T \xrightarrow{d} \kappa$ jointly with

$$B_T(\kappa_T) \xrightarrow{d} B(\kappa) = \begin{bmatrix} \sigma_\varepsilon^2 z_\varepsilon' z_\varepsilon & \sigma_\varepsilon \upsilon_2' \Sigma_{V_2}^{1/2} \\ \sigma_\varepsilon \Sigma_{V_2}^{1/2} \upsilon_2 & \Sigma_{V_2}^{1/2} \upsilon_1 \Sigma_{V_2}^{1/2} \end{bmatrix} - \kappa \begin{bmatrix} \sigma_\varepsilon^2 & \sigma_{V_2\varepsilon}' \\ \sigma_{V_2\varepsilon} & \Sigma_{V_2} \end{bmatrix}$$
$$= \Gamma'(\Xi - \kappa \bar{\Sigma}_2)\Gamma, \qquad \textbf{(B.23)}$$

where $\bar{\Sigma}_2 = \begin{bmatrix} 1 & \rho_2' \\ \rho_2 & I_{m_2} \end{bmatrix}$, $\Gamma = diag(\sigma_\varepsilon, \Sigma_{V_2}^{1/2})$, and $\rho_2 = \sigma_\varepsilon^{-1} \Sigma_{V_2}^{-1/2} \sigma_{V_2\varepsilon}$. Furthermore, the convergence of $B_T(\kappa_T)$ is uniform in κ over compact sets, hence the solution of $|\mathbf{B}_T(\kappa_T)| = 0$, converge to the solution of $|\Xi - \kappa \bar{\Sigma}_2| = 0$. Thus $\kappa_T \xrightarrow{d} \bar{\kappa}$, where $\bar{\kappa}$ is the smallest root of $|\Xi - \kappa \bar{\Sigma}_2| = 0$. ■

https://doi.org/10.1017/S0266466614000462 Published online by Cambridge University Press