Weak identification and confidence sets for covariances between errors and endogenous regressors: finite-sample and asymptotic theory*

Firmin Doko †
Université de Montréal

Jean-Marie Dufour ‡
McGill University

First version: September 2007
Revised: May 2008, November 2008
This version: March 2008
Compiled: March 30, 2009, 6:11pm

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

† Université de Montréal, Département de Sciences Économiques, C.P. 6128, succ. Centre-ville, Montréal, QC H3C 3J7. TEL: (1) 514 343 6111 ext. 5283; Fax: (1) 514 343 7221; e-mail : f.doko.tchatoka@umontreal.ca.

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

ABSTRACT

In this paper, we focus on structural models and propose a finite-and large-sample projection-based techniques for building confidence sets for the endogeneity parameter between errors and regressors allowing for the presence of weak identification. First, we show that our procedure is robust to weak identification and provide analytic forms of the confidence sets for endogeneity parameter. Second, we provide necessary and sufficient conditions under which such confidence sets are bounded in both finite-and large-sample. Finally, after formulating a general asymptotic framework which allows one to take into account a possible heteroskedasticity and/or autocorrelation of model residuals, we show that our procedure is asymptotically robust to these problems (heteroskedasticity and/or autocorrelation).

Key words: Anderson-Rubin test statistic; projection-based techniques; confidence sets; endogeneity parameter; weak identification.
JEL classification: C3; C12; C15; C52.

Contents

List of Tables

List of Assumptions, Propositions and Theorems

1. Introduction

Inference and confidence sets construction about covariance matrices are important problems in econometrics. For example, one may wish to build confidence sets on some linear transformation of the covariance matrix between errors and regressors. However, standard specification tests proposed by Durbin-Wu-Hausman (DWH) [see Durbin (1954), Wu (1973), Hausman (1978), Revankar and Hartley (1973)]; LR tests of Hwang (1980) and Smith (1984), and LM tests of Engle (1982) can not be used to produce confidence sets for covariances. In fact, all these procedures require a separate estimation for each null hypothesis tested so that it is difficult to construct confidence sets for the covariances of interest because covariance estimations or they standard errors are not typically produced.

Dufour (1987) proposed a generalized Wald-type procedure which allows the construction of confidence sets for linear transformations of covariance matrices under Gaussian distribution of model residuals. Doko and Dufour (2009) extend the setup in Dufour (1987) to non Gaussian distributions. However, the procedure in both Dufour (1987) and Doko and Dufour (2009) is only valid in large-sample. Furthermore, the procedure is based on the assumption that the instruments are strong, *i.e.* model parameters are strongly identified. More precisely, the rank condition for identifiability of models parameters in simultaneous equations is assumed to be satisfied. But, in the last decade, more interest is put on problems caused by "weak instruments" – *i.e.* situations where "instruments" are poorly correlated with endogenous explanatory variables. This raises the question of the robustness and consistency of the procedures in Dufour (1987) and Doko and Dufour (2009) in presence of weak instruments. For example, if $[\text{rank}(\Pi_2) < G]$ in Dufour (1987, Assumption 4) and Doko and Dufour (2009, eq. (2.10)), then the matrices in Dufour (1987, eq. (3.6)) and Doko and Dufour (2009, eq. (3.4)) have not full columns rank. In consequence, the key results in Dufour (1987, Theorem 1) and Doko and Dufour (2009, Theorem 3.1) are misleading in presence of weak instruments (identification). So, an interesting question is: can we propose a valid procedure for building confidence sets for covariances even in presence of weak identification?

In this paper, we focus on structural models and propose a finite-and large-sample projection-based techniques [Dufour (1997), Abdelkhalek and Dufour (1998), Dufour and Jasiak (2001)] for building confidence sets for the endogeneity parameter between errors and regressors allowing for the presence of weak instruments. We distinguish between two setups: the finite-sample setup and the asymptotic one. In both setups, we provide analytic forms of the confidence sets for endogeneity parameter and give necessary and sufficient conditions under which these confidence sets are bounded. Moreover, we show that our procedure is robust to weak identification in both setups. Furthermore, after formulating a general asymptotic framework which allows one to take into account a possible heteroskedasticity and/or autocorrelation of model residuals, we propose modified Anderson and Rubin (1949, AR) statistics which are asymptotically robust to these problems, this means that our projection-based procedure is asymptotically robust to heteroskedasticity and/or autocorrelation. Further, following Andrews (1991) and Andrews (1992), we propose HAC-type estimators which estimate consistently the covariance matrix of model residuals.

The paper is organized as follows. Section 2 formulates the model considered. Section 3 presents the finite-sample theory and Section 4, the asymptotic theory. We conclude in Section

5. Proofs are presented in the Appendix.

2. Model

We consider the common simultaneous equation model described by the following assumptions:

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

where y is a $T \times 1$ vector of observations on the dependent variable, Y is a $T \times G$ matrix of observations on the explanatory endogenous variables $(G \geq 1)$, X_1 is a $T \times k_1$ matrix of observations on the included exogenous variables, $u = \begin{bmatrix} u_1 & \ldots & u_T \end{bmatrix}'$ is a vector of structural disturbances, β and γ are $G \times 1$ and $k_1 \times 1$ vectors of unknown coefficients.

Let

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

where X_2 is a $T \times k_2$ matrix of observations on the excluded exogenous variables, Π_1 and Π_2 are $k_1 \times G$ and $k_2 \times G$ matrices of unknown coefficients, $V = \begin{bmatrix} V_1 & \ldots & V_T \end{bmatrix}'$ is a $T \times G$ matrix of disturbances.

Assume that

$$X = [X_1 : X_2] \in \mathbb{R}^{T\times k} \text{ has full-column rank} \tag{2.3}$$

where $k = k_1 + k_2$. The usual necessary and sufficient condition for identification of this

$$\mathrm{rank}(\Pi_2) = G\,. \tag{2.4}$$

If $\mathrm{rank}(\Pi_2) < G$, then β is not identified and the instruments X_2 are weak.

The reduced form for $[y, Y]$ can be written as

$$y = X_1\pi_1 + X_2\pi_2 + v\,, \tag{2.5}$$

$$Y = X_1\Pi_1 + X_2\Pi_2 + V\,, \tag{2.6}$$

where $\pi_1 = \gamma + \Pi_1\beta,\ \pi_2 = \Pi_2\beta,$ and $v = u + V\beta = [v_1, \ldots, v_T]'$. Let

$$M = M_X = I - X(X'X)^{-1}X', \quad M_1 = M_{X_1} = I - X_1(X_1'X_1)^{-1}X_1'. \tag{2.7}$$

Then, we can see that

$$M_1 - M = M_1X_2(X_2'M_1X_2)^{-1}X_2'M_1 = M_1MM_1. \tag{2.8}$$

Assume that

$$u = Va + \varepsilon\,, \tag{2.9}$$

where the $G \times 1$ vector of unknown coefficients a is the parameter which characterizes the endogeneity of $Y,\ \varepsilon$ is independent of V with mean zero and variance σ_ε^2.

Our main object in this paper is to build confidence sets for a. Note that neither u nor V are known, eq.(2.9) is not typically useful for achieving this goal. In what follows, we shall distinguish between two setups: finite-sample and asymptotic theories.

3. Finite-sample theory

In this section, we assume that

$$u \sim N\left(0,\, \sigma_u^2 I_T\right) \quad \text{is independent of } X\,. \tag{3.1}$$

where u is defined in (2.1). Let

$$U = [u,\, V] = [U_1,\, \ldots\, ,\, U_T]' \tag{3.2}$$

and suppose that the vectors $U_t = [u_t,\, V_t']'$, $t = 1,\, \ldots\, ,\, T$, have mean zero and the same nonsingular covariance matrix:

$$E[U_t] = \mu = [0,\, 0']'\,, \tag{3.3}$$

$$E[U_t U_t'] = \Sigma = \begin{bmatrix} \sigma_u^2 & \delta' \\ \delta & \Sigma_V \end{bmatrix} > 0\,, \quad t = 1,\, \ldots\, ,\, T\,. \tag{3.4}$$

From (2.9) and (3.1) - (3.4), we have

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2 I_T)\,, \quad t = 1,\, \ldots,\, T\,, \tag{3.5}$$

$$a = \Sigma_V^{-1}\delta\,, \quad \sigma_\varepsilon^2 = \sigma_u^2 - \delta'\Sigma_V^{-1}\delta\,. \tag{3.6}$$

We see from (3.6) that $a = 0$ iff $\delta = 0$, which means that the exogeneity of Y in model (2.1) - (2.2) can be assessed by testing whether $a = 0$ in (2.9).

Let us consider the joint hypothesis

$$H_0 : \beta = \beta_0,\, a = a_0\,, \tag{3.7}$$

where β_0 and a_0 are given $G \times 1$ vectors. We want to build confidence sets for a by using (3.7) since as mentioned before, neither u nor V are produced so that eq.(2.9) can be used to achieve this goal. To overcome such difficulty, we shall proceed in four steps:

1. build a confidence set for the structural parameter β;
2. build a confidence set for the transformed parameter $\theta = \beta + a$;
3. build a joint confidence set for (β, θ);
4. get a confidence set for $g(a)$ by using projection techniques [see Dufour (1997), Abdelkhalek and Dufour (1998), Dufour and Jasiak (2001)].

3.1. Inference on structural parameter

In this subsection, we consider the hypothesis

$$H_{\beta_0} : \beta = \beta_0 \,, \tag{3.8}$$

where β_0 is a given $G \times 1$ vectors. The Anderson and Rubin (1949, AR) test for H_{β_0} involves considering the transformed equation

$$y - Y\beta_0 = X_1\pi_1^0 + X_2\pi_2^0 + v^0 \,, \tag{3.9}$$

where $\pi_1^0 = \gamma + \mathit{\Pi}_1(\beta - \beta_0)$, $\pi_2^0 = \mathit{\Pi}_2(\beta - \beta_0)$ and $v^0 = u + V(\beta - \beta_0)$. If any restriction is imposed on γ (which is typically the case), H_{β_0} can then be assessed by testing $H_{\pi_2} : \pi_2^0 = 0$ using the standard F-statistic [say $AR(\beta_0)$]. Under H_{π_2}, we have

$$AR(\beta_0) = \frac{(y - Y\beta_0)'(M_1 - M)(y - Y\beta_0)/k_2}{(y - Y\beta_0)'M(y - Y\beta_0)/(T - k)} \,, \tag{3.10}$$

where the projection matrices M_1 and M are defined in (2.7). Under the assumption (3.1) and the null hypothesis H_{π_2}, we have

$$AR(\beta_0) \sim F(k_2, T - k) \,, \tag{3.11}$$

even if $\mathrm{rank}(\mathit{\Pi}_2) < G$. Which means that tests based on $AR(\beta_0)$ are robust to weak instruments. A confidence set for β with level $1 - \alpha_1$ is given by

$$C_\beta(\alpha_1) = \{\beta_0 : AR(\beta_0) \leq F_{\alpha_1}(k_2, T - k)\} \,, \tag{3.12}$$

where $F_{\alpha_1}(k_2, T-k)$ is the $1-\alpha_1$ quantile of the F distribution with $(k_2, T-k)$ degrees of freedom. From Dufour and Taamouti (2005), the inequality $AR(\beta_0) \leq F_{\alpha_1}(k_2, T - k)$ is equivalent to

$$\beta_0' A \beta_0 + b'\beta_0 + c \leq 0 \,, \tag{3.13}$$

where

$$A \;=\; Y'HY \,, \quad b = -2Y'Hy \,, \quad c = y'Hy \,, \tag{3.14}$$

$$H \;\equiv\; H_{AR} = M_1 - [1 + k_2 F_{\alpha_1}(k_2, T - k)/(T - k)]M \,. \tag{3.15}$$

So, (3.12) becomes

$$C_\beta(\alpha_1) = \left\{\beta_0 : \beta_0' A \beta_0 + b'\beta_0 + c \leq 0\right\} \,. \tag{3.16}$$

Note that the confidence set defined by (3.16) is a quadric confidence set.

3.2. Inference on a transformation of structural and endogeneity parameters

We now focus on the problem of testing the hypothesis

$$H_{\theta_0} : \theta = \theta_0 \,, \tag{3.17}$$

where $\theta = \beta + a$ and θ_0 is a given $G \times 1$ vectors. If we Substitute (2.9) into (2.1), we get

$$y = Y\beta + X_1\gamma + Va + \varepsilon\,, \tag{3.18}$$

where ε is defined in (2.9). Equation (3.18) illustrates that the existence of correlation between Y and u may be viewed as a problem of omitted variables. If the matrix V were observed, we would test any set of linear restriction on the coefficients β, γ and a in (3.18) by standard F-test, and these tests would be exact in finite-sample. In particular, linear hypotheses regarding the parameter a could be tested. Furthermore, if Σ_V were known, the transformation $\delta = \Sigma_V a$ would allow to test some linear restrictions on δ by standard F-test. The difficulty, of course, is that neither V nor Σ_V are known. This makes difficult the direct use of F-type tests in (3.18). To deal with such difficulty, we proceed as follow.

Let

$$\hat{\Pi} = (X'X)^{-1}X'Y\,, \tag{3.19}$$

be the OLS estimator of $\Pi = [\Pi_1'\,,\,\Pi_2']'$ from (2.2) and $\hat{V} = \hat{Y} - X\hat{\Pi}$ the corresponding residuals. If we replace the disturbance matrix V by $\hat{V}$ in (3.18), we have

$$y = Y\beta + X_1\gamma + \hat{V}a + \varepsilon^*\,, \tag{3.20}$$

where $\varepsilon^* = X(\hat{\Pi} - \Pi)a + \varepsilon$. If we estimate (3.19) by OLS, the estimators $\hat{\beta}$ and $\hat{\gamma}$ obtained are the 2SLS estimators of β and γ in (3.18) - (2.2) [see Dufour (1987)]. Furthermore, $Y = \hat{Y} + \hat{V}$, hence, (3.20) can also be written as

$$y = \hat{Y}\beta + X_1\gamma + \hat{V}\theta + \varepsilon^*\,, \tag{3.21}$$

where $\theta = \beta + a$. Note that the estimation of (3.21) using ordinary least squares method provides estimators of $\hat{\theta}$, $\hat{\beta}$ and $\hat{\gamma}$ for θ, β and γ. Further, the estimators $\hat{\beta}$ and $\hat{\gamma}$ obtained are the 2SLS estimators of β and γ in (3.18) - (2.2). It is worthwhile to note that instead of replacing V by $\hat{V}$ in (3.18), one may replace V in by $V = Y - X_1\Pi_1 - X_2\Pi_2$. By doing so, (3.18) becomes

$$y = Y\theta + X_1(\gamma - \Pi_1 a) + X_2(-\Pi_2 a) + \varepsilon \tag{3.22}$$

or equivalently

$$y - Y\theta_0 = Y\psi + X_1\pi_1^* + X_2\pi_2^* + \varepsilon \tag{3.23}$$

where $\psi = \theta - \theta_0$, $\pi_1^* = \gamma - \Pi_1 a$, $\pi_2^* = -\Pi_2 a$ and the disturbance vector ε is independent of all the regressors. So, H_{θ_0} can be assessed by testing whether $\psi = 0$ in (3.22). We also observe that $\pi_1^* = \gamma - \Pi_1 a$ and $\pi_2^* = -\Pi_2 a$. Hence, if any restriction is imposed on γ, a is identifiable if and only if $rank(\Pi_2) = G$, which is also the necessary and sufficient condition for identification of β. We conclude that a is not identified when we have weak instruments. So, using standard Wald-type statistics to build confidence sets for a or some transformations of a may lead to inaccurate results.

In what follows, we test H_{θ_0} by using (3.22). Let

$$Z = [Y\,,\,X_1\,,\,X_2]\,, \quad \bar{M} \equiv M_Z = I - Z(Z'Z)^{-1}Z'\,. \tag{3.24}$$

We can show that

$$\bar{M} = M - P_{MY} = M M_{MY} M \,, \quad M_{MY} = I - P_{MY} \,, \tag{3.25}$$
$$P_{MY} = MY(Y'MY)^{-1}Y'M \,, \tag{3.26}$$

where M is defined in (2.7). The Anderson and Rubin (1949, AR) test statistic for H_{θ_0} in (3.22) is given by

$$AR(\theta_0) = \frac{\varepsilon(\theta_0)' P_{MY} \varepsilon(\theta_0)/G}{\varepsilon(\theta_0)' \bar{M} \varepsilon(\theta_0)/(T-G-k)} \,, \tag{3.27}$$

where $\varepsilon(\theta_0) = (y - Y\theta_0)$. Under the assumption (3.1) and the null hypothesis $H(\theta_0)$, all the conditions of the classical linear model are satisfied and

$$AR(\theta_0) \sim F(G, T-G-k) \tag{3.28}$$

irrespective of the rank of the matrix Π_2, which means that tests based on $AR(\theta_0)$ are robust to weak instruments. A confidence set with level $1-\alpha_2$ for $\theta = \beta + a$ can be obtained by inverting the statistic $AR(\theta_0)$:

$$C_\theta(\alpha_2) = \{\theta_0 : AR(\theta_0) \leq F_{\alpha_2}(G, T-G-k)\} \;. \tag{3.29}$$

A simple algebraic calculation shows that

$$C_\theta(\alpha_2) = \left\{\theta_0 : \theta_0' \tilde{A} \theta_0 + \tilde{b}' \theta_0 + \tilde{c} \leq 0\right\} \;. \tag{3.30}$$

where

$$\tilde{A} = Y' \tilde{H} Y \,, \tilde{b}' = -2Y' \tilde{H} y \,, \tilde{c} = y' \tilde{H} y \,, \tag{3.31}$$
$$\begin{aligned} \tilde{H} &= M - [1 + G F_{\alpha_2}(G, T-G-k)/(T-G-k)] \bar{M} \\ &= P_{MY} - G F_{\alpha_2}(G, T-G-k)/(T-G-k) \bar{M} \,. \end{aligned} \tag{3.32}$$

Note that the confidence set in (3.30) is a quadric confidence set.

3.3. Joint inference on structural and endogeneity parameters

We focus in this subsection on the joint parameter (β, θ). Let consider the confidence sets $C_\beta(\alpha_1)$ and $C_\theta(\alpha_2)$ in the previous subsections. By Bonferroni's inequality, we have

$$P\left[\beta \in C_\beta(\alpha_1) \,, \, \theta \in C_\theta(\alpha_2)\right] \geq 1 - \alpha_1 - \alpha_2 \,. \tag{3.33}$$

If we choose $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$, then (3.33) becomes

$$P\left[\beta \in C_\beta(\alpha_1) \,, \, \theta \in C_\theta(\alpha_2)\right] \geq 1 - \alpha \tag{3.34}$$

So, the set

$$C_{(\beta,\theta)}(\alpha) = \{(\beta_0', \theta_0')' : \beta_0 \in C_\beta(\alpha_1)\ ,\ \theta_0 \in C_\theta(\alpha_2)\} \tag{3.35}$$

is the confidence set with level $1-\alpha$ for the joint parameter (β, θ). Note $C_{(\beta,\theta)}(\alpha)$ defined in (3.35) is simultaneous in the sense of Scheffé. From (3.16) and (3.30), $C_{(\beta,\theta)}(\alpha)$ can be expressed:

$$C_{(\beta,\theta)}(\alpha) = \{(\beta_0', \theta_0')' : \beta_0' A \beta_0 + b'\beta_0 + c \leq 0\ ,\ \theta_0' \tilde{A} \theta_0 + \tilde{b}'\theta_0 + \tilde{c} \leq 0\}\,. \tag{3.36}$$

It is worthwhile to note that even though the simultaneous confidence sets $C_\beta(\alpha_1)$ and $C_\theta(\alpha_2)$ may be interpreted as a confidence sets based on inverting LR-type tests or as a profile likelihood confidence sets [see Meeker and al. (1998)], the confidence sets $C_{(\beta,\theta)}(\alpha)$ is not (strictly speaking) LR-type confidence sets. $C_{(\beta,\theta)}(\alpha)$ is obtained by taking the intersection of two quadrics.

3.4. Projection-based confidence sets for the endogeneity parameter

We study in this section, the problem of building confidence sets for the endogeneity parameter a or any transformation $g(a)$ of this parameter. The method used here to achieve this goal is projection techniques [see Dufour (1990), Dufour and Jan (1998), Dufour and Jasiak (2001)].

Let $\phi = (\beta', \theta')' \in \mathbb{R}^{2G}$ and consider the transformation $g(\phi) \in \mathbb{R}^G$. Since $\phi \in C_\phi(\alpha)$ entails $g(\phi) \in g[C_\phi(\alpha)]$, we have

$$P\{g(\phi) \in g[C_\phi(\alpha)]\} \geq P\{\phi \in C_\phi(\alpha)\} \geq 1 - \alpha\,, \tag{3.37}$$

where

$$g[C_\phi(\alpha)] = \{g(\phi) \in \mathbb{R}^G : \phi \in C_\phi(\alpha)\}\,. \tag{3.38}$$

We see that the confidence set $g[C_\phi(\alpha)]$ has level $1-\alpha$. Note that (3.37) holds for any transformation $g(.)$. Therefore, $g(.)$ may be discontinuous or non differentiable transformation. In particular when $g(\phi) = a$, $g[C_\phi(\alpha)]$ can be interpreted as the projection of $C_\phi(\alpha)$ on the space spanned by the columns of a. In this case, we have

$$g[C_\phi(\alpha)] = \{a \in \mathbb{R}^G : \phi \in C_\phi(\alpha)\} \equiv C_a(\alpha)\,. \tag{3.39}$$

So, we can get confidence sets for each component $a_j, j = 1, \ldots, G$ of a by minimizing and maximizing $a_j, j = 1, \ldots, G$ subject to the restriction $\phi \in C_\phi(\alpha)$. Let

$$a_{jL} = \inf\{a_j : \phi \in C_\phi(\alpha)\} \quad \text{and} \quad a_{jU} = \sup\{a_j : \phi \in C_\phi(\alpha)\}\,, \quad j = 1, \ldots, G\,. \tag{3.40}$$

Then, we have

$$P[a_{jL} \leq a_j \leq a_{jU}] \geq P[\phi \in C_\phi(\alpha)] \geq 1 - \alpha\,, \quad j = 1, \ldots, G\,. \tag{3.41}$$

Hence, $[a_{jL}, a_{jU}]$ is a confidence interval with level $1-\alpha$ for a_j, $j = 1, \ldots, G$. Furthermore, the confidence intervals $[a_{jL}, a_{jU}]$, $j = 1, \ldots, G$, are simultaneous at level $1-\alpha$, *i.e.* the corresponding G-dimensional confidence box contains the true vector $(a_1, \ldots, a_G)$ with probability (at least)

$1-\alpha$.

In what follows, we focus on the case where $g(\phi) = a$ and characterize the sets $C_a(\alpha)$ defined in (3.39). Let

$$\begin{aligned} \mathcal{Q}(\beta, a) &= \theta'\tilde{A}\theta + \tilde{b}'\theta + \tilde{c} = a'\tilde{A}a + \tilde{b}'a + \tilde{c} + \beta'\tilde{A}\beta + (2\tilde{A}a + \tilde{b})'\beta\,, && (3.42)\\ f(\beta) &= \beta' A\beta + b'\beta + c\,. && (3.43) \end{aligned}$$

Then, we can see that the confidence set $C_a(\alpha)$ in (3.39) is obtained by minimizing (3.42) subject to (3.43), *i.e.*

$$C_a(\alpha) = \left\{a \in \mathbb{R}^G : \; \mathcal{Q}^*(a) \leq 0\right\}\,, \qquad (3.44)$$

where

$$\mathcal{Q}^*(a) = \min_{\substack{\beta \\ s.c\ f(\beta)\leq 0}} \mathcal{Q}(\beta, a)\,. \qquad (3.45)$$

Notations

For any matrix Λ, Λ^- denotes any generalized inverse of Λ and $\mathcal{H}_\Lambda = \Lambda^-\Lambda$. Let

$$\begin{aligned} \bar{a} &= \frac{1}{2}\mathcal{H}_{\tilde{A}}A^-b - \mathcal{H}_{\tilde{A}}(I-\mathcal{H}_A)\beta_{0*} - \frac{1}{2}\tilde{A}^-\tilde{b} + (I-\mathcal{H}_{\tilde{A}})a_*\,, && (3.46)\\ \tilde{c}^* &= \tilde{c} - \frac{1}{4}\tilde{b}'\tilde{A}^-\tilde{b}' - \frac{1}{2}\beta_{0*}'(I-\mathcal{H}_A')\tilde{A}(I-\mathcal{H}_A)\beta_{0*}\,, && (3.47)\\ c^* &= c - \frac{1}{4}b'A^-b' + \frac{1}{2}b'(I-\mathcal{H}_A)\beta_{0*}\,, \bar{\beta} = -\frac{1}{2}A^-b + (I-\mathcal{H}_A)\beta_{0*}\,, && (3.48) \end{aligned}$$

where β_{0*} and a_* are any arbitrary vectors in $\mathbb{R}^G$, A, b, and c are defined in (3.14), and $\tilde{A}$, $\tilde{b}$, $\tilde{c}$ are defined in (3.31). Let also defined

$$\begin{aligned} \Delta_{\lambda_1} &= \tilde{A} + \tilde{A}\tilde{\mathcal{A}}\tilde{A} - 2\tilde{A}A_{\lambda_1}^-\tilde{A}\,, && (3.49)\\ \tilde{b}_{\lambda_1} &= \tilde{b} + \frac{1}{2}\tilde{A}\tilde{\mathcal{A}}\tilde{A}\tilde{b} + \frac{\lambda_1}{2}\tilde{A}\tilde{\mathcal{A}}\tilde{A}b + \tilde{\mathcal{A}}\tilde{A}\tilde{b} + \frac{\lambda_1}{2}\tilde{A}\tilde{\mathcal{A}}b - 2\tilde{A}\tilde{\mathcal{A}}^*\beta_{30} - 2\tilde{A}A_{\lambda_1}^-\tilde{b}\\ &\quad -\lambda_1\tilde{A}A_{\lambda_1}^-b + 2\tilde{A}(I-\mathcal{H}_{A_{\lambda_1}})\beta_{30}\,, && (3.50)\\ \tilde{c}_{\lambda_1} &= \tilde{c} + \frac{\tilde{b}'\tilde{\mathcal{A}}\tilde{b}}{4} + \frac{3\lambda_1\tilde{b}'\tilde{\mathcal{A}}b}{4} + \frac{\lambda_1^2 b'\tilde{\mathcal{A}}b}{4} - \tilde{b}'\tilde{\mathcal{A}}\beta_{30} - \frac{3\lambda_1 b'\tilde{\mathcal{A}}^*\beta_{30}}{2} - \frac{\tilde{b}'A_{\lambda_1}^-\tilde{b}}{2}\\ &\quad -\frac{\lambda_1\tilde{b}'A_{\lambda_1}^-b}{4} + \tilde{b}'(I-\mathcal{H}_{A_{\lambda_1}})\beta_{30} + \beta_{30}'(I-\mathcal{H}_{A_{\lambda_1}}')\tilde{A}(I-\mathcal{H}_{A_{\lambda_1}})\beta_{30}\,, && (3.51)\\ \Gamma_{\lambda_1} &= \tilde{A}\mathcal{A}\tilde{A}, \quad b_{\lambda_1} = \frac{1}{2}\tilde{A}\mathcal{A}(I+\tilde{A})(\tilde{b}+b) - \tilde{A}A_{\lambda_1}^-b - 2\tilde{A}\mathcal{A}^*\beta_{30}\,, && (3.52)\\ c_{\lambda_1} &= c + \frac{\tilde{b}'\mathcal{A}\tilde{b}}{4} + \frac{3\lambda_1\tilde{b}'\mathcal{A}b}{4} + \frac{\lambda_1^2 b'\mathcal{A}b}{4} - \lambda_1 b'\mathcal{A}\beta_{30} - \tilde{b}'\mathcal{A}^*\beta_{30} + b'(I-\mathcal{H}_{A_{\lambda_1}})\beta_{30} \end{aligned}$$

$$+\beta_{30}'(I-\mathcal{H}_{A_{\lambda_1}}')A(I-\mathcal{H}_{A_{\lambda_1}})\beta_{30}-\frac{\lambda_1 b'A_{\lambda_1}^{-}b}{2}-\frac{b'A_{\lambda_1}^{-}\tilde{b}}{2}\,, \tag{3.53}$$

$$\tilde{\mathcal{A}} = A_{\lambda_1}\tilde{A}A_{\lambda_1}, \quad \mathcal{A}=A_{\lambda_1}AA_{\lambda_1}, \quad \tilde{\mathcal{A}}^* = A_{\lambda_1}\tilde{A}(I-\mathcal{H}_{A_{\lambda_1}})$$

$$\mathcal{A}^* = A_{\lambda_1}A(I-\mathcal{H}_{A_{\lambda_1}})\,, \tag{3.54}$$

β_{30} is any arbitrary $G\times 1$ vector and

$$\lambda_1=\min_{\lambda}\vartheta_\lambda,\ \ \vartheta_\lambda=\{\lambda : |\tilde{A}+\lambda A|=0\}\,. \tag{3.55}$$

Finally, we denote by

$$\Delta_\mu=[\tilde{b}'(I-\mathcal{H}_A)\beta_{10}]^2-4(\tilde{b}'A^{-}\tilde{b})c_0, \quad c_0=c-\frac{b'A^{-}b}{4}+\frac{b'(I-\mathcal{H}_A)\beta_{10}}{2}\,, \tag{3.56}$$

where β_{10} is any arbitrary $G\times 1$ vector. For any matrix $\mathcal{K}$, $\mathcal{K}\geq 0$ means that $\mathcal{K}$ is positive semidefinite (p.s.d.).

The sets $C_a(\alpha)$ take one of the following form of Theorem **3.1**.

Theorem 3.1 PROJECTION-BASED CONFIDENCE SETS FOR ENDOGENEITY. *Assume that the assumptions* (2.1) - (2.3), (3.1) - (2.9) *hold. If furthermore* $\beta=\beta_0$ *and* $a=a_0$, *where* β_0 *and* a_0 *are* $G\times 1$ *constant vectors, then the sets* $C_a(\alpha)$ *take one of the following forms* :

(I) *if* $2A\beta+b=0$, *then*

(I_1) *if* $\tilde{A}$ *is semidefinite positive and* $\tilde{A}\neq 0$,

$$C_a(\alpha)=\begin{cases}\{a\in\mathbb{R}^G : a=\bar{a}\} & \text{if} \quad \tilde{c}^*\leq 0 \quad \text{and} \quad c^*\leq 0\\ \emptyset \quad \text{otherwise}\,, & \end{cases} \tag{3.57}$$

(I_2) *if* $\tilde{A}=0$,

$$C_a(\alpha)=\begin{cases}\left\{a\in\mathbb{R}^G : \tilde{b}'a+\tilde{c}+\tilde{b}'\bar{\beta}\leq 0\right\} & \text{if} \quad c^*\leq 0\\ \emptyset \quad \text{otherwise}\,, & \end{cases} \tag{3.58}$$

(I_3) *if* $\tilde{A}$ *is not semidefinite positive,*

$$C_a(\alpha)=\begin{cases}\left\{a\in\mathbb{R}^G : a'\tilde{A}a+\bar{b}'a+\bar{c}\leq 0\right\} & \text{if} \quad c^*\leq 0\\ \emptyset \quad \text{otherwise} & \end{cases} \tag{3.59}$$

where $\bar{b}=\tilde{b}+2\tilde{A}\bar{\beta}$ *and* $\bar{c}=\tilde{c}+\bar{\beta}'\tilde{A}\bar{\beta}+\tilde{b}'\bar{\beta}$,

(II) *if* $2A\beta+b\neq 0$, *then*

(II_1) *if $\tilde{A}$ and A are semidefinite positive,*

$$C_a(\alpha) = \left\{ a \in \mathbb{R}^G : a'\Delta_{\lambda_1} a + \tilde{b}'_{\lambda_1} a + \tilde{c}_{\lambda_1} \leq 0, \text{ and } a'\Gamma_{\lambda_1} a + b'_{\lambda_1} a + c_{\lambda_1} \leq 0 \right\}, \quad (3.60)$$

(II_2) *if $\tilde{A} = 0$,*

$$C_a(\alpha) = \begin{cases} \emptyset \quad \text{if} \quad [\tilde{b} \neq 0] \text{ or } [\tilde{b} = 0,\, \tilde{c} > 0] \text{ or } [A = 0, b = 0, c > 0] \\ \quad \text{or } [A \text{ is positive definite and } d = \frac{1}{4} b' A^{-1} b - c < 0] \\ \quad \text{or } [\Delta_\mu < 0,\ A \geq 0] \\ \\ \left\{ a \in \mathbb{R}^G : \tilde{b}' a - [\tilde{b}' A^- \tilde{b}] \mu^+ + \tilde{c} - \frac{\tilde{b}' A^- b}{2} + \tilde{b}'(I - \mathcal{H}_A)\beta_{10} \leq 0 \right\} \\ \quad \text{if} \quad [\Delta_\mu \geq 0,\ A \geq 0] \\ \\ \mathbb{R}^G \quad \text{if } A \text{ is not positive semidefinite}, \end{cases} \quad (3.61)$$

where μ^+ is the positive solution of

$$[\tilde{b}' A^- \tilde{b}] \mu^2 - [\tilde{b}'(I - \mathcal{H}_A)\beta_{10}] \mu + c_0 = 0\,, \quad (3.62)$$

(II_3) *if $\tilde{A}$ or A is not semidefinite positive, $C_a(\alpha)$ is bounded if it is empty, otherwise, $C_a(\alpha)$ is unbounded.*

We can now deduce the necessary and sufficient conditions under which the sets $C_a(\alpha)$ are bounded. Corollary **3.2** below gives these conditions.

Corollary 3.2 N.S. CONDITIONS FOR BOUNDED SETS. *Assume that the assumptions of Theorem* **3.1**, *then $C_a(\alpha)$ defined above is bounded if and only if A and $\tilde{A}$ are positive definite. Otherwise, $C_a(\alpha)$ is bounded if it is empty.*

In the Appendix A, we use the results of Theorem **3.1** to show Corollary **3.2**. However, Corollary **3.2** can be showed by considering the following five cases:

(i) all the eigenvalues of A and $\tilde{A}$ are positive ($\lambda_i > 0$, $\tilde{\lambda}_i > 0$; $i = 1, \ldots, G$), that is, A and $\tilde{A}$ are positive definite,

(ii) one of the matrices A and $\tilde{A}$ is positive definite and the other is negative definite; that is [($\lambda_i > 0$ and $\tilde{\lambda}_i < 0$) or vis-versa; $i = 1, \ldots, G$],

(iii) one of the matrices A and $\tilde{A}$ is positive (or negative) definite and the other is neither positive nor negative definite,

(iv) all the eigenvalues of A and $\tilde{A}$ are negative ($\lambda_i < 0$, $\tilde{\lambda}_i < 0$; $i = 1, \ldots, G$), that is, A and $\tilde{A}$ are negative definite,

(v) A and $\tilde{A}$ have both positive and negative eigenvalues, that is, A and $\tilde{A}$ are neither positive nor negative definite.

case-(i) : A and $\tilde{A}$ are positive definite matrices, hence If $\lambda_i > 0$, $\tilde{\lambda}_i > 0$; $i = 1, \ldots, G$. The inequalities (3.16) and (3.30) become

$$\left(\frac{z_1}{\gamma_1}\right)^2 + \left(\frac{z_2}{\gamma_2}\right)^2 + \ldots + \left(\frac{z_G}{\gamma_G}\right)^2 \leq d\,, \tag{3.63}$$

$$\left(\frac{\tilde{z}_1}{\tilde{\gamma}_1}\right)^2 + \left(\frac{\tilde{z}_2}{\tilde{\gamma}_2}\right)^2 + \ldots + \left(\frac{\tilde{z}_G}{\tilde{\gamma}_G}\right)^2 \leq \tilde{d}\,, \tag{3.64}$$

where $\gamma_i = \sqrt{1/\lambda_i}$ and $\tilde{\gamma}_i = \sqrt{1/\tilde{\lambda}_i}$, $i = 1, \ldots, G$, $z = P(\beta - \tilde{\beta})$, $\tilde{z} = \tilde{P}(\theta - \tilde{\theta})$, P and $\tilde{P}$ are orthogonal matrices such that

$$A = P'DP \quad \text{and} \quad \tilde{A} = \tilde{P}'\tilde{D}\tilde{P}\,, \tag{3.65}$$

$\tilde{\beta} = -\frac{1}{2}A^{-1}b$, $\tilde{\theta} = -\frac{1}{2}\tilde{A}^{-1}\tilde{b}$, $d = \frac{1}{4}b'A^{-1}b - c$, $\tilde{d} = \frac{1}{4}\tilde{b}'\tilde{A}^{-1}\tilde{b} - \tilde{c}$. So, if $d = \tilde{d} = 0$, then $C_{(\beta,\theta)} = \{(\tilde{\beta}, \tilde{\theta})\}$, thus $C_a = \{\tilde{\theta} - \tilde{\beta}\}$. If $d < 0$ or $\tilde{d} < 0$, $C_{(\beta,\theta)}$ is empty, hence C_a is empty. If $d > 0$ and $\tilde{d} > 0$, $C_{(\beta,\theta)}$ is the intersection of two ellipsoids. Hence, C_a is bounded or empty.

case-(ii) : one of the matrices A and $\tilde{A}$ is positive definite and the other is negative definite, let us say A is positive definite and $\tilde{A}$ is negative definite. That is, $\lambda_i > 0$ and $\tilde{\lambda}_i < 0$ for all $i = 1, \ldots, G$. In this case, (3.16) and (3.30) can be written as

$$\left(\frac{z_1}{\gamma_1}\right)^2 + \left(\frac{z_2}{\gamma_2}\right)^2 + \ldots + \left(\frac{z_G}{\gamma_G}\right)^2 \leq d\,, \tag{3.66}$$

$$\left(\frac{\tilde{z}_1}{\tilde{\gamma}_1}\right)^2 + \left(\frac{\tilde{z}_2}{\tilde{\gamma}_2}\right)^2 + \ldots + \left(\frac{\tilde{z}_G}{\tilde{\gamma}_G}\right)^2 \geq -\tilde{d}\,, \tag{3.67}$$

where $\gamma_i = \sqrt{1/\lambda_i}$ and $\tilde{\gamma}_i = \sqrt{-1/\tilde{\lambda}_i}$, $i = 1, \ldots, G$. If $d = \tilde{d} = 0$, then $C_{(\beta,\theta)} = \{(\tilde{\beta}, \tilde{\theta})\}$ and $C_a = \{\tilde{\theta} - \tilde{\beta}\}$. If $d < 0$, $C_{(\beta,\theta)}$ is empty, then C_a is also empty. If $d > 0$, $C_{(\beta,\theta)}$ is bounded if and only if it is empty and C_a is also bounded only if it is empty.

case-(iii) : We distinguish two subcases:

(iii-a) A is positive definite and $\tilde{A}$ is neither positive nor negative definite. Assume that $\tilde{\lambda}_j > 0$ for $j = 1, \ldots, g$ and $\tilde{\lambda}_j < 0$ for $j = g + 1, \ldots, G$. (3.16) and (3.30) may be written as

$$\left(\frac{z_1}{\gamma_1}\right)^2 + \left(\frac{z_2}{\gamma_2}\right)^2 + \ldots + \left(\frac{z_G}{\gamma_G}\right)^2 \leq d\,, \tag{3.68}$$

$$\left(\frac{\tilde{z}_1}{\tilde{\gamma}_1}\right)^2 + \left(\frac{\tilde{z}_2}{\tilde{\gamma}_2}\right)^2 + \ldots + \left(\frac{\tilde{z}_g}{\tilde{\gamma}_g}\right)^2 - \left(\frac{\tilde{z}_{g+1}}{\tilde{\gamma}_{g+1}}\right)^2 - \ldots - \left(\frac{\tilde{z}_G}{\tilde{\gamma}_G}\right)^2 \leq \tilde{d}\,, \qquad (3.69)$$

where $\gamma_i = \sqrt{1/\lambda_i}$, g is the number of positive eigenvalues of $\tilde{A}$, $\tilde{\gamma}_i = \sqrt{1/\tilde{\lambda}_i}$ for $j = 1, \ldots, g$, $\tilde{\gamma}_i = \sqrt{-1/\tilde{\lambda}_i}$ for $j = g+1, \ldots, G$. Clearly, it appears that $C_{(\beta,\theta)}$ is bounded only when it is empty. So, C_a is also bounded only when it is empty.

(iii-b) if A is negative definite and $\tilde{A}$ is neither positive nor negative definite, as in (iii-a), we have

$$\left(\frac{z_1}{\gamma_1}\right)^2 + \left(\frac{z_2}{\gamma_2}\right)^2 + \ldots + \left(\frac{z_G}{\gamma_G}\right)^2 \geq -d\,, \qquad (3.70)$$

$$\left(\frac{\tilde{z}_1}{\tilde{\gamma}_1}\right)^2 + \left(\frac{\tilde{z}_2}{\tilde{\gamma}_2}\right)^2 + \ldots + \left(\frac{\tilde{z}_g}{\tilde{\gamma}_g}\right)^2 - \left(\frac{\tilde{z}_{g+1}}{\tilde{\gamma}_{g+1}}\right)^2 - \ldots - \left(\frac{\tilde{z}_G}{\tilde{\gamma}_G}\right)^2 \leq \tilde{d}\,, \qquad (3.71)$$

where $\gamma_i = \sqrt{-1/\lambda_i}$ and g is the number of positive eigenvalues of $\tilde{A}$, $\tilde{\gamma}_i = \sqrt{1/\tilde{\lambda}_i}$ for $j = 1, \ldots, g$, $\tilde{\gamma}_i = \sqrt{-1/\tilde{\lambda}_i}$ for $j = g+1, \ldots, G$. So, $C_{(\beta,\theta)}$ is bounded if and only if it is empty. Thus, C_a is also bounded if it is empty set.

case-(iv) : A and $\tilde{A}$ are negative definite matrices. Hence, we have If $\lambda_i < 0$, and $\tilde{\lambda}_i < 0$; $i = 1, \ldots, G$. So, the inequalities (3.16) and (3.30) become

$$\left(\frac{z_1}{\gamma_1}\right)^2 + \left(\frac{z_2}{\gamma_2}\right)^2 + \ldots + \left(\frac{z_G}{\gamma_G}\right)^2 \geq -d\,, \qquad (3.72)$$

$$\left(\frac{\tilde{z}_1}{\tilde{\gamma}_1}\right)^2 + \left(\frac{\tilde{z}_2}{\tilde{\gamma}_2}\right)^2 + \ldots + \left(\frac{\tilde{z}_G}{\tilde{\gamma}_G}\right)^2 \geq -\tilde{d}\,, \qquad (3.73)$$

where $\gamma_i = \sqrt{-1/\lambda_i}$ and $\tilde{\gamma}_i = \sqrt{-1/\tilde{\lambda}_i}$, $i = 1, \ldots, G$. Since (3.72) and (3.73) hold as soon as any $|z_i|$ and $|\tilde{z}_i|$ are large enough, $C_{(\beta,\theta)}$ is unbounded. So, C_a is unbounded set. In particular, if $d \geq 0$ and $\tilde{d} \geq 0$, we have $C_{(\beta,\theta)} = \mathbb{R}^{2G}$ and $C_a = \mathbb{R}^G$.

case-(v) : A and $\tilde{A}$ are not positive or negative definite. Hence, A and $\tilde{A}$ have both positive and negative eigenvalues, say $\lambda_i > 0$ for $i = 1, \ldots, p$ and $\lambda_i < 0$ for $i = p+1, \ldots, G$, $\tilde{\lambda}_j > 0$ for $j = 1, \ldots, g$ and $\tilde{\lambda}_j < 0$ for $j = g+1, \ldots, G$. (3.16) and (3.30) may be written as

$$\left(\frac{z_1}{\gamma_1}\right)^2 + \left(\frac{z_2}{\gamma_2}\right)^2 + \ldots + \left(\frac{z_p}{\gamma_p}\right)^2 - \left(\frac{z_{p+1}}{\gamma_{p+1}}\right)^2 - \ldots - \left(\frac{z_p}{\gamma_p}\right)^2 \leq d\,, \qquad (3.74)$$

$$\left(\frac{\tilde{z}_1}{\tilde{\gamma}_1}\right)^2 + \left(\frac{\tilde{z}_2}{\tilde{\gamma}_2}\right)^2 + \ldots + \left(\frac{\tilde{z}_g}{\tilde{\gamma}_g}\right)^2 - \left(\frac{\tilde{z}_{g+1}}{\tilde{\gamma}_{g+1}}\right)^2 - \ldots - \left(\frac{\tilde{z}_G}{\tilde{\gamma}_G}\right)^2 \leq \tilde{d}\,, \qquad (3.75)$$

where p and g are the numbers of positive eigenvalues of A and $\tilde{A}$ respectively, $\gamma_i = \sqrt{1/\lambda_i}$

for $i = 1, \ldots, p$, $\gamma_i = \sqrt{-1/\lambda_i}$ for $i = p+1, \ldots, G$ and $\tilde{\gamma}_i = \sqrt{1/\tilde{\lambda}_i}$ for $j = 1, \ldots, g$, $\tilde{\gamma}_i = \sqrt{-1/\tilde{\lambda}_i}$ for $j = g{+}1, \ldots, G$. Then, for arbitrary given values $z_1, \ldots, z_p, \tilde{z}_1, \ldots, \tilde{z}_g$ and d, $\tilde{d}$, (3.74) and (3.75) will hold if any of the values z_i, z_j, $i = p+1, \ldots, G$ and $j = g+1, \ldots, G$, are small enough (as $|z_i| \to \infty$ and $|\tilde{z}_i| \to \infty$). Consequently, each component of $(z, \tilde{z})$ is unbounded and similarly for each component of (β, θ) . So, $C_{(\beta,\theta)}$ and C_a are unbounded.

Putting together these different cases gives the results of Corollary **3.2**.

4. Asymptotic theory

In this section, we extend the above procedure to large-sample setup. We assume that (2.1) - (2.3) hold. Let

$$u = Va + \varepsilon\,, \tag{4.1}$$

where we assume that

$$E[V'\varepsilon] = 0, \quad \text{and} \quad E[\varepsilon\varepsilon'] = \Sigma_\varepsilon = [\sigma_{ij}]_{1\leq i,j\leq T} > 0\,. \tag{4.2}$$

If $\Sigma_\varepsilon = \sigma_\varepsilon^2 I_T$, then the errors $\varepsilon_t,\ t = 1, \ldots, T$ are homoskedastic. If $\Sigma_\varepsilon = diag[\sigma_{11}, \sigma_{22}, \ldots, \sigma_{TT}]$, then $\varepsilon_t,\ t = 1, \ldots, T$ are heteroskedastic. And finally if

$$\Sigma_\varepsilon = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \gamma_\varepsilon(1) & \ldots & \gamma_\varepsilon(T-1) \\ \gamma_\varepsilon(1) & \ddots & \ldots & \vdots \\ \vdots & \ldots & \ddots & \vdots \\ \gamma_\varepsilon(T-1) & \ldots & \ldots & 1 \end{bmatrix},$$

then $\varepsilon_t,\ t = 1, \ldots, T$ are autocorrelated. Note that we can have both heteroskedasticity and autocorrelation, $\gamma_\varepsilon(.)$ is the Autocovariance Function of ε. From (4.1), the covariance matrix of u is given by

$$\Sigma_u = [\sigma_{ij}^u]_{1\leq i,j\leq T} > 0\ , \sigma_{ij}^u = a'\Sigma_V a + \sigma_{ij}\,, \tag{4.3}$$

where we assume that V as mean zero and covariance matrix Σ_V. Clearly, ε is the only source of heteroskedastic and/or autocorrelated of the structural errors u.

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

$$\frac{1}{T}X'\Sigma_u^{-1}u \xrightarrow{p} 0\,,\ \frac{1}{T}X'\Sigma_u^{-1}X \xrightarrow{p} \Omega_X > 0\,, \tag{4.4}$$

$$\frac{1}{T}Z'\Sigma_\varepsilon^{-1}\varepsilon \xrightarrow{p} 0\,,\ \frac{1}{T}Z'\Sigma_\varepsilon^{-1}Z \xrightarrow{p} \Omega_Z > 0\,, \tag{4.5}$$

$$\frac{1}{T}X'\Sigma_u^{-1}\Sigma_\varepsilon\Sigma_u^{-1}X \xrightarrow{p} \Delta_{X\varepsilon} > 0\,,\ \frac{1}{T}Z'\Sigma_\varepsilon^{-1}\Sigma_u\Sigma_\varepsilon^{-1}Z \xrightarrow{p} \Delta_{Zu} > 0\,, \tag{4.6}$$

$$\frac{1}{\sqrt{T}}X'\Sigma_u^{-1}[u,\,V,\,\varepsilon] \xrightarrow{L} [S_u^x,\,S_V^x,\,S_\varepsilon^x]\,, \tag{4.7}$$

$$\frac{1}{\sqrt{T}}Z'\Sigma_\varepsilon^{-1}[u,\,V,\,\varepsilon] \xrightarrow{L} [S_u^z,\,S_V^z,\,S_\varepsilon^z]\,, \tag{4.8}$$

$$\mathrm{vec}[S_u^x,\,S_V^x,\,S_\varepsilon^x] \sim N\left[0\,,\,\Sigma_{S^x}\right]\,,\mathrm{vec}[S_u^z,\,S_V^z,\,S_\varepsilon^z] \sim N\left[0\,,\,\Sigma_{S^z}\right]\,, \tag{4.9}$$

$$S_\varepsilon^x \text{ is independent with } S_V^x \text{ and } S_\varepsilon^z \text{ is independent with } S_V^z \tag{4.10}$$

$$S_u^x \sim N\left[0,\,\Omega_X\right]\,,\quad S_u^z \sim N\left[0,\,\Delta_{Zu}\right]\,, \tag{4.11}$$

$$S_\varepsilon^x \sim N\left[0,\,\Delta_{X\varepsilon}\right]\,,\quad S_\varepsilon^x \sim N\left[0,\,\Omega_Z\right]\,, \tag{4.12}$$

where $X = [X_1\,,\,X_2]$ and $Z = [Y\,,\,X_1\,,\,X_2]$.

Consider again equations (3.9) and (3.23). Since $\Sigma_u > 0$ and $\Sigma_\varepsilon > 0$, we can multiply (3.9) by $\Sigma_u^{-1/2}$ and (3.23) by $\Sigma_\varepsilon^{-1/2}$. By proceeding so, we get

$$\begin{aligned} \Sigma_u^{-1/2}(y - Y\beta_0) &= \Sigma_u^{-1/2}X_1\pi_1^0 + \Sigma_u^{-1/2}X_2\pi_2^0 + \Sigma_u^{-1/2}\upsilon^0\,, && (4.13)\\ &\textit{i.e.} \\ y_* - Y_*\beta_0 &= X_{*1}\pi_1^0 + X_{*2}\pi_2^0 + \upsilon_1 && (4.14)\end{aligned}$$

for equation (3.9) and

$$\begin{aligned} \Sigma_\varepsilon^{-1/2}(y - Y\theta_0) &= \Sigma_\varepsilon^{-1/2}Y\psi + \Sigma_\varepsilon^{-1/2}X_1\pi_1^* + \Sigma_\varepsilon^{-1/2}X_2\pi_2^* + \Sigma_\varepsilon^{-1/2}\varepsilon\,, && (4.15)\\ &\textit{i.e.} \\ y^* - Y_*\theta_0 &= Y^*\psi + X_1^*\pi_1^* + X_2^*\pi_2^* + \varepsilon_1 && (4.16)\end{aligned}$$

for equation (3.23), where $\psi = \theta - \theta_0$, $\pi_1^* = \gamma - \Pi_1 a$, $\pi_2^* = -\Pi_2 a$. For any fixed matrix Λ and any random matrix $\mathcal{Z}$, let

$$P_{\mathcal{Z}}(\Lambda) = \Lambda^{-1/2}\mathcal{Z}(\mathcal{Z}'\Lambda^{-1}\mathcal{Z})^{-1}\mathcal{Z}'\Lambda^{-1/2},\ M_{\mathcal{Z}}(\Lambda) = I - P_{\mathcal{Z}}(\Lambda)\,. \tag{4.17}$$

The Anderson and Rubin (1949, AR) test-statistics for testing the hypotheses $H_{\pi_2} : \pi_2^0 = 0$ and $H_\psi : \psi = 0$ respectively in (4.14) and (4.16) are given by

$$AR^{he}(\beta_0) = \frac{\upsilon^0(\beta_0)'\hat{\Sigma}_u^{-1/2}[P_X(\hat{\Sigma}_u) - P_{X_1}(\hat{\Sigma}_u)]\hat{\Sigma}_u^{-1/2}\upsilon^0(\beta_0)/k_2}{\upsilon^0(\beta_0)'\hat{\Sigma}_u^{-1/2}M_X(\hat{\Sigma}_u)\hat{\Sigma}_u^{-1/2}\upsilon^0(\beta_0)/(T - k_1 - k_2)}\,, \tag{4.18}$$

and

$$AR^{he}(\theta_0) = \frac{\varepsilon(\theta_0)'\hat{\Sigma}_\varepsilon^{-1/2}[P_Z(\hat{\Sigma}_\varepsilon) - P_X(\hat{\Sigma}_\varepsilon)]\hat{\Sigma}_\varepsilon^{-1/2}\varepsilon(\theta_0)/G}{\varepsilon(\theta_0)'\hat{\Sigma}_\varepsilon^{-1/2}M_Z(\hat{\Sigma}_\varepsilon)\hat{\Sigma}_\varepsilon^{-1/2}\varepsilon(\theta_0)/(T - k - G)}\,, \tag{4.19}$$

where $\upsilon^0(\beta_0) = y - Y\beta_0$ and $\varepsilon(\theta_0) = y - Y\theta_0$. Due to the structure of the errors (heteroskedasticity and/or autocorrelation), the covariance matrices Σ_u and Σ_ε can be estimate consistently by using the HAC estimators that have been proposed in the literature–see Levine (1983), White (1984, pp. 147-161), White and Domowitz (1984), Gallant (1987, pp. 533, 551, 573), Newey and West (1987), Andrews (1991), Andrews (1992). In this paper, we estimate Σ_u and Σ_ε as in Andrews (1991) and Andrews (1992) Econometrica papers.

Let

$$J_{\upsilon,T} = \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T} E(\upsilon_s^0 \upsilon_t^{0\prime}) = \sum_{j=-T+1}^{T-1} \gamma_\upsilon(j)\,, \tag{4.20}$$

$$\gamma_\upsilon(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} E(\upsilon_t^0 \upsilon_{t-j}^{0\prime}) & \text{for } j \geq 0 \\ \\ \frac{1}{T}\sum_{t=-j+1}^{T} E(\upsilon_{t+j}^0 \upsilon_t^{0\prime}) & \text{for } j < 0\,, \end{cases} \tag{4.21}$$

$$J_{\varepsilon,T} = \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T} E(\varepsilon_s \varepsilon_t^{\prime}) = \sum_{j=-T+1}^{T-1} \gamma_\varepsilon(j)\,, \tag{4.22}$$

$$\gamma_\varepsilon(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} E(\varepsilon_t \varepsilon_{t-j}^{\prime}) & \text{for } j \geq 0 \\ \\ \frac{1}{T}\sum_{t=-j+1}^{T} E(\varepsilon_{t+j} \varepsilon_t^{\prime}) & \text{for } j < 0\,. \end{cases} \tag{4.23}$$

We consider the following class of estimators of $J_{\upsilon,T}$ and $J_{\varepsilon,T}$

$$\hat{J}_{\upsilon,T} = \hat{J}_{\upsilon,T}(S_T) = \frac{T}{T-r}\sum_{j=-T+1}^{T-1} \kappa\left(\frac{j}{S_T}\right)\hat{\gamma}_\upsilon(j)\,, \tag{4.24}$$

$$\hat{\gamma}_\upsilon(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \hat{\upsilon}_t^0 \hat{\upsilon}_{t-j}^{0\prime} & \text{for } j \geq 0 \\ \\ \frac{1}{T}\sum_{t=-j+1}^{T} \hat{\upsilon}_{t+j}^0 \hat{\upsilon}_t^{0\prime} & \text{for } j < 0 \end{cases} \tag{4.25}$$

$$\hat{J}_{\varepsilon,T} = \hat{J}_{\varepsilon,T}(\tilde{S}_T) = \frac{T}{T-r}\sum_{j=-T+1}^{T-1} \tilde{\kappa}\left(\frac{j}{\tilde{S}_T}\right)\hat{\gamma}_\varepsilon(j)\,, \tag{4.26}$$

$$\hat{\gamma}_\varepsilon(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}^{\prime} & \text{for } j \geq 0 \\ \\ \frac{1}{T}\sum_{t=-j+1}^{T} \hat{\varepsilon}_{t+j} \hat{\varepsilon}_t^{\prime} & \text{for } j < 0 \end{cases} \tag{4.27}$$

where $\hat{\upsilon}_t^0 \equiv \hat{\upsilon}_t^0(\hat{\pi}_1^0, \hat{\pi}_2^0)$, $\hat{\varepsilon}_t = \hat{\varepsilon}_t(\hat{\psi},\, \hat{\pi}_1^*, \hat{\pi}_2^*)$, $\kappa(.)$ and $\tilde{\kappa}(.)$ are real-valued kernel and $S_T,\ \tilde{S}_T,$ are band-with parameters [see Andrews (1991) and Andrews (1992)]. So,

$$\hat{J}_{v,T} = \hat{\Sigma}_u \quad \text{and} \quad \hat{J}_{\varepsilon,T} = \hat{\Sigma}_\varepsilon\,. \tag{4.28}$$

are consistent estimators of Σ_u and Σ_ε, *i.e.*,

$$\underset{T\to\infty}{\text{plim}}\,(\hat{J}_{v,T}) = \underset{T\to\infty}{\text{plim}}\,(\hat{\Sigma}_u) = \Sigma_u \quad \text{and} \quad \underset{T\to\infty}{\text{plim}}\,(\hat{J}_{\varepsilon,T}) = \underset{T\to\infty}{\text{plim}}\,(\hat{\Sigma}_\varepsilon) = \Sigma_\varepsilon\,. \tag{4.29}$$

Lemma 4.1 ASYMPTOTIC DISTRIBUTIONS OF AR STATISTICS. *Under the assumptions* (2.1) - (2.3), (4.1) - (4.12), *if* $\beta = \beta_0$ *and* $\theta = \theta_0$ *where* β_0 *and* θ_0 *are* $G \times 1$ *constant vectors, then*

$$AR^{he}(\beta_0) \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2)\,, \tag{4.30}$$

$$AR^{he}(\theta_0) \xrightarrow{L} \frac{1}{G}\chi^2(G)\,, \tag{4.31}$$

irrespective of whether the instruments are weak or strong.

In Consequence, the asymptotic confidence sets for β and θ with level $1-\alpha_1$ and $1-\alpha_2$ are given by

$$C_\beta^\infty(\alpha_1) = \left\{\beta_0 : AR^{he}(\beta_0) \le \frac{1}{k_2}\chi^2_{\alpha_1}(k_2)\right\}\,, \tag{4.32}$$

$$C_\theta^\infty(\alpha_2) = \left\{\theta_0 : AR^{he}(\theta_0) \le \frac{1}{G}\chi^2_{\alpha_2}(G)\right\}\,, \tag{4.33}$$

where $\chi^2_{\alpha_1}(k_2)$ and $\chi^2_{\alpha_2}(G)$ are respectively the $1-\alpha_1$ and $1-\alpha_2$ quantiles of the χ^2 distributions with k_2 and G degrees of freedom. As in finite-sample setup, the confidence sets $C_\beta^\infty(\alpha_1)$ and $C_\theta^\infty(\alpha_2)$ can be written as

$$C_\beta^\infty(\alpha_1) = \{\beta_0 : \beta_0' A_h \beta_0 + b_h' \beta_0 + c_h \le 0\}\,, \tag{4.34}$$

$$C_\theta^\infty(\alpha_2) = \{\theta_0 : \theta_0' \tilde{A}_h \theta_0 + \tilde{b}_h' \theta_0 + \tilde{c}_h \le 0\} \tag{4.35}$$

where $A_h = Y'\hat{\Sigma}_u^{-1}H_\infty\hat{\Sigma}_u^{-1}Y$, $\tilde{A}_h = Y'\hat{\Sigma}_\varepsilon^{-1}\tilde{H}_\infty\hat{\Sigma}_\varepsilon^{-1}Y$, $b_h' = -2Y'\hat{\Sigma}_u^{-1}H_\infty\hat{\Sigma}_u^{-1}y$, $\tilde{b}_h' = -2Y'\hat{\Sigma}_\varepsilon^{-1}\tilde{H}_\infty\hat{\Sigma}_\varepsilon^{-1}y$, $c_h = y'\hat{\Sigma}_u^{-1}H_\infty\hat{\Sigma}_u^{-1}y$, $\tilde{c}_h = y'\hat{\Sigma}_\varepsilon^{-1}\tilde{H}_\infty\hat{\Sigma}_\varepsilon^{-1}y$, and

$$H_\infty = M_{X_1}(\hat{\Sigma}_u) - [1 + \frac{1}{T-k}\chi^2_{\alpha_1}(k_2)]M_X(\hat{\Sigma}_u),$$

$$\tilde{H}_\infty = M_X(\hat{\Sigma}_\varepsilon) - [1 + \frac{1}{T-G-k}\chi^2_{\alpha_2}(G)]M_Z(\hat{\Sigma}_\varepsilon)\,.$$

By Bonferroni inequality,

$$C^\infty_{(\beta,\theta)}(\alpha) = \{(\beta_0', \theta_0')' : \beta_0' A_h \beta_0 + b_h' \beta_0 + c_h \le 0\ ,\ \theta_0' \tilde{A}_h \theta_0 + \tilde{b}_h' \theta_0 + \tilde{c}_h \le 0\}\,, \tag{4.36}$$

is an asymptotic confidence set with level $1-\alpha$ for the joint parameter(β, θ), where $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$. So, a projection-based asymptotic confidence set for the endogeneity parameters a is

$$C_a^{\infty}(\alpha) = \{a \in \mathbb{R}^G : (\beta', \theta')' \in C_{(\beta,\theta)}^{\infty}(\alpha)\}. \tag{4.37}$$

As in the finite-sample setup, the Theorem **4.2** below gives the forms of $C_a^{\infty}(\alpha)$.

Theorem 4.2 ASYMPTOTIC PROJECTION-BASED CONFIDENCE SETS FOR ENDOGENEITY. *Assume that the assumptions* (2.1) - (2.3), (4.1) - (4.12) *hold. If furthermore* $\beta = \beta_0$ *and* $a = a_0$, *where* β_0 *and* a_0 *are* $G \times 1$ *constant vectors, then the sets* $C_a^{\infty}(\alpha)$ *take one of the following forms* :

(I) *if* $2A_h\beta + b_h = 0$, *then*

(I_1) *if* $\tilde{A}_h$ *is semidefinite positive and* $\tilde{A}_h \neq 0$,

$$C_a^{\infty}(\alpha) = \begin{cases} \{a \in \mathbb{R}^G : a = \bar{a}_h\} & \text{if} \quad \tilde{c}_h^* \leq 0 \quad \text{and} \quad c_h^* \leq 0 \\ \emptyset \quad \text{otherwise}, \end{cases} \tag{4.38}$$

(I_2) *if* $\tilde{A}_h = 0$,

$$C_a^{\infty}(\alpha) = \begin{cases} \left\{a \in \mathbb{R}^G : \tilde{b}_h' a + \tilde{c}_h + \tilde{b}_h' \bar{\beta}_h \leq 0\right\} & \text{if} \quad c_h^* \leq 0 \\ \emptyset \quad \text{otherwise}, \end{cases} \tag{4.39}$$

(I_3) *if* $\tilde{A}_h$ *is not semidefinite positive,*

$$C_a^{\infty}(\alpha) = \begin{cases} \left\{a \in \mathbb{R}^G : a'\tilde{A}_h a + \bar{b}_h' a + \bar{c}_h \leq 0\right\} & \text{if} \quad c_h^* \leq 0 \\ \emptyset \quad \text{otherwise} \end{cases} \tag{4.40}$$

where $\bar{b}_h = \tilde{b}_h + 2\tilde{A}_h\bar{\beta}_h$ *and* $\bar{c}_h = \tilde{c}_h + \bar{\beta}_h'\tilde{A}_h\bar{\beta}_h + \tilde{b}_h'\bar{\beta}_h$,

(II) *if* $2A_h\beta + b_h \neq 0$, *then*

(II_1) *if* $\tilde{A}_h$ *and* A_h *are semidefinite positive,*

$$C_a^{\infty}(\alpha) = \left\{a \in \mathbb{R}^G : a'\Delta_{\lambda_{1h}} a + \tilde{b}_{\lambda_{1h}}' a + \tilde{c}_{\lambda_{1h}} \leq 0,\ a'\Gamma_{\lambda_{1h}} a + b_{\lambda_{1h}}' a + c_{\lambda_{1h}} \leq 0\right\}, \tag{4.41}$$

(II_2) *if* $\tilde{A}_h = 0$,

$$C_a^{\infty}(\alpha) = \begin{cases} \emptyset \quad \textit{if} \quad [\tilde{b}_h \neq 0] \;\textit{or}\; [\tilde{b}_h = 0,\, \tilde{c}_h > 0] \;\textit{or}\; [A_h = 0, b_h = 0, c_h > 0] \\ \quad \textit{or}\; [A_h \textit{ is positive definite and } d_h = \frac{1}{4} b_h' A_h^{-1} b_h - c_h < 0] \\ \quad \textit{or}\; [\Delta_{\mu h} < 0,\, A_h \geq 0] \\ \\ \left\{ a \in \mathbb{R}^G : \tilde{b}_h' a - [\tilde{b}_h' A_h^- \tilde{b}_h] \mu_h^+ + \tilde{c}_h - \frac{\tilde{b}_h' A_h^- b_h}{2} + \tilde{b}_h' (I - \mathcal{H}_{A_h}) \beta_{10h} \leq 0 \right\} \\ \quad \textit{if} \quad [\Delta_{\mu h} \geq 0,\, A_h \geq 0] \\ \\ \mathbb{R}^G \quad \textit{if } A_h \textit{ is not positive semidefinite}\,, \end{cases} \tag{4.42}$$

where μ_h^+ *is the positive solution of*

$$[\tilde{b}_h' A_h^- \tilde{b}_h] \mu_h^2 - [\tilde{b}_h' (I - \mathcal{H}_{A_h}) \beta_{10h}] \mu_h + c_{0h} = 0\,, \tag{4.43}$$

(II_3) *if* $\tilde{A}_h$ *or* A_h *is not semidefinite positive,* $C_a^{\infty}(\alpha)$ *is bounded if it is empty, otherwise,* $C_a^{\infty}(\alpha)$ *is unbounded.*

where the indexation h *denotes the variables defined in* (3.46) - (3.62) *by replacing* $A,\ \tilde{A},\ \tilde{b},\ b,\ \tilde{c},\ c,$ *by* $A_h,\ \tilde{A}_h,\ \tilde{b}_h,\ b_h,\ \tilde{c}_h,\ c_h$.

The above theorem shows that the projection techniques are asymptotically valid even in presence of heteroskedasticity and/or autocorrelation. This means that the projection techniques are robust to heteroskedasticity and/or autocorrelation. Furthermore, these results hold whether the instruments are strong or weak. So, the procedure is asymptotically robust to weak instruments and heteroskedasticity and/or autocorrelation.

Corollary **4.3** below gives the necessary and sufficient conditions under which the sets $C_a^{\infty}(\alpha)$ are bounded.

Corollary 4.3 N.S. CONDITIONS FOR BOUNDED SETS: ASYMPTOTIC SETUP. *Assume that the assumptions of Theorem* **4.2**, *then* $C_a^{\infty}(\alpha)$ *defined above is bounded if and only if* A_h *and* $\tilde{A}_h$ *are positive definite. Otherwise,* $C_a^{\infty}(\alpha)$ *is bounded if it is empty.*

As in the finite-sample framework, if the concentration matrices A_h and $\tilde{A}_h$ are positive definite, then the projection-based confidence set $C_a^{\infty}(\alpha)$ is bounded. Otherwise, $C_a^{\infty}(\alpha)$ is bounded if it is empty. If the rank condition for identification (2.4) fails, β and a are not identified and the concentration matrices A_h and $\tilde{A}_h$ are not positive definite. So, the confidence sets $C_a^{\infty}(\alpha)$ are unbounded when they are non empty.

5. Conclusion

In this paper, we provide a finite-and large-sample confidence sets for endogeneity parameter between errors and endogenous regressors by using projection-based techniques [Dufour (1997), Abdelkhalek and Dufour (1998), Dufour and Jasiak (2001)]. We show that the procedure proposed is robust to weak instruments in both finite-and large-sample and provide the analytic forms of the confidence sets. We also provide necessary and sufficient conditions under which such confidence sets are bounded. Furthermore, we show that the procedure is asymptotically valid even in presence of heteroskedasticity and/or autocorrelation.

A. Proofs

PROOF OF THEOREM **3.1** Define the Lagrangian of the problem (3.45) as

$$\mathcal{L}(\beta, a, \lambda) = \mathcal{Q}(\beta, a) + \lambda f(\beta)\,. \tag{A.1}$$

The F.O.C are

$$\begin{aligned} \frac{\partial \mathcal{Q}(\beta, a)}{\partial \beta} + \lambda \frac{\partial f(\beta)}{\partial \beta} &= 0 & \text{(A.2)} \\ \lambda f(\beta) &= 0 & \text{(A.3)} \\ \lambda \geq 0, \quad f(\beta) &\leq 0 & \text{(A.4)} \end{aligned}$$

i.e.

$$\begin{aligned} 2\tilde{A}\beta + 2\tilde{A}a + \tilde{b} + \lambda(2A\beta + b) &= 0 & \text{(A.5)} \\ \lambda f(\beta) &= 0 & \text{(A.6)} \\ \lambda \geq 0, \quad f(\beta) = \beta' A\beta + b'\beta + c &\leq 0\,. & \text{(A.7)} \end{aligned}$$

The bordered hessian for binding constraint is given by

$$\bar{H}_{br} = \begin{bmatrix} 0 & 2\beta' A + b' \\ 2A\beta + b & 2\tilde{A} + 2\lambda A \end{bmatrix}. \tag{A.8}$$

We shall distinguish between 2 cases:

1. $2A\beta + b = 0$ and
2. $2A\beta + b \neq 0\,.$

For any matrix N, let $\mathcal{H}_N = N^- N$, where N^- is any generalized inverse of N.

Case 1: suppose that $2A\beta + b = 0$ *i.e* $\beta = \bar{\beta} = -\frac{1}{2}A^- b + (I - \mathcal{H}_A)\beta_{0*}$, where β_{0*} in any arbitrary $G \times 1$ vector and $AA^- b = b$ (consistency). Then, eqs.(A.5)- (A.8) becomes

$$\begin{aligned} 2\tilde{A}\beta + 2\tilde{A}a + \tilde{b} &= 0\,, & \text{(A.9)} \\ \lambda f(\beta) &= 0\,, & \text{(A.10)} \\ \lambda \geq 0, \quad f(\beta) = \beta' A\beta + b'\beta + c &\leq 0\,, & \text{(A.11)} \end{aligned}$$

$$\bar{H}_{br} = \begin{bmatrix} 0 & 0' \\ 0 & 2\tilde{A} + 2\lambda A \end{bmatrix}. \tag{A.12}$$

We will now consider between 2 subcases: (a) $\lambda = 0$ and (b) $\lambda > 0$.

(a) Suppose that $\lambda = 0$. Then, from (A.9)- (A.12), we have

$$2\tilde{A}\beta + 2\tilde{A}a + \tilde{b} \;=\; 0\,, \tag{A.13}$$
$$f(\beta) = \beta' A\beta + b'\beta + c \;\leq\; 0\,, \tag{A.14}$$

$$\bar{H}_{br} \;=\; \begin{bmatrix} 0 & 0' \\ 0 & 2\tilde{A} \end{bmatrix}. \tag{A.15}$$

$(I_1.a)$: If $\tilde{A}$ is positive semidefinite and $\tilde{A} \neq 0$, the bordered based approach in Magnus and Neudecker (1998, Theorem 12) is satisfied. Furthermore, from (A.13), we have

$$\begin{aligned} a &= \tilde{A}^{-}(-\tilde{A}\beta - \frac{1}{2}\tilde{b}) + (I - \mathcal{H}_{\tilde{A}})a_* = -\mathcal{H}_{\tilde{A}}\beta - \frac{1}{2}\tilde{A}^{-}\tilde{b} + (I - \mathcal{H}_{\tilde{A}})a_* \\ &= \frac{1}{2}\mathcal{H}_{\tilde{A}}A^{-}b - \mathcal{H}_{\tilde{A}}(I - \mathcal{H}_A)\beta_{0*} - \frac{1}{2}\tilde{A}^{-}\tilde{b} + (I - \mathcal{H}_{\tilde{A}})a_* = \bar{a}\,, \end{aligned} \tag{A.16}$$

where $\tilde{A}\tilde{A}^{-}(-\tilde{A}\beta - \frac{1}{2}\tilde{b}) = -\tilde{A}\beta - \frac{1}{2}\tilde{b}$ and a_* is an arbitrary $G \times 1$ vector. Substituting (A.16) in both the objective and constraint functions gives

$$Q(\beta, a) \;=\; \tilde{c} - \frac{1}{4}\tilde{b}'\tilde{A}^{-}\tilde{b}' - \frac{1}{2}\beta_{0*}'(I - \mathcal{H}_A')\tilde{A}(I - \mathcal{H}_A)\beta_{0*} \equiv \tilde{c}^*\,, \tag{A.17}$$
$$f(\beta) \;=\; c - \frac{1}{4}b'A^{-}b' + \frac{1}{2}b'(I - \mathcal{H}_A)\beta_{0*} \equiv c^*\,. \tag{A.18}$$

So, we have

$$C_a(\alpha) = \begin{cases} \{a \in \mathbb{R}^G : a = \bar{a},\ \beta_{0*} \in \mathbb{R}^G,\ a_* \in \mathbb{R}^G\} & \text{if} \\ \tilde{c}^* \leq 0 \quad \text{and} \quad c^* \leq 0 \\ \emptyset \quad \text{otherwise}\,. \end{cases}$$

$(I_2.a)$ If $\tilde{A} = 0$, we see immediately that

$$C_a(\alpha) = \begin{cases} \left\{a \in \mathbb{R}^G : \tilde{b}'a + \tilde{c} + \tilde{b}\bar{\beta} \leq 0\right\} & \text{if} \quad c^* \leq 0 \\ \emptyset \quad \text{otherwise}\,. \end{cases} \tag{A.19}$$

$(I_3.a)$ If $\tilde{A}$ is not positive semidefinite, this entails that $\tilde{A} \neq 0$. So, we can write

$$\begin{aligned} Q(a, \bar{\beta}) &= a'\tilde{A}a + (\tilde{b} + 2\tilde{A}\bar{\beta})'a + \tilde{c} + \bar{\beta}'\tilde{A}\bar{\beta} + \tilde{b}'\bar{\beta} \\ &= a'\tilde{A}a + \bar{b}'a + \bar{c}\,, \end{aligned} \tag{A.20}$$

so that

$$C_a(\alpha) = \begin{cases} \left\{a \in \mathbb{R}^G : a'\tilde{A}a + \bar{b}'a + \bar{c} \le 0\right\} & \text{if} \quad c^* \le 0 \\ \emptyset \quad \text{otherwise} \end{cases} \tag{A.21}$$

is bounded if and only if it is $\emptyset$ [we will check this in Corollary **3.2**], where $\bar{b} = \tilde{b} + 2\tilde{A}\bar{\beta}$ and $\bar{c} = \tilde{c} + \bar{\beta}'\tilde{A}\bar{\beta} + \tilde{b}'\bar{\beta}$.

(b) Suppose now that $\lambda > 0$. Then eqs.(A.9)- (A.12) become

$$2\tilde{A}\beta + 2\tilde{A}a + \tilde{b} \quad = \quad 0\,, \tag{A.22}$$

$$f(\beta) = \beta'A\beta + b'\beta + c \quad = \quad 0\,, \tag{A.23}$$

$$\bar{H}_{br} \quad = \quad \begin{bmatrix} 0 & 0' \\ 0 & 2\tilde{A} + 2\lambda A \end{bmatrix}. \tag{A.24}$$

We note that the solution β of (A.22)- (A.23) does not involve any λ.
$(I_1.b)$ If $\tilde{A} \neq 0$ is positive semidefinite, we can choose $\lambda > 0$ sufficiently small such that $\tilde{A} + \lambda A$ is is positive semidefinite. So, as $(I_1.a)$, we get

$$C_a(\alpha) = \begin{cases} \left\{a \in \mathbb{R}^G : a = \bar{a},\ \beta_{0*} \in \mathbb{R}^G,\ a_* \in \mathbb{R}^G\right\} \quad \text{if} \\ \tilde{c}^* \le 0 \quad \text{and} \quad c^* = 0 \\ \\ \emptyset \quad \text{otherwise}\,. \end{cases}$$

$(I_2.b)$ If $\tilde{A} = 0$, it is easy to see that

$$C_a(\alpha) = \begin{cases} \left\{a \in \mathbb{R}^G : \tilde{b}'a + \tilde{c} + \tilde{b}\bar{\beta} \le 0\right\} & \text{if} \quad c^* = 0 \\ \emptyset \quad \text{otherwise}\,. \end{cases} \tag{A.25}$$

$(I_3.b)$ If $\tilde{A}$ is not positive semidefinite, $C_a(\alpha)$ can be written as

$$C_a(\alpha) = \begin{cases} \left\{a \in \mathbb{R}^G : a'\tilde{A}a + \bar{b}'a + \bar{c} \le 0\right\} & \text{if} \quad c^* = 0 \\ \\ \emptyset \quad \text{otherwise} \end{cases} \tag{A.26}$$

and is bounded if and only if it is $\emptyset$. otherwise, $C_a(\alpha)$ is unbounded.

By putting the results of $(I.a)$ and $(I.b)$ together, we get the results of Theorem **3.1**-(I).

Case 2 : Assume that $2A\beta + b \neq 0$. We consider 3 subcases: (II.a) $\tilde{A} = 0$; (II.b) $\tilde{A} \neq 0$ and A are

positive definite; (II.c) $\tilde{A}$ or A is not positive definite.
(II.a) Suppose that $\tilde{A} = 0$, then $Q(\beta, a) = \tilde{b}'a + \tilde{b}'\beta + \bar{c}$ is linear in β. In this case we have

$$C_a(\alpha) = \left\{ a \in \mathbb{R}^G : \tilde{b}'a + \tilde{b}'\beta + \bar{c} \leq 0\,, \quad \beta \quad \text{satisfies} \quad \beta' A \beta + b'\beta + c \leq 0 \right\} \tag{A.27}$$

If $\lambda = 0$, we have

$$C_a(\alpha) = \begin{cases} \emptyset \quad \text{if} \quad [\tilde{b} \neq 0] \text{ or } [\tilde{b} = 0,\, \tilde{c} > 0] \text{ or } [A \text{ is positive definite and } d = \frac{1}{4} b' A^{-1} b - c < 0] \\ \quad \text{or } [A = 0,\, b = 0,\, c > 0] \\ \\ \mathbb{R}^G \quad \text{otherwise}\,. \end{cases} \tag{A.28}$$

If $\lambda > 0,$ this entails that $\tilde{b} \neq 0$, so that (A.5)- (A.8) give

$$A\beta = -\tfrac{b}{2} - \tfrac{\tilde{b}}{2\lambda} \tag{A.29}$$

$$\beta = -\tfrac{A^{-}b}{2} - \tfrac{A^{-}\tilde{b}}{2\lambda} + (I - \mathcal{H}_A)\beta_{10} \equiv \beta_\lambda \tag{A.30}$$

$$f(\beta_\lambda) = \beta_\lambda' A \beta_\lambda + b'\beta_\lambda + c = 0, \tag{A.31}$$

where β_{10} is any arbitrary $G \times 1$ vector. The bordered hessian at this point is given by

$$\bar{H}_{br} = \begin{bmatrix} 0 & -\frac{\tilde{b}'}{\lambda} \\ -\frac{\tilde{b}}{\lambda} & 2\lambda A \end{bmatrix} = \begin{bmatrix} 0 & -\mu\tilde{b}' \\ -\mu\tilde{b} & \frac{1}{\mu}A \end{bmatrix}, \tag{A.32}$$

where $\mu = \frac{1}{2\lambda}$. Hence $\beta_\lambda = -\frac{A^{-}b}{2} - \mu A^{-}\tilde{b} + (I - \mathcal{H}_A)\beta_{10}$ and by substituting this in (A.31), we get

$$[\tilde{b}' A^{-} \tilde{b}]\mu^2 - [\tilde{b}'(I - \mathcal{H}_A)\beta_{10}]\mu + c_0 = 0, \quad \mu > 0\,, \tag{A.33}$$

where $c_0 = c - \frac{b'A^{-}b}{4} + \frac{b'(I - \mathcal{H}_A)\beta_{10}}{2}$. By substituting β_λ in the objective function, we also get

$$Q(\beta_\lambda, a) = \tilde{b}'a - [\tilde{b}'A^{-}\tilde{b}]\mu + \tilde{c} - \frac{\tilde{b}'A^{-}b}{2} + \tilde{b}'(I - \mathcal{H}_A)\beta_{10}, \quad \mu > 0\,. \tag{A.34}$$

So, we see that

$$C_a(\alpha) = \begin{cases} \emptyset \quad \text{if} \quad \Delta_\mu = [\tilde{b}'(I - \mathcal{H}_A)\beta_{10}]^2 - 4(\tilde{b}'A^{-}\tilde{b})c_0 < 0 \quad \text{and } A \geq 0 \\ \\ \left\{ a \in \mathbb{R}^G : \tilde{b}'a - [\tilde{b}'A^{-}\tilde{b}]\mu + \tilde{c} - \frac{\tilde{b}'A^{-}b}{2} + \tilde{b}'(I - \mathcal{H}_A)\beta_{10} \leq 0 \right\} \\ \quad \text{if} \quad \Delta_\mu \geq 0 \quad \text{and} \quad A \geq 0 \\ \\ \mathbb{R}^G \quad \text{if } A \text{ is not positive semidefinite}\,, \end{cases} \tag{A.35}$$

where μ^+ is the positive solution of (A.33). Putting (A.28)-(A.35) gives Theorem **3.1**-(II$_2$).
(II.b) $\tilde{A}$ and A are positive semidefinite. This entails that $\tilde{A} + \lambda A \geq 0$ for any $\lambda \geq 0$. Define

$$\vartheta_\lambda = \{\lambda : |\tilde{A} + \lambda A| = 0\} \quad \text{and} \quad \lambda_1 = \min_{\lambda} \vartheta_\lambda \,. \tag{A.36}$$

Since $\tilde{A}$ and A are positive semidefinite, λ_1 in (A.36) is non negative and insures the semidefinite positivity of $\tilde{A} + \lambda A$, that is $\tilde{A} + \lambda_1 A \geq 0$. Of course, if the constraint is not binding, $\lambda_1 = 0$. So, from (A.5)- (A.8), we have

$$\begin{aligned} (\tilde{A} + \lambda_1 A)\beta &= -\tilde{A}a - \frac{\tilde{b}}{2} - \frac{\lambda_1 b}{2} && \text{(A.37)} \\ \beta' A\beta + b'\beta + c &\leq 0, && \text{(A.38)} \end{aligned}$$

or

$$\beta_{\lambda_1} = -A^-_{\lambda_1}\tilde{A}a - \frac{A^-_{\lambda_1}\tilde{b}}{2} - \frac{\lambda_1 A^-_{\lambda_1} b}{2} + (I - \mathcal{H}_{A^-_{\lambda_1}})\beta_{30} \,, \tag{A.39}$$

$$\beta'_{\lambda_1} A\beta_{\lambda_1} + b'\beta_{\lambda_1} + c \leq 0, \tag{A.40}$$

where β_{30} is an arbitrary $G \times 1$ vector and $A_{\lambda_1} = \tilde{A} + \lambda_1 A$. The bordered hessian at this point is

$$\bar{H}_{br} = \begin{bmatrix} 0 & 0 \\ 0 & 2(\tilde{A} + \lambda_1 A) \end{bmatrix} \geq 0 \,, \tag{A.41}$$

which insures that β_{λ_1} in (A.39) is a minimum. Substituting (A.39) in (A.40) and the objective function gives

$$a'\Gamma_{\lambda_1} a + b'_{\lambda_1} a + c_{\lambda_1} \leq 0 \tag{A.42}$$

$$Q(a;\lambda_1) = a'\Delta_{\lambda_1} a + \tilde{b}'_{\lambda_1} a + \tilde{c}_{\lambda_1} \,, \tag{A.43}$$

where

$$\begin{aligned} \Delta_{\lambda_1} &= \tilde{A} + \tilde{A}\tilde{\mathcal{A}}\tilde{A} - 2\tilde{A}A^-_{\lambda_1}\tilde{A} \,, && \text{(A.44)} \\ \tilde{b}_{\lambda_1} &= \tilde{b} + \frac{1}{2}\tilde{A}\tilde{\mathcal{A}}\tilde{A}\tilde{b} + \frac{\lambda_1}{2}\tilde{A}\tilde{\mathcal{A}}\tilde{A}b + \tilde{\mathcal{A}}\tilde{A}\tilde{b} + \frac{\lambda_1}{2}\tilde{\mathcal{A}}\tilde{A}b - 2\tilde{A}\tilde{\mathcal{A}}^*\beta_{30} - 2\tilde{A}A^-_{\lambda_1}\tilde{b} \\ &\quad -\lambda_1\tilde{A}A^-_{\lambda_1}b + 2\tilde{A}(I - \mathcal{H}_{A_{\lambda_1}})\beta_{30} \,, && \text{(A.45)} \\ \tilde{c}_{\lambda_1} &= \tilde{c} + \frac{\tilde{b}'\tilde{\mathcal{A}}\tilde{b}}{4} + \frac{3\lambda_1\tilde{b}'\tilde{\mathcal{A}}b}{4} + \frac{\lambda_1^2 b'\tilde{\mathcal{A}}b}{4} - \tilde{b}'\tilde{A}\beta_{30} - \frac{3\lambda_1 b'\tilde{\mathcal{A}}^*\beta_{30}}{2} - \frac{\tilde{b}'A^-_{\lambda_1}\tilde{b}}{2} \\ &\quad -\frac{\lambda_1\tilde{b}'A^-_{\lambda_1}b}{4} + \tilde{b}'(I - \mathcal{H}_{A_{\lambda_1}})\beta_{30} + \beta'_{30}(I - \mathcal{H}'_{A_{\lambda_1}})\tilde{A}(I - \mathcal{H}_{A_{\lambda_1}})\beta_{30} \,, && \text{(A.46)} \end{aligned}$$

$$\begin{aligned}
\Gamma_{\lambda_1} &= \tilde{A}\mathcal{A}\tilde{A}, \quad b_{\lambda_1} = \frac{1}{2}\tilde{A}\mathcal{A}(I+\tilde{A})(\tilde{b}+b) - \tilde{A}A^{-}_{\lambda_1}b - 2\tilde{A}\mathcal{A}^{*}\beta_{30}\,, && \text{(A.47)}\\
c_{\lambda_1} &= c + \frac{\tilde{b}'\mathcal{A}\tilde{b}}{4} + \frac{3\lambda_1\tilde{b}'\mathcal{A}b}{4} + \frac{\lambda_1^2 b'\mathcal{A}b}{4} - \lambda_1 b'\mathcal{A}\beta_{30} - \tilde{b}'\mathcal{A}^{*}\beta_{30} + b'(I-\mathcal{H}_{A_{\lambda_1}})\beta_{30}\\
&\quad + \beta_{30}'(I-\mathcal{H}'_{A_{\lambda_1}})A(I-\mathcal{H}_{A_{\lambda_1}})\beta_{30} - \frac{\lambda_1 b'A^{-}_{\lambda_1}b}{2} - \frac{b'A^{-}_{\lambda_1}\tilde{b}}{2}\,, && \text{(A.48)}\\
\tilde{\mathcal{A}} &= A_{\lambda_1}\tilde{A}A_{\lambda_1}, \quad \mathcal{A} = A_{\lambda_1}AA_{\lambda_1}, \quad \tilde{\mathcal{A}}^{*} = A_{\lambda_1}\tilde{A}(I-\mathcal{H}_{A_{\lambda_1}})\\
\mathcal{A}^{*} &= A_{\lambda_1}A(I-\mathcal{H}_{A_{\lambda_1}})\,, && \text{(A.49)}
\end{aligned}$$

So, we have

$$C_a(\alpha) = \Big\{a \in \mathbb{R}^G : a'\Delta_{\lambda_1}a + \tilde{b}'_{\lambda_1}a + \tilde{c}_{\lambda_1} \leq 0, \text{ and } a'\Gamma_{\lambda_1}a + b'_{\lambda_1}a + c_{\lambda_1} \leq 0\Big\}. \tag{A.50}$$

(II.c) Suppose now that $\tilde{A}$ or A is not positive definite. C_β and C_θ are unbounded or empty. Consequently, C_a is unbounded or empty. □

PROOF OF COROLLARY **3.2**

(A) Suppose first that $2A\beta + b = 0$. We see from Theorem **3.1**-(I$_1$) and (I$_2$) that the sets $C_a(\alpha)$ are non empty bounded if $\tilde{A}$ is p.d and $\tilde{c}^* \leq 0$ and $c^* \leq 0$. We can verify that this corresponds to A p.d. Otherwise, $C_a(\alpha)$ is bounded if it is empty. Now, if $\tilde{A}$ is not p.s.d, from Theorem **3.1**-(I$_3$), we can write

$$\begin{aligned}
Q(a,\bar{\beta}) &= a'\tilde{A}a + (\tilde{b} + 2\tilde{A}\bar{\beta})'a + \tilde{c} + \bar{\beta}'\tilde{A}\bar{\beta} + \tilde{b}'\bar{\beta}\\
&= a'\tilde{A}a + \bar{b}'a + \bar{c} = (a-\tilde{a})'\tilde{A}(a-\tilde{a}) - \tilde{d}\,, && \text{(A.51)}
\end{aligned}$$

where $\tilde{a} = -\frac{1}{2}\tilde{A}^{-1}\bar{b}$, $\tilde{d} = \frac{1}{4}\bar{b}'\tilde{A}^{-1}\bar{b} - \bar{c}$. Moreover, since $\tilde{A} = Y'\tilde{H}Y$ is a real symmetric matrix, hence

$$\tilde{A} = \tilde{P}'\tilde{D}\tilde{P} \tag{A.52}$$

where $\tilde{P}$ is an orthogonal matrix and $\tilde{D}$ is a diagonal matrix whose elements are the eigenvalues of $\tilde{A}$. So we can reexpress $Q(a,\bar{\beta}) \leq 0$ as

$$\left(\frac{\tilde{z}_1}{\tilde{\gamma}_1}\right)^2 + \left(\frac{\tilde{z}_2}{\tilde{\gamma}_2}\right)^2 + \ldots + \left(\frac{\tilde{z}_g}{\tilde{\gamma}_g}\right)^2 - \left(\frac{\tilde{z}_{g+1}}{\tilde{\gamma}_{g+1}}\right)^2 - \ldots - \left(\frac{\tilde{z}_G}{\tilde{\gamma}_G}\right)^2 \leq \tilde{d}\,, \tag{A.53}$$

where $\tilde{\lambda}_i$ are the eigenvalues of $\tilde{A}$, $\tilde{z} = \tilde{P}(a-\tilde{a})$, $\tilde{\gamma}_i = \sqrt{1/\tilde{\lambda}_i}$ for $j = 1,\ldots,g$, $\tilde{\gamma}_i = \sqrt{-1/\tilde{\lambda}_i}$ for $j = g+1,\ldots,G$ and g is the number of positive eigenvalues of $\tilde{A}$. Then, for arbitrary given values $\tilde{z}_1,\ldots,\tilde{z}_g$ and $\tilde{d}$, inequalities (A.53) will hold if any of the values $\tilde{z}_i$, $i = 1,\ldots,g$, is small enough (as $|\tilde{z}_j| \to \infty$). Consequently, C_a is unbounded if $c^* \leq 0$. If $c^* > 0$, C_a is empty.

(B) suppose now that $2A\beta + b \neq 0$. We can see form Theorem **3.1**-(II$_2$) and (II$_3$) that if $\tilde{A}$ or A are not p.s.d, C_a is bounded only if it is empty. Moreover, if $\tilde{A}$ and A are both p.d, Theorem **3.1**-(II$_1$) stays that C_a is an intersection of two ellipsoids, hence, is bounded or empty.

$\square$

PROOF OF LEMMA **4.1** Note first that the AR-type statistics defined in (4.18) - (4.19) can be written as

$$AR^{he}(\beta_0) = \frac{\upsilon_1(\beta_0)'[P_{X^*} - P_{X_1^*}]\upsilon_1(\beta_0)/k_2}{\upsilon_1(\beta_0)'M_{X^*}\upsilon_1(\beta_0)/(T-k_1-k_2)}\,, \tag{A.54}$$

and

$$AR^{he}(\theta_0) = \frac{\varepsilon_1(\theta_0)'[P_{Z^*} - P_{X^{**}}]\varepsilon_1(\theta_0)/G}{\varepsilon_1(\theta_0)'M_{Z^*}\varepsilon_1(\theta_0)/(T-k-G)}\,, \tag{A.55}$$

where $\upsilon_1(\beta_0) = \hat{\Sigma}_u^{-1/2}\upsilon^0(\beta_0) \equiv \upsilon_1$, $\varepsilon_1(\theta_0) = \hat{\Sigma}_\varepsilon^{-1/2}\varepsilon(\theta_0) \equiv \varepsilon_1$, $X^* = [X_1^*,\, X_2^*] = \hat{\Sigma}_u^{-1/2}X$, $X^{**} = [X_1^{**},\, X_2^{**}] = \hat{\Sigma}_\varepsilon^{-1/2}X,\ Z^* = [Y^{**},\, X^{**}] = \hat{\Sigma}_\varepsilon^{-1/2}Y,\ P_{\mathcal{Z}} = \mathcal{Z}(\mathcal{Z}'\mathcal{Z})^{-1}\mathcal{Z}'$ and $M_{\mathcal{Z}} = I - P_{\mathcal{Z}},$ for any matrix $\mathcal{Z}$. Now, consider the denominators of (A.54), we have

$$\frac{\upsilon_1' M_{X^*}\upsilon_1}{T-k} = \frac{\upsilon_1'\upsilon_1}{T-k} - \frac{T}{T-k}\frac{\upsilon_1'X^*}{T}\left(\frac{X^{*\prime}X^*}{T}\right)^{-1}\frac{X^{*\prime}\upsilon_1}{T}\,. \tag{A.56}$$

Under H_{β_0}, we have $\upsilon^0 = u$, hence $\frac{\upsilon_1'\upsilon_1}{T-k} = \frac{u'\hat{\Sigma}_u^{-1}u}{T-k} = \frac{u'(\hat{\Sigma}_u^{-1} - \Sigma_u^{-1})u}{T-k} + \frac{u'\Sigma_u^{-1}u}{T-k} \xrightarrow{p} 1,\ \frac{X^{*\prime}\upsilon_1}{T} = \frac{X'\hat{\Sigma}_u^{-1}u}{T} = \frac{X'(\hat{\Sigma}_u^{-1} - \Sigma_u^{-1})u}{T} + \frac{X'\Sigma_u^{-1}u}{T} \xrightarrow{p} 0,\ \frac{X^{*\prime}X^*}{T} = \frac{X'\hat{\Sigma}_u^{-1}X}{T} = \frac{X'(\hat{\Sigma}_u^{-1} - \Sigma_u^{-1})X}{T} + \frac{X'\Sigma_u^{-1}X}{T} \xrightarrow{p} \Omega_X > 0$ because $\underset{T\to\infty}{\text{plim}}\,(\hat{\Sigma}_u) = \Sigma_u$. So, we have $\frac{\upsilon_1'M_{X^*}\upsilon_1}{T-k} \xrightarrow{p} 1.$ The numerator of (A.54) can be written as

$$\begin{aligned}
&\upsilon_1'[P_{X^*} - P_{X_1^*}]\upsilon_1/k_2 \\
=\ & \left[\begin{array}{cc} \frac{u'\Sigma_u^{-1}X}{\sqrt{T}} & \frac{u'\Sigma_u^{-1}X_1}{\sqrt{T}} \end{array}\right]\left[\begin{array}{cc} \left(\frac{X'\hat{\Sigma}_u^{-1}X}{T}\right)^{-1} & 0 \\ 0 & -\left(\frac{X_1'\hat{\Sigma}_u^{-1}X_1}{T}\right)^{-1} \end{array}\right]\left[\begin{array}{c} \frac{X'\Sigma_u^{-1}u}{\sqrt{T}} \\ \frac{X_1'\Sigma_u^{-1}u}{\sqrt{T}} \end{array}\right] && \text{(A.57)}\\
\xrightarrow{L}\ & S_u^{x\prime}\Omega_X^{-1}S_u^x - S_u^{x_1'}\Omega_{X_1}^{-1}S_u^{x_1} && \text{(A.58)}
\end{aligned}$$

because $\left[\begin{array}{cc} \left(\frac{X'\hat{\Sigma}_u^{-1}X}{T}\right)^{-1} & 0 \\ 0 & \left(\frac{X_1'\hat{\Sigma}_u^{-1}X_1}{T}\right)^{-1} \end{array}\right] \xrightarrow{p} \left[\begin{array}{cc} \Omega_X^{-1} & 0 \\ 0 & \Omega_{X_1}^{-1} \end{array}\right]$, where $S_u^x = \left[\begin{array}{c} S_u^{x_1} \\ S_u^{x_2} \end{array}\right]$ is partitioned according to $[X_1,\ X_2\,]$. We easily see that (A.58) can be written as

$$S_u^{x\prime}\Omega_X^{-1}S_u^x - S_u^{x_1'}\Omega_{X_1}^{-1}S_u^{x_1} = S_u^{x\prime}\Omega_X^{-1/2}\Omega\Omega_X^{-1/2}S_u^x\,, \tag{A.59}$$

where $\Omega = I_k - \Omega_X^{1/2}\left[\begin{array}{cc} \Omega_{X_1}^{-1} & 0 \\ 0 & 0 \end{array}\right]\Omega_X^{1/2}$. Moreover, $\Omega_X^{-1/2}S_u^x \sim N[0\,, I_k]$ and by noting that $\Omega^2 = \Omega$, hence Ω is idempotent, we have

$$S_u^{x\prime}\Omega_X^{-1/2}\Omega\Omega_X^{-1/2}S_u^x \sim \chi^2(m)\,,$$

where $m = rank(\Omega)$. Furthermore, we have

$$Trace(\Omega) = k - Trace\left(\Omega_X^{1/2}\begin{bmatrix}\Omega_{X_1}^{-1} & 0\\ 0 & 0\end{bmatrix}\Omega_X^{1/2}\right)$$

and

$$\begin{aligned}Trace\left(\Omega_X^{1/2}\begin{bmatrix}\Omega_{X_1}^{-1} & 0\\ 0 & 0\end{bmatrix}\Omega_X^{1/2}\right) &= Trace\left(\Omega_X\begin{bmatrix}\Omega_{X_1}^{-1} & 0\\ 0 & 0\end{bmatrix}\right)\\ &= Trace\left(\begin{bmatrix}I_{k_1} & 0\\ \Omega_{X_2X_1} & 0\end{bmatrix}\right) = k_1\,,\end{aligned}$$

where $\Omega_X = \begin{bmatrix}\Omega_{X_1} & \Omega'_{X_2X_1}\\ \Omega_{X_2X_1} & \Omega_{X_2}\end{bmatrix}$. So, we have $Trace(\Omega) = k_2$. Since Ω is idempotent, its eigenvalues λ_j are 1 or 0 [see Magnus and Neudecker (1998, Theorem 7, p.14)]. Because, $Trace(\Omega) = \sum_{j=1}^{k} = k_2$ then Ω has k_2 eigenvalues equal 1. So, $m = rank(\Omega) = k_2$ and $S_u^{x\prime}\Omega_X^{-1/2}\Omega\Omega_X^{-1/2}S_u^x \sim \chi^2(k_2)$, *i.e.*, $AR^{he}(\beta_0) \xrightarrow{L} \frac{1}{k_2}\chi^2(k_2)$. By following the same steps as above, we can show that $AR^{he}(\theta_0) \xrightarrow{L} \frac{1}{G}\chi^2(G)$.

□

PROOF OF THEOREM **4.2** Same proof as in Theorem **3.1**. □

PROOF OF COROLLARY **4.3** Same proof as in Corollary **3.2**. □

References

Abdelkhalek, T., Dufour, J.-M., 1998. Statistical inference for computable general equilibrium models with applications to a model of the moroccan economy., pp. 520–534.

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

Andrews, D. W. K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.

Andrews, D. W. K., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica 60(4), 953–966.

Doko, F., Dufour, J.-M., 2009. Inference on structural covariances with non gaussian distributions: A generalization of linear wald methods. Technical report, C.R.D.E., Université de Montréal.

Dufour, J.-M., 1987. Linear Wald methods for inference on covariances and weak exogeneity tests in structural equations. In: I. B. MacNeill, G. J. Umphrey, eds, Advances in the Statistical Sciences: Festschrift in Honour of Professor V.M. Joshi's 70th Birthday. Volume III, Time Series and Econometric Modelling. D. Reidel, Dordrecht, The Netherlands, pp. 317–338.

Dufour, J.-M., 1990. Exact tests and confidence sets in linear regressions with autocorrelated errors. Econometrica 58, 475–494.

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

Dufour, J.-M., Jan, K., 1998. Exact inference methods for first-order autoregressive distributed lag models. Econometrica 66, 9–104.

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

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

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

Engle, R. F., 1982. A general approach to lagrange multiplier diagnostics. Journal of Econometrics 20, 83–104.

Gallant, A. R., ed., 1987. Nonlinear Statistical Model. New York: Wiley, New York, New York.

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

Hwang, H.-S., 1980. Test of independence between a subset of stochastic regressors and disturbances. International Economic Review 21, 749–760.

Levine, D., 1983. A remark on serial correlation in maximum likelihood. JOE 23, 337–342.

Magnus, J. R., Neudecker, H., 1998. Matrix Differential Calculus with Applications in Statistics and Econometrics, Revised Edition. John Wiley & Sons, New York.

Meeker, W. Q., al., 1998. Accelaration degradation tests: Modeling and analysis. Technometrics 40(2), 89–99.

Newey, W., West, K., 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Revankar, N. S., Hartley, M. J., 1973. An independence test and conditional unbiased predictions in the context of simultaneous equation systems. International Economic Review 14, 625–631.

Smith, R. J., 1984. A note on likelihood ratio tests for the independence between a subset of stochastic regressors and disturbances. International Economic Review 25, 263–269.

White, H., Domowitz, I., 1984. Nonlinear regression with dependent observations. Econometrica 52, 143–161.

White, H., ed., 1984. Asymptotic Theory for Econometricians. New York: Academic Press, New York, New York.

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