Paolo Paruolo

Design of vector autoregressive processes for invariant statistics

2005/6

UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

http://eco.uninsubria.it

I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

The views expressed in the working papers reflect the opinions of the authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

Design of vector autoregressive processes for invariant statistics

Paolo Paruolo[*]

September 23, 2005

Abstract

This paper discusses the Monte Carlo (MC) design of Gaussian Vector Autoregressive processes (VAR) for the evalutation of invariant statistics. We focus on the case of cointegrated (CI) I(1) processes, linear and invertible transformations and CI rank likelihood ratio (LR) tests. It is found that all VAR of order 1 can be reduced to a system of independent or recursive subsystems, of computational dimension at most equal to 2. The results are applied to the indexing of the distribution of LR test statistics for CI rank under local alternatives. They are also extended to the case of VAR processes of higher order.

Keywords: Invariance, Vector autoregressive process, Monte Carlo, Likelihood ratio test, Cointegration.
JEL code: C32.

Contents

1   Introduction

This paper discusses the design of Gaussian VAR for MC simulations. We consider CI systems, integrated of order 1, I(1), and LR test statistics for CI rank. These tests are based on eigenvalues (squared canonical correlations) associated with reduced rank regression (RRR), which are invariant with respect to linear and invertible transformations of the variables. We show how this invariance can be used to reduce the dimension of the MC design.

MC experiments for the evaluation of the finite sample properties of LR CI tests has been considered by several authors, see e.g. Gonzalo (1994), Toda (1994, 1995), Saikkonen and Lütkepohl (2000), Johansen (2002), Nielsen (2004), Cavaliere, Fanelli and Paruolo (2005). The idea of using invariance to reduce the design dimension for a VAR or order 1 can be found in Johansen (2002) and Nielsen (2004), respectively for the case of CI rank equal to 1 and for the bivariate case.

In this paper we extend these results for any dimension of the CI rank, any number of variables in the system and any number of lags. We exploit the fact that I(1) VAR processes are closed under linear and invertible transformations, which form a group. This generates an associated group of transformations on the parameters. A classical result on invariance then assures that the distribution of statistics that are invariant to transformations on the variables only depend on the maximal invariant function for the transformation of the parameters.

One of the major limitations of MC simulations is their lack of generality. Each data generating process (DGP) gives information that is confined to the particular values chosen for the parameters. However if the statistics under investigation are invariant with respect to some group of transformations on the process, then the MC result of a single DGP cover all the processes traced by the orbit of a group of associated transformations acting on the parameters.

This observation thus can be used both to avoid simulation of two DGP on the same orbit (for which the distribution of the statistic of interest is the same) and to cover the parameter space uniformly, with a smaller set of DGP. We show how the MC design can be reduced both for VAR or order 1, VAR(1) – a popular choice – and VAR of higher order.

Despite the persistent increase in computing power, one issue that remains of concern in design of MC is the computational speed of various operations involved in data simulation. We also consider this issue and discuss the possibility to use of univariate, bivariate or multivariate recursions to generate the VAR. We show that a VAR(1) can be reduced to a system of independent or block-triangular subsystems, each at most of bivariate dimension.

We emphasize that the MC design reduction discussed in this paper for VAR processes equally apply to the VAR part of VARMA processes. We note that the simulation of the MA part does not involve recursions, and it is hence less time-consuming than the VAR part. We also remark that the invariance results discussed here with reference to MC design are analytic, and could be used also in analytic work on the distributions.

We give an application of the VAR(1) results to the indexing of the local alternative for LR CI rank tests. We also discuss how VAR(1) results can be extended to processes of order q, VAR(q), using the companion form. We find that a sizable MC reduction can be obtained by invariance with respect to the companion form parametrization, while this is not possible in general with respect to the standard I(1) Equilibrium Correction (EC) parametrization.

The rest of the paper is organized as follows. Section 2 formulates the problem.

Section 3 presents properties of VAR and RRR that are used in the Section 4 to define the design of VAR processes of order 1. An application of these results to the indexing of the limit distribution of LR tests for CI rank under a local alternative is reported in Section 5. Section 6 extends results to VAR(q) processes. Section 7 concludes. All proofs are placed in the Appendix.

We use the notation $\mathrm{diag}\,(A_1, ..., A_n)$ to indicate a matrix with blocks A_1, ..., A_n on the main diagonal and let $\mathrm{dg}(A)$ be the matrix with off–diagonal elements equal to 0 and diagonal elements equal to the ones on the diagonal of A.

2 Invariance and design reduction

In this Section we present a generic formulation of the problem, using the well known ideas of invariance and groups, see e.g. Lehmann (1986) Chapter 6.

Let X_t^* be a $n^* \times 1$ vector of stochastic processes with time index $t \in \mathcal{T}$, and define $X^* := \{X_t^*, t \in \mathcal{T}\}$. In particular we consider $\mathcal{T} := \{1, ..., T\}$ in discrete time situations and $\mathcal{T} := [0, 1]$ in the continuous time case. The stochastic process X^* lives on a probability space $(\mathcal{X}, \mathfrak{S}, \mathbb{P})$. $\mathbb{P}$ is the probability measure for X^*, i.e. $\mathbb{P}(X^* \in A)$, $A \in \mathfrak{S}$, gives the probability that X^* belong to A.

$\mathbb{P}$ is assumed to belong to some parametric family $\mathbb{P} = \mathbb{P}_\theta$ that is easy to simulate by MC, where $\theta \in \Theta$ are a finite dimensional parameter vector and parameter space, respectively. We assume that θ gives an identified parametrization, in the sense that $\mathbb{P}_{\theta^\circ} \neq \mathbb{P}_\theta$ whenever $\theta \neq \theta^\circ$. Let $\mathcal{P} := \{\mathbb{P}_\theta, \theta \in \Theta\}$ denote the associated class of probability measures.

2.1 Invariant statistic

We assume that interest lies with the distribution of a statistic $h(X^*)$ that is invariant with respect to some transformation g of X^*, $h(X^*) = h(g(X^*))$. The distribution of h is assumed to be analytically intractable, and that therefore it must be estimated by MC simulation. This is accomplished by sampling s i.i.d. draws X^{*i} from $\mathbb{P}_\theta$ for a fixed value of θ, and by estimating the distribution function of h by the MC empirical distribution function $\widehat{\mathrm{Pr}}_\theta(h(X^*) \leq x) = s^{-1} \sum_{i=1}^s 1\,(h(X^{*i}) \leq x)$, for any x and large s. Here $1(\cdot)$ is the indicator function. Θ is taken to be a subset of $\Re^{d_\Theta}$; it is also assumed that the MC design θ takes on values obtained by discretization of each coordinate in θ.

We take the MC design dimension as d_{Θ_1}.[1] The goal of the paper is to find ways to reduce the dimension of the parameter space Θ using the invariance of h; this is later called MC design (dimension) reduction. In the following we take g to represent linear invertible transformation of the process X^* defined as $X^\circ := g\,(X^*) := \{g\,(X_t^*), t \in \mathcal{T}\}$; g has matrix representation $X^\circ := HX^*$, where H is a square invertible matrix.

2.2 Transformed process

The transformed process X° has probability distribution $\mathbb{P}^\circ$ that is assumed still to belong to the class $\mathcal{P}$; in other words there exists a value $\theta^\circ \in \Theta$ such that $\mathbb{P}^\circ = \mathbb{P}_{\theta^\circ}$, i.e. the class $\mathcal{P}$ is closed under the transformation induced by g. This defines a function $\bar{g}$ that maps θ into θ°, $\theta^\circ = \bar{g}(\theta)$.

---

[1] This is an approximate indicator of the complexity of the MC simulation, because Θ is typically *not* a product space.

We assume that the parameter set Θ is preserved by g, in the sense that $\bar{g}(\theta) \in \Theta$ for all g, and that for any θ° there exists a θ such that $\theta^\circ = \bar{g}(\theta)$. Let $\mathcal{G}$ be a class of transformations g defined above and let G be the smallest class containing $\mathcal{G}$ such that $g_1, g_2 \in G$ implies $g_1 \circ g_2 \in G$ and $g^{-1} \in G$; then G is a group and the induced set of transformations $\bar{g}$ also forms a group $\bar{G}$, see Lehmann (1986) Lemma 1 page 283.

The set of points $g(x)$ for fixed $X^* = x$ and all $g \in G$ defines an orbit in $\mathcal{X}$. The orbits define a partition of the sample space $\mathcal{X}$ into sets of the form $\mathcal{X}_x := \{X^* \in \mathcal{X} : X^* = g(x), g \in G\}$. Similarly the set of points $\bar{g}(\theta)$ for fixed θ and all $\bar{g} \in \bar{G}$ defines an orbit in Θ. The G orbits define a partition of the sample space $\mathcal{X}$ into sets of the form $\mathcal{X}_x := \{X^* \in \mathcal{X} : X^* = g(x), g \in G\}$, and similarly for $\bar{G}$ orbits in Θ.

2.3 MC design reduction

We next recall the definition of maximal invariant, using X^* and the group G, noting that these definition equally apply to θ and $\bar{G}$. Recall that the function $h(X^*)$ is called invariant under G, if $h(X^*) = h(g(X^*))$ for all $X^* \in \mathcal{X}$ and $g \in G$. A statistic $h(X^*)$ is called maximal invariant if $h(X^{*1}) = h(X^{*2})$ implies $X^{*1} = g(X^{*2})$ for some $g \in G$. The maximal invariant statistic $h(X^*)$ is constant on the orbits, i.e. $h(X^*) = h(x)$ for all $X^* \in \mathcal{X}_x$, and takes different values on different orbits.

The following classical results, see Lehmann (1986) Theorem 3 in Chapter 6, gives the key to MC design reduction; it is reported here without proof for ease of later reference.

Theorem 1 *If $h(X)$ is invariant under G, and $\psi(\theta)$ is maximal invariant under the induced group $\bar{G}$, then the distribution of $h(X)$ depends only on $\psi(\theta)$.*

This result allows to partition Θ through the maximal invariant function $\psi(\theta)$ into sets of the form $\Theta_{\psi^\circ} := \{\theta \in \Theta : \psi(\theta) = \psi^\circ\}$ and to simulate just one process $\mathbb{P}_\theta$ for each Θ_{ψ°. In this way all points θ that lie on the same orbit of ψ are represented by a single MC experiment.

Note that this implies a reduction in the MC design dimension; in fact the parameter space is reduced from Θ to $F := \{\psi = \psi(\theta), \theta \in \Theta\} \subset \Re^{d_F}$, which has dimension not greater than Θ. By 'MC design reduction' we mean $d_{\Theta_1} - d_F$; this will be computed in the following sections for the case of I(1) VAR processes.

3 VARs and Invariance

In this section we collect properties of VAR processes and on the invariance properties of RRR. Subsection 3.1 defines the MC designs considered in the paper. Subsection 3.2 illustrates the properties of VAR processes under linear transformations. Subsection 3.3 presents invariance properties of the eigenvalues in RRR.

3.1 A class of probability measures

In this subsection we define the main class $\mathcal{P}$ of probability measures implied by Gaussian I(1) VAR processes, which enjoys a property of being closed closed under linear invertible transformations, similarly to unrestricted VARs.

We consider a $p \times 1$ process X_t generated by a VAR(q),

$$A(L)X_t = \mu D_t + \varepsilon_t \tag{1}$$

with deterministic part μD_t, i.i.d. innovations $\varepsilon_t \sim N(0, \Omega)$, and autoregressive polynomial $A(L) := -\sum_{i=0}^{q} A_i L^i$, $A_0 := -I$, L being the lag operator.

It is well known, see Johansen (1988), that the coefficients A_i, $i = 0, 1..., q$ can be linearly mapped into the set of coefficients $A(1) = -\sum_{j=0}^{q} A_j$, $\Psi_i := \sum_{j=0}^{i} A_j$, $i = 1, ..., q-1$, that characterize the equilibrium correction form, EC, see eq. (3) below. This map is a 1 to 1, so one can take either the set A_i, $i = 1, ..., q$ or the set $A(1)$, Ψ_i, $i = 1, ..., q-1$ to represent the AR polynomial $A(L)$.

We here take $X_t^* = X_t$, $X^* = \{X_1, ..., X_T\}$ in the notation of the previous section; $\mathbb{P}_\theta$ is a Gaussian measure on X^* induced by (1). The parameter θ is defined below in terms of $A(1)$ and $\Psi := (\Psi_1 : ... : \Psi_{q-1})$.

We assume that the VAR process X_t satisfies Granger's I(1) representation theorem, GRT, as given by Theorem 4.2 in Johansen (1996). Specifically the assumptions of GRT are the following:

(a) $|A(z)|$ has roots either at $z = 1$ or $|z| > 1$;

(b) $A(1)$ has rank r, $0 \leq r < p$, so that it allows representation $A(1) = -\alpha\beta'$ for α, β full column rank $p \times r$ matrices;

(c) $\alpha_\perp' \dot{A}(1)\beta_\perp = \alpha_\perp' (\Psi(i_{q-1} \otimes I_p) - I_p) \beta_\perp$ has full rank $p - r$, where $\dot{A}(z) = dA(z)/dz$, i_n is a $n \times 1$ vector of ones, and $\otimes$ is Kronecker's product.

We consider all VAR processes X_t in (1) that satisfy GRT. Moreover we assume that μD_t can be decomposed as $\mu_1 D_{1t} + \mu_2 D_{2t}$, where μ_i is $p \times m_i$ and D_{it} is $m_i \times 1$, $i = 1, 2$, and $\mu_1 = \alpha\rho_1'$. We define as parameters $\theta := (\alpha, \beta, \rho_1, \Psi, \mu_2, \Omega)$ satisfying (a) (b) (c) above, and Ω positive definite. Note that α, β in (b) are not identified because the decomposition $A(1) = -\alpha\beta'$ is not unique; we hence consider pairs $(\alpha^1, \beta^{1\prime})$, and $(\alpha^2, \beta^{2\prime})$ equivalent when $\alpha^1\beta^{1\prime} = \alpha^2\beta^{2\prime}$, this defines an equivalence relation and equivalence classes. The parametrization θ is identified up to these equivalence classes. This defines an identified parametrization with parameter space Θ_1.

We call the associated set of probability measures $\mathcal{P}_1 := \{\mathbb{P}_\theta, \theta \in \Theta_1\}$. We find that the dimension of Θ_1 is

$$d_{\Theta_1} = (2p - r + m_1)r + p(n + m_2) + \frac{1}{2}p(p+1), \qquad (2)$$

where $n := p(q-1)$.

In the next subsection we show that $\mathcal{P}_1$ is closed under the action of linear and invertible transformations g on X_t.

3.2 Transformations

In this subsection we define the linear transformations g. We consider the transformation g of the form $X_t^\circ := HX_t$, with H of dimension $p \times p$ and invertible. The following result shows that $\mathcal{P}_1$ is closed under the action of g, and defines $\bar{g}$ in this case.

Theorem 2 *Let X_t be a process with probability measure $\mathbb{P}_\theta \in \mathcal{P}_1$, $\theta := (\alpha, \beta, \rho_1, \Psi, \mu_2, \Omega) \in \Theta_1$; the transformed process $X_t^\circ := HX_t$ with H is square and invertible has probability measure $\mathbb{P}_{\theta^\circ} \in \mathcal{P}_1$, where $\theta^\circ = \bar{g}(\theta) := (H\alpha, H'^{-1}\beta, \rho_1, H\Psi(I_{q-1} \otimes H^{-1}), H^{-1}\mu_2, H\Omega H') \in \Theta_1$.*

We note that Theorem 2 implies that g preserves Θ_1: $\bar{g}(\theta) \in \Theta_1$ for all $\theta \in \Theta_1$ and given $\theta^\circ \in \Theta_1$ one can find a $\theta \in \Theta_1$ such that $\theta^\circ = \bar{g}(\theta)$. To verify the latter claim, simply note that $\bar{g}$ can be inverted to give $\theta = \bar{g}^{-1}(\theta^\circ) = (H^{-1}\alpha^\circ, H'\beta^\circ, \rho_1^\circ, H^{-1}\Psi^\circ(I_{q-1} \otimes H), H\mu_2^\circ, H^{-1}\Omega^\circ H'^{-1})$, which still belongs to Θ_1, such that $\theta^\circ = \bar{g}(\theta)$.

Consider now the class of transformation G of the form g given above. It is well known that the class G of invertible linear transformation forms a group, see e.g. Lehmann (1986) Appendix 1, Example 2. Hence also $\bar{G}$, the set of implied transformations $\bar{g}$ defined in Theorem 2 on θ is a group by Lemma 1 in Lehmann (1986), Chapter 6.

We next consider the statistical procedure of reduced rank regression and its invariance properties.

3.3 Invariance

In this subsection we describe the invariance properties of statistics based on RRR. For simplicity we assume that the statistical model (3) select the correct number of lags q; the model may be written in EC form:

$$\Delta X_t = \alpha \beta' X_{t-q} + \Psi U_{t-1} + \mu D_t + \varepsilon_t \tag{3}$$

where X_t and ε_t are $p \times 1$, $U_{t-1} := (\Delta X_{t-1}, ..., \Delta X_{t-q+1})'$ is $p(q-1) \times 1$, D_t is a vector of deterministic terms. Here we have chosen the EC form with level variables dated $t-q$, as in Johansen (1988). It is well known that the level term can be dated in $t-j$ where j can be chosen equal to 1, 2, ..., q. [2]

In model (3) α and β are $p \times j$ matrices (not necessarily of full column rank). Partition also D_t as $D_t := (D_{1t}' : D_{2t}')'$ and $\mu := (\mu_1 : \mu_2)$ where $\mu_1 = \alpha \rho_1' \in \text{col}(\alpha)$. We indicate the corresponding statistical model as $H(j)$, which can be put in the RRR format

$$Z_{0t} = \alpha \beta^{*\prime} Z_{1t} + \Upsilon Z_{2t} + \varepsilon_t \tag{4}$$

with $Z_{0t} := \Delta X_t$, $Z_{1t} := (X_{t-q}' : D_{1t}')'$, $Z_{2t} := (U_{t-1}' : D_{2t}')'$, $\beta^* := (\beta' : \rho_1')'$. If D_{it} is set equal to 0, it is understood that D_{it} is dropped from the definition of Z_{it}, $i = 1, 2$.

Given a sample $X := \{X_t, t = 1, ..., T\}$, let $\ell(\theta, j)$ indicate the Gaussian log-likelihood of model $H(j)$ and let $\ell(j) := \max_\theta \ell(\theta, j)$. The LR test of $H(j)$ within $H(l)$, for $j < l$ can be written as (see Johansen, 1996)[3]

$$LR(j, l) := -2(\ell(j) - \ell(l)) = -T \sum_{i=j+1}^{l} \log(1 - \widehat{\lambda}_i), \tag{5}$$

where $\widehat{\lambda}_i$ is the i-th largest solution of the eigenvalue problem

$$\left| \widehat{\lambda} S_{11} - S_{10} S_{00}^{-1} S_{01} \right| = 0. \tag{6}$$

and $S_{ij} := M_{ij.2} := M_{ij} - M_{i2} M_{22}^{-1} M_{2j}$, $M_{ij} := T^{-1} \sum_{t=1}^{T} Z_{it} Z_{jt}'$. We indicate the eigenvalue problem (6) with the notation $\text{RRR}(Z_{0t}, Z_{1t}; Z_{2t})$. We let $S := (S_{ij})_{i,j=0,1}$ and $M = (M_{ij})_{i,j=0,1,2}$ be matrices with blocks S_{ij} and M_{ij} respectively.

---

[2] See also Mulargia et al. (1992) for an application where the EC form with level term measured in $t-q$ is appropriate.

[3] Usually l in (5) is either taken as $j+1$ or p; Cavaliere, Fanelli and Paruolo (2005) consider the class of tests $LR(j, l)$ for any $l = j+1, ..., p$. All these tests are functions of the eigenvalues $\widehat{\lambda}_i$, and hence inherit their invariance properties.

The invariance property of the eigenvalues $\widehat{\Lambda} := \{\widehat{\lambda}_i\}_{i=1}^p$ as canonical correlations are well known, see Anderson (1984, Theorem 12.2.2); this property is inherited by $LR(j,l)$ as function of $\widehat{\Lambda}$.

Theorem 3 *Consider the eigenvalues $\widehat{\lambda}_i$ in $RRR(Z_{0t}, Z_{1t}; Z_{2t})$ with Z_{it} defined in (3). The eigenvalues $\widehat{\lambda}_i$ are invariant with respect to the following joint transformation of the variables Z_{it}:*

$$Z_{0t} \mapsto H_0 Z_{0t} + H_{02} Z_{2t}, \qquad Z_{1t} \mapsto H_1 Z_{1t} + H_{12} Z_{2t}, \qquad Z_{2t} \mapsto H_2 Z_{2t} \qquad (7)$$

where H_i, $i = 0, 1, 2$ are square invertible matrices. Moreover any function of $S := (S_{ij})_{i,j=0,1}$ which is invariant with respect to the transformation (7) is a function of $\widehat{\Lambda}$.

In particular we consider the transformation $X_t^\circ := HX_t$ and the associated transformations (7) with $H_0 = H$, $H_1 = \text{diag}(H, I_{m_1})$, $H_2 = \text{diag}\left((I_{q-1} \otimes H), I_{m_2}\right)$ with $H_{02} = H_{12} = 0$.

Incidentally we recall that Theorem 3 implies that $RRR(Z_{0t}, Z_{1t}; Z_{2t})$ is invariant with respect to the choice of lag of the level term. Let in fact Z_{1t}° contain the level term X_{t-j} with a given $j \in \{1, 2, ..., q-1\}$; then Theorem 3 gives $RRR(Z_{0t}, Z_{1t}; Z_{2t}) = RRR(Z_{0t}, Z_{1t}^\circ; Z_{2t})$ because $Z_{1t}^\circ = Z_{1t} + H_{12}Z_{2t}$ with $H_{12} := \left((a_j' \otimes I_p) : 0\right)$ where a_j is a $q - 1 \times 1$ vector with $j - 1$ leading zeros and remaining entries equal to 1.

The next section considers MC design reductions for VAR(1) processes.

4   VAR(1) design

In this section we consider the class $\mathcal{P}_1$ with $q = 1$, see (1). We assume that also the statistical model (3) selects $q = 1$. U_{t-1} is dropped form (3) which becomes

$$\Delta X_t = \alpha \begin{pmatrix} \beta' & \rho_1' \end{pmatrix} \begin{pmatrix} X_{t-1} \\ D_{1t} \end{pmatrix} + \mu_2 D_{2t} + \varepsilon_t \qquad (8)$$

Note that $\theta := (\alpha, \beta, \rho_1, \mu_2, \Omega)$ and Θ_1 has MC design dimension $d_{\Theta_1} = (2p - r + m_1) r + pm_2 + \frac{1}{2}p(p+1)$. Here we have used the identification condition on α, β.

In Subsection 4.1 we state how this design dimension can be reduced by means of invariance. In Subsection 4.2 we compute the design reduction. In Subsection 4.3 we discuss further reductions that are possible for special cases. In Subsection 4.4 we compare the present design with an alternative one based on covariances.

4.1   MC design reduction

In this subsection we present the main result. We first define a function $\psi(\theta)$ that is invariant with respect to $\bar{g} \in \bar{G}$; we then show that ψ is maximally invariant. We then apply Theorem 1, and obtain that the distribution of the eigenvalues $\widehat{\Lambda}$ depends only on $\psi(\theta)$. Hence we can restrict attention to the simulation of one DGP for each value of $\psi(\theta) \in F$, instead than of each value of $\theta \in \Theta$.

We first define the following function of $\theta := (\alpha, \beta, \rho_1, \mu_2, \Omega)$,

$$\psi(\theta) := (J, \kappa, \gamma, \zeta, \Phi)(\theta)$$

where:

1. $(J, R)(\theta)$ is the real Jordan pair from the real Jordan decomposition of $\alpha_1^{(1)} + I$, where $\alpha_1^{(1)} := (\beta'\Omega\beta)^{-1/2}\beta'\alpha(\beta'\Omega\beta)^{1/2}$; J is a $r \times r$ real Jordan matrix J and R is the nonsingular transformation that satisfy $\alpha_1^{(1)} + I = RJR^{-1}$. For details on the real Jordan decomposition see e.g. Horn and Johnson (1985) p. 152, Theorem 3.4.5;

2. $(Q, \kappa)(\theta)$, is the QR pair in the QR decomposition of $(\xi'\xi)^{1/2}\eta'$, where $\xi\eta'$ is the rank decomposition of $\alpha_2^{(1)} := (\beta'_\perp \Omega^{-1}\beta_\perp)^{-1/2}\beta'_\perp \Omega^{-1}\alpha(\beta'\Omega\beta)^{1/2}$, of rank $j \le \min(r, p-r)$. κ is a $j \times r$ upper triangular matrix, with positive entries on the main diagonal;

3. $\gamma(\theta) := a^{-1/2}R^{-1}(\beta'\Omega\beta)^{-1/2}\rho_1'$, where $\mu_1 = \alpha\rho_1'$; γ is a $r \times m_1$ matrix;

4. $\zeta(\theta) := H\mu_2$, where $H = H_2 H_1$, and

$$H_1 := \left(\beta\left(\beta'\Omega\beta\right)^{-1/2} : \Omega^{-1}\beta_\perp\left(\beta'_\perp\Omega^{-1}\beta_\perp\right)^{-1/2}\right)', \tag{9}$$

$$H_2 := \operatorname{diag}\left(a^{-1/2}R^{-1}, H_3\right), \qquad H_3 := \left(\xi\left(\xi'\xi\right)^{-1/2}Q : \xi_\perp\left(\xi'_\perp\xi_\perp\right)^{-1/2}\right). \tag{10}$$

ζ is $p \times m_2$.

5. $\Phi(\theta) := a^{-1}R^{-1}R^{-1\prime}$ where a is the first element on the main diagonal of $R^{-1}R^{-1\prime}$. Φ is a $r \times r$ positive definite symmetric matrix, with a 1 in the first entry on the main diagonal. Φ contains $r(r+1)/2 - 1$ free elements.

Let F indicate the parameter space of ψ when θ varies in Θ. We next define the map $\varphi : F \longmapsto \Theta$, $\varphi(\psi) := (\alpha, \beta, \rho_1, \mu_2, \Omega)(\psi)$, that maps ψ back into θ as follows:

$$\alpha(\psi) = ((J - I)' : \kappa' : 0)', \qquad \beta(\psi) = (I_r : 0_{r \times p-r})'$$
$$\rho_1(\psi) = \gamma, \qquad \mu_2(\psi) = \zeta, \qquad \Omega(\psi) = \operatorname{diag}(\Phi, I_{p-r}). \tag{11}$$

We also use the notation $f(\theta) := \varphi \circ \psi(\theta)$.

Theorem 4 *The function $\psi(\theta)$ is invariant with respect to the action of $\bar{g} \in \bar{G}$. The point $f(\theta) := \varphi(\psi(\theta))$ is on the same orbit as θ, i.e. $f(\theta) = \bar{g}(\theta)$ for some $\bar{g} \in \bar{G}$. Moreover ψ is maximally invariant, i.e. $\psi(\theta^\circ) = \psi(\theta)$ for $\theta^\circ, \theta \in \Theta_1$ implies $\theta^\circ = \bar{g}(\theta)$ for some $\bar{g} \in \bar{G}$.*

Theorem 4 partitions Θ through the maximal invariant function $\psi(\theta)$ into sets of the form $\Theta_{\psi^\circ} := \{\theta \in \Theta : \psi(\theta) = \psi^\circ\}$. One can simulate just one process $\mathbb{P}_\theta$ choosing one $\theta = \varphi(\psi^\circ)$ for each Θ_{ψ°, because the distribution of $\widehat{\Lambda}$ is constant over Θ_{ψ° by Theorem 1. In this way all points θ that lie in Θ_{ψ° are represented by a single MC experiment. The next corollary describes the representative DGP in Θ_{ψ°. Here and hereafter we omit zero entries for readability.

Corollary 5 *In order to draw from the distribution $\mathbb{P}_{\varphi(\psi)}$ for a fixed value of $\psi = (J, \kappa, \gamma, \zeta, \Phi)$, one can simulate a Gaussian VAR(1) process W_t with dynamics*

$$\begin{cases} W_{1t} = & J\left(W_{1t-1} + \gamma D_{1t}\right) + \zeta_1 D_{2t} + v_{1t} \\ \Delta W_{2t} = & \kappa\left(W_{1t-1} + \gamma D_{1t}\right) + \zeta_2 D_{2t} + v_{2t} \\ \Delta W_{3t} = & \zeta_3 D_{2t} + v_{3t} \end{cases}, \tag{12}$$

$$\Sigma := \operatorname{var}\begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix} = \begin{pmatrix} \Phi & & \\ & I_r & \\ & & I_{p-2r} \end{pmatrix}$$

9

where $\zeta := (\zeta_1' : \zeta_2' : \zeta_3')'$, W_{1t} *is* $r \times 1$, W_{2t} *is* $j \times 1$, W_{3t} *is* $p - r - j \times 1$ *and* $\upsilon_t := H\varepsilon_t \sim N(0, \Sigma)$. *The eigenvalues in* J *are less of equal to 1 in modulus by the assumption that all the measures in* $\mathcal{P}_1$ *satisfy GRT.*

The representative DGP is thus (12) with a fixed value of ψ. We collect observations and comments about (12) in the following remarks.

1. The matrix J is associated with stationary dynamics. The eigenvalues in J are all less than 1 in modulus because of assumption (a) in GRT. The latter can be seed to be equivalent for VAR(1) processes to $|eig(I + \beta'\alpha)| < 1$, see Johansen (1996) Exercise 4.12, where $eig(\cdot)$ indicates a generic eigenvalue of the argument. We note here that $eig(I + \beta'\alpha) = eig(\alpha_1^{(1)} + I)$.

2. We next describe the form of J. Let λ_k be a generic eigenvalue (less than 1 in modulus) of $\alpha_1^{(1)} + I$ with dimension n_k of the corresponding Jordan block, $r = \sum_k n_k$ see e.g. Horn and Johnson (1985) p. 126, Theorem 3.1.11.

   The eigenvalues λ_k are not necessarily real. Let $\lambda_s, ..., \lambda_r$ be the ordered real eigenvalues with $1 \leq s \leq r$. Let $\lambda_j := a_j + ib_j = c_j (\cos \omega_j + i \sin \omega_j)$ be a generic complex eigenvalue, where a_j, b_j are the real and imaginary parts, c_j is the modulus and $\omega_j = \arg(\lambda_j)$, $0 \leq \omega_j < 2\pi$. Because the matrix $\alpha_1^{(1)} + I$ is real, complex eigenvalues λ_j always appear with their complex conjugate $\bar{\lambda}_j = a_j - ib_j$.[4]

   Let h be the number of complex pairs of eigenvalues with separate Jordan blocks in the Jordan canonical form (not necessarily distinct). The form of J is

   $$\underset{r \times r}{J} := \operatorname{diag}\left(C_{n_1}(a_1, b_1), ..., C_{n_h}(a_h, b_h), J_{n_s}(\lambda_s), ..., J_{n_r}(\lambda_r)\right), \text{ where}$$

   $$\underset{2k \times 2k}{C_k(a, b)} := \begin{pmatrix} C(a, b) & I_2 & & \\ & \ddots & \ddots & \\ & & \ddots & I_2 \\ & & & C(a, b) \end{pmatrix}, \tag{13}$$

   $$\underset{2 \times 2}{C(a_j, b_j)} := \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} =: compl(a_j, b_j) = c_j \begin{pmatrix} \cos \omega_j & \sin \omega_j \\ -\sin \omega_j & \cos \omega_j \end{pmatrix}$$

   $$\underset{k \times k}{J_k(\lambda)} := \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

   where we have reported dimensions below the matrices on the l.h.s.

3. We observe that the matrix J has a upper-block-triangular structure. The real blocks are upper triangular, while the complex blocks $C_k(a, b)$ are block-upper-triangular, with bivariate blocks. We call this structure 2-block triangular, or block triangular of dimension 2.

---

[4]One can note that the tripler (c_j, ω_j, n_j) completely characterizes $C_{n_j}(c_j \cos \omega_j, c_j \sin \omega_j)$, and it is associated in a simple way with stationary ($c_j < 1$), explosive ($c_j > 1$) and unit ($c_j = 1$) roots. For unit roots, the order of integration is associated with n_j, i.e. the sub-process obtained by selecting the block of variables corresponding fo $C_{n_j}(c_j \cos \omega_j, c_j \sin \omega_j)$ is integrated of order $I(n_j)$, see e.g. Bauer and Wagner (2005).

4. Complex Jordan blocks $C_k(a, b)$ are 2-block triangular. Hence one needs to use bivariate algorithms to produce a simulation of each bivariate block given the following ones in the ordering of $C_k(a, b)$ in (13). Hence one cannot in general simulate all I(1) VAR(1) processes only using univariate procedures, unless all eigenvalues in J are real.

5. The block-triangular structures in the real Jordan decomposition of W_{1t} can speed calculations considerably with respect to non-block-triangular forms. This is well known for the computation of solutions to linear systems. Here it is even more important given that recursions are needed in order to generate autoregressive series.

6. The matrix κ is a loading matrix of the differences of the second set ΔW_{2t} of j variables on the lagged values of W_{1t}. Also here one can adopt a block-recursive approach in computations; one can first generate W_1^*, a $T \times r$ matrix, and v_2^* a $T \times j$ matrix of errors, with t-th rows respectively equal to W_{1t}' and v_{2t}' respectively; next compute $W_{1,lagged}^*$ as $(0 : W_{11} : \ldots : W_{1T-1})'$ and $e_2 = W_{1,lagged}^* \kappa' + v_2^*$; finally calculate the $T \times j$ matrix W_2 as the cumulative sum of e_2. This just requires matrix multiplication and a single cumulative sum.

7. Simulation of W_3 is independent of W_1, W_2 (both stochastically and computationally), and can be performed as cumulative sum of a $T \times p - r - j$ matrix with t-th row equal to v_{3t}'.

8. Combining 5., 6., 7. one sees that recursions are needed only for the W_{1t} block, where the upper-(block-)triangular form of J implies savings in computing time.

We next illustrate simple cases for J, for $r = 2, 3$. If $r = 2$, one may have the following cases.

i. Real eigenvalues:
$$J = \begin{pmatrix} \lambda_1 & c \\ & \lambda_2 \end{pmatrix}$$
where $c = 0$ or $c = 1$; if $c = 1$ then $\lambda_1 = \lambda_2$.

ii. Complex conjugate eigenvalues: $\lambda_1 = a + bi$, $\lambda_2 = a - bi$ with
$$J = C(a, b) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

If $r = 3$, one may have 0 or 2 (conjugate) complex eigenvalues. These are the possible situations:

iii. One real, two complex conjugate eigenvalues:
$$J = \begin{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} & \\ & \lambda_3 \end{pmatrix}$$

iv. Three real eigenvalues:
$$J = \begin{pmatrix} \lambda_1 & c_1 & 0 \\ 0 & \lambda_2 & c_2 \\ 0 & 0 & \lambda_3 \end{pmatrix} \tag{14}$$
where c_1 and c_2 can be either 0 or 1. If $c_1 = 1$ then $\lambda_1 = \lambda_2$, if $c_2 = 1$, then $\lambda_2 = \lambda_3$. Finally $c_1 = c_2 = 1$ implies $\lambda_1 = \lambda_2 = \lambda_3$.

$p - r$	$r = 0$	1	2	3	4	5	6	7
1	3	4	4	7	9	14	18	25
2	5	9	12	18	23	31	38	48
3	8	15	21	30	38	49	59	72
4	12	22	31	43	54	68	81	97
5	17	30	42	57	71	88	104	123
6	23	39	54	72	89	109	128	150
7	30	49	67	88	108	131	153	178

Table 1: Lower bound m_0 for MC design dimension reduction.

4.2 Dimension comparison

This subsection calculates MC design reduction $d_{\Theta_1} - d_F$ associated with (12). We observe that the deterministic parts of (8) and (12) have the same number of parameters: γ, ζ have the same number of free elements as ρ_1 and μ_2. We hence restrict attention to the case of no deterministics for dimension comparisons.

We consider the representation of the various elements J, κ, Φ within ψ, in order to calculate d_F. Consider first the Jordan matrix J; observe that each eigenvalue λ_i in J can be either real or complex, where complex eigenvalues are conjugate. We hence group eigenvalues in pairs, associating complex conjugate numbers.

Each generic pair, indicated as (λ_1, λ_2), can be represented in $\Re^3$: the first coordinate represents $\Re(\lambda_1)$, the second one $\Re(\lambda_2)$, the third one $\Im(\lambda_1)$. When the eigenvalues are real, the third coordinate is equal to 0. When (λ_1, λ_2) are complex conjugate, the first and second coordinate are equal, $\Im(\lambda_1)$ is reported on the third axis and $\Im(\lambda_2) = -\Im(\lambda_1)$ is a function of $\Im(\lambda_1)$.

This implies that $2\lfloor r/2 \rfloor + r$ coordinates are needed to represent the entries on the main diagonal of J, where $\lfloor \cdot \rfloor$ indicates the largest lower integer value. The matrix J contains also 0 or 1 in each entry on the first super-diagonal. One hence needs $r - 1$ indicators that take values in $\{0, 1\}$. For dimension comparisons, we embed the discrete set $\{0, 1\}$ in $\Re$ and count a total of $2\lfloor r/2 \rfloor + 2r - 1$ coordinates in ψ to represent J.

The matrix κ has $j(j+1)/2 + \max(r - j, 0)^2$ elements different from 0, where $0 \le j \le \min(r, p - r)$ takes all different values on Θ. For dimension comparisons we reserve $r(r+1)/2$ entries in ψ to represent κ; this choice is conservative, in the sense that when $p - r < r$, fewer coordinates are really needed. Finally we recall that Φ corresponds to $r(r+1)/2 - 1$ coordinates in ψ.

The above gives $d_F = 2\lfloor r/2 \rfloor + 2r - 1 + r(r+1) - 1 + rm_1 + pm_2$; this must be compared with d_{Θ_1} from (2). In both dimensions we have also added the dimension of the deterministic components for completeness. One finds the following corollary.

Corollary 6 *If $r < p$ then*

$$m_0 := d_{\Theta_1} - d_F = 3(p - r - 1)r + (p - r)^2/2 + \left(r^2/2 - \lfloor r/2 \rfloor\right) + (p/2 - \lfloor r/2 \rfloor) + 2 > 0,$$

where $\lfloor \cdot \rfloor$ indicates the largest lower integer value.

Table 1 reports the lower bound of the MC design reduction m_0; it is seen that m_0 can be substantial.

4.3 Special cases and further reductions

The (maximal) invariant function ψ in Subsection 4.1 provides a partition of the parameter space Θ_1. For special subsets of Θ_1, there exist other invariant functions

of θ that imply an even greater MC design reduction. These special cases are treated in this subsection.

In particular we discuss two nested special cases. Let $\Theta_2 := \{\theta : J(\theta) = \mathrm{diag}\}$ indicate the subset of Θ of points that correspond to a diagonal Jordan matrix J in $\psi(\theta)$. Define also $\Theta_3 := \{\theta : J(\theta) = cI_r\}$ as the subset of Θ of points that correspond to a Jordan matrix J proportional to the identity matrix. Let $\mathcal{P}_i := \{\mathbb{P}_\theta, \theta \in \Theta_i\}$ be the corresponding subsets of probability measures, $i = 2, 3$. Note that $\Theta_3 \subset \Theta_2 \subset \Theta_1$; moreover for $\theta \in \Theta_2$ or Θ_3, J has all blocks of dimension one and only real eigenvalues.

Let $\psi_2(\psi)$ indicate the function $\psi_2(\psi) := (J, \kappa^*, \gamma^*, \zeta^*, \Phi^*)(\psi)$ that leaves the J coefficients in ψ unaltered. κ is mapped into $\kappa^* := H_4 \kappa H_3^{-1}$ and $\Phi := (\phi_{ij})_{i,j=1,\dots,r}$ is mapped into the correlation matrix $\Phi^* = H_3 \Phi H_3$ where $H_3 := \mathrm{diag}(s_1 \phi_{11}^{-1/2}, \dots, s_r \phi_{rr}^{-1/2})$, $H_4 := \mathrm{diag}(s_1, \dots, s_j)$, ϕ_{ii} are the diagonal elements of Φ; moreover $s_i := \mathrm{sign}(\phi_{1i})$, where $\mathrm{sign}(a) := 1$ if $a \geq 0$ and $\mathrm{sign}(a) := -1$ otherwise, $a \in \Re$. Moreover $\gamma^* := H_3 \gamma$, $\zeta^* := \mathrm{diag}(H_3, H_4, I_{p-r-j})\zeta$.

The matrix Φ^*

$$\Phi^* = \begin{pmatrix} 1 & \phi_{12}^* & \dots & & \phi_{1r}^* \\ \phi_{21}^* & \ddots & & & \\ \vdots & & \ddots & & \phi_{r-1,r}^* \\ \phi_{r1}^* & \dots & & \phi_{r,r-1}^* & 1 \end{pmatrix}$$

has non-negative correlations in the first row and column. The choice of the first variable for the non-negativity restriction is arbitrary. Note also that κ^* has the same characteristics as κ, i.e. it is upper triangular with non-negative entries on the main diagonal, because H_3, H_4 are diagonal.

Because ψ_2 is a function of $\psi(\theta)$, with a slight abuse of notation we also indicate by $\psi_2(\theta)$ the composite map $\psi_2 \circ \psi(\theta)$. Note that $\psi_2(\theta)$ is $\bar{g} \in \bar{G}$ invariant, being a function of ψ. Remark that Θ_2 is preserved by $g \in G$. In fact $\bar{g}(\theta) \in \Theta_2$ for all $\theta \in \Theta_2$ because $J(\theta) = J(\bar{g}(\theta))$ is invariant; moreover given $\theta^\circ \in \Theta_2$ one can find a $\theta \in \Theta_2$ such that $\theta^\circ = \bar{g}(\theta)$. To verify the latter claim, simply note that $\bar{g}$ can be inverted to give $\theta = \bar{g}^{-1}(\theta^\circ)$, which still belongs to Θ_2 because of the invariance of J under $\bar{g} \in \bar{G}$, such that $\theta^\circ = \bar{g}(\theta)$.

Define also $f_2(\theta) := \varphi(\psi_2(\theta))$, which maps back ψ_2 in θ-space. We next state results concerning $\mathcal{P}_2$.

Theorem 7 $f_2(\theta) := \varphi(\psi_2(\theta))$ *is on the same orbit as θ for all $\theta \in \Theta_2$, i.e. $f_2(\theta) = \bar{g}(\theta)$ for some $\bar{g} \in \bar{G}$. Moreover ψ_2 is maximally invariant on Θ_2, i.e. $\psi_2(\theta^\circ) = \psi_2(\theta)$ for $\theta^\circ, \theta \in \Theta_2$ implies $\theta^\circ = \bar{g}(\theta)$ for some $\bar{g} \in \bar{G}$. Hence Φ can be replaced by a correlation matrix Φ^* in the MC design (12) when $\theta \in \Theta_2$.*

Note that the above simplification reduces the MC design dimension by $r - 1$ by eliminating the variances on the main diagonal of Φ; moreover it restricts the $r - 1$ elements on the first row and column to the half-line $\Re_0^+$.

We next consider the case of $\mathcal{P}_3$, when J is a scalar multiple of the identity, i.e. $J = \lambda I_r$. Let $\psi_3(\psi)$ indicate a function $\psi_3(\psi) := (J, \kappa^*, \gamma^*, \zeta^*, \Phi^*)(\psi)$ that leaves J in ψ unaltered. κ is mapped into $\kappa^* := \kappa V^{-1}$ and Φ is mapped into the identity matrix I_r where V is an upper triangular matrix with positive elements on the main diagonal that satisfies $V \Phi V' = I_r$, see Lemma 16 in the Appendix. Note also that κ^* has the same characteristics as κ, i.e. it is upper triangular with non-negative entries on the main diagonal, because V is upper triangular with positive entries on the main diagonal. Moreover $\gamma^* := V\gamma$, $\zeta^* := \mathrm{diag}(V, I_{p-r})\zeta$.

Again, because ψ_3 is a function of $\psi(\theta)$, with a slight abuse of notation we also indicate by $\psi_3(\theta)$ the composite map $\psi_3 \circ \psi(\theta)$. Note that $\psi_3(\theta)$ is $\bar{g} \in \bar{G}$ invariant,

being a function of ψ. Note also that Θ_3 is preserved by $g \in G$. In fact $\bar{g}(\theta) \in \Theta_3$ for all $\theta \in \Theta_3$ because $J(\theta) = J(\bar{g}(\theta))$ is $\bar{g} \in \bar{G}$ invariant; moreover given $\theta^{\circ} \in \Theta_3$ one can find a $\theta \in \Theta_3$ such that $\theta^{\circ} = \bar{g}(\theta)$. To verify the latter claim, simply note that $\bar{g}$ can be inverted to give $\theta = \bar{g}^{-1}(\theta^{\circ})$, which still belongs to Θ_3 because of the invariance of J under $\bar{g} \in \bar{G}$, such that $\theta^{\circ} = \bar{g}(\theta)$.

Define also $f_3(\theta) := \varphi(\psi_3(\theta))$, which maps back in θ-space. We next state results concerning $\mathcal{P}_3$.

Theorem 8 $f_3(\theta) := \varphi(\psi_3(\theta))$ *is on the same orbit as* θ *for all* $\theta \in \Theta_3$. *Moreover* ψ_3 *is maximally invariant on* Θ_3, *i.e.* $\psi_3(\theta^{\circ}) = \psi_3(\theta)$ *for* $\theta^{\circ}, \theta \in \Theta_3$ *implies* $\theta^{\circ} = \bar{g}(\theta)$ *for some* $\bar{g} \in \bar{G}$. *Hence* Φ *can be replaced by an identity matrix* I *in the MC design* (12) *when* $\theta \in \Theta_3$.

Note that Theorem 8 reduces the design dimension by $r(r+1)/2 - 1$ with respect to (12) by eliminating Φ.

Finally note that the above further reduction can be applied to single real Jordan blocks that are diagonal; this implies a comparatively smaller reduction in the number of covariance parameters in Φ.

4.4 Covariance versus AR parametrization

An alternative parametrization, which eliminates κ at the expense of creating extra non-zero covariances, is analyzed in this subsection. This alternative is explored because eliminating AR coefficients implies reducing recursions and hence saves computer time. The cost of this elimination is the creation of some extra non-zero covariances of the errors, which are less expensive to generate in terms of computer time. However we find that this alternative parametrization has higher dimension than $\dim(F)$ in (12); moreover it is less well interpretable. The covariance parametrization is hence less attractive than the AR one given previously. This subsection presents these results.

The covariance parametrization is obtained by a suitable g transformation of W_t in (12). Observe that $I - J$ is nonsingular, thanks to GRT. We consider the transformation $Y_t = HW_t$, with

$$
H = \begin{pmatrix} I_r & & \\ \kappa(I-J)^{-1} & I_j & \\ & & I_{p-r-j} \end{pmatrix}
$$

One finds $Y_{1t} = W_{1t}$, $Y_{3t} = W_{3t}$ and

$$
\Delta Y_{2t} = \zeta_2^* D_{2t} + v_{2t}^*
$$

where $\zeta_2^* := \kappa(I-J)^{-1}\zeta_1 + \zeta_2$, $v_{2t}^* := \kappa(I-J)^{-1}v_{1t} + v_{2t}$. The covariance matrix of the errors $v_t^* := Hv_t$ is equal to $\Sigma^* := (\Sigma_{ij}^*)_{i,j=1,2,3} := \mathrm{var}(v_t^*)$ where

$$
\Sigma^* := \begin{pmatrix} \Phi & \Phi(I-J')^{-1}\kappa' & \\ \kappa(I-J)^{-1}\Phi & I_j + \kappa(I-J)^{-1}\Phi(I-J')^{-1}\kappa' & \\ & & I_{p-r-j} \end{pmatrix}. \tag{15}
$$

Note that ΔY_{2t}, unlike ΔW_{2t}, does not contain the term κW_{1t-1}. The resulting system has no AR coefficients in the equations for ΔY_{2t}, i.e. it has eliminated κ. This was done at the expense of introducing variances for the second block of errors corresponding to Y_{2t} (of dimension $j \times 1$) and between the first block Y_{1t} and Y_{2t}.

14

The alternative parametrization $\psi^*(\theta)$ thus obtained is given by $\psi^*(\theta) := (J, \gamma, \zeta^*, \Sigma_{11}^*, \Sigma_{12}^*, \Sigma_{22}^*)$, where $\Sigma_{11}^* = \Phi$ and $\Sigma_{12}^*, \Sigma_{22}^*$ are defined in (15). It can then be proved that this alternative parametrization $\psi^*(\theta)$ is maximally invariant in Θ_1; this is omitted for brevity.

We now wish to compare design dimensions in (12) and (15) where, without loss of generality, we abstract from the coefficients to D_{1t} and D_{2t}. $J, \gamma, \zeta^*, \Sigma_{11}^* = \Phi$ are present both in W_t and Y_t. The extra variances and covariances $\Sigma_{12}^*, \Sigma_{22}^*$ in (15) (when unrestricted) generate $rj + j(j+1)/2$ dimensions, which should be compared with the number of free elements in κ; the latter is bounded by rj. Hence the covariance parametrization (15) gives a higher MC design dimension than (12).

Note also that the AR parametrization is amenable to a directional dependence interpretation, i.e. ΔW_{2t} adjusts to W_{1t-1} through κ. On the contrary the covariances in (15) give a-directional measures of associations between the errors in Y_{1t} and Y_{2t}. The parametrization (12) should hence be preferred both because it gives a lower MC design dimension and because the AR coefficients in κ are more directly interpretable than the covariances in (15).

5 An application to asymptotics for local alternatives

In this section we show how the results in Section 4 developed for VAR(1) processes can be used to index the asymptotic distribution of LR CI rank test, under local alternatives.

The local alternative is defined substituting $\alpha\beta'$ with $\alpha\beta' + T^{-1}\alpha_1\beta_1'$, where α_1, β_1 are $p \times s$ full column rank matrices, $\alpha_1 \in \mathrm{col}(\alpha_\perp)$, $\beta_1 \in \mathrm{col}(\beta_\perp)$. Let $eig_i(N)$ the i-th largest eigenvalue, in case N has all real eigenvalues.

Define $K(t)$, $t \in \mathcal{T} := [0,1]$ as the diffusion

$$K(t) = ab' \int_0^t K(u)du + V(t), \qquad t \in [0,1] \tag{16}$$
$$a := \alpha_\perp'\alpha_1 \qquad b' := \beta_1'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}.$$

Here $V(t)$ is a $(p-r)\times 1$ Brownian motion with covariance $\Sigma := \alpha_\perp'\Omega\alpha_\perp = \mathbb{E}(V(1)V(1)')$. Let $\mathbb{P}_\theta$ indicate the probability measure defined by (16), where $\theta := (a, b, \Sigma)$ and let $\mathbb{E}(\cdot)$ denote expectations with respect to it. Here a and b are $p-r \times s$ matrices of full column rank that satisfy $\mathrm{rk}(b'a) = s$, and Σ is a positive definite symmetric matrix of dimension $p - r$. This parameter space is identical to Θ_1 (apart from dimensions) when $q = 1$ and all the deterministic terms ρ_1 and μ_2, are canceled.[5] Indicate this parameter space as Θ_4 and let $\mathcal{P}_4 := \{\mathbb{P}_\theta, \theta \in \Theta_4\}$.

If $|eig(I_r + \beta'\alpha)| < 1$, see Johansen (1996 Chapter 14) and Cavaliere, Fanelli, Paruolo (2005), then as $T \to \infty$,

$$LR(r,l) \xrightarrow{w} V_{l-r} := \sum_{i=1}^{l-r} eig_i(N(K)),$$

$$N(K) := \left(\int_0^1 K(u)K(u)'du\right)^{-1} \int_0^1 K(u)(dK(u))' \Sigma^{-1} \int_0^1 (dK(u))K(u)'$$

---

[5] The present results easily generalize when the local asymptotics are performed with deterministic terms as in Saikkonen and Lütkepohl (2000).

Eq. (16) can be seen as a continuos time analog of the VAR(1). We here show in the following three subsections how the results developed for a VAR(1) apply also in this continuous time case.

5.1 Transformations

We consider the transformation g that takes the process $K(t)$ into $HK(t)$, for H square and invertible; g defined the group G. The following is the analogue of Theorem 2.

Theorem 9 *Let $K(t)$ be a process with probability measure $\mathbb{P}_\theta \in \mathcal{P}_4$, $\theta := (a,\, b,\, \Sigma) \in \Theta_4$; the transformed process $HK(t)$ with H square and invertible has probability measure $\mathbb{P}_{\theta^\circ} \in \mathcal{P}_4$, where $\theta^\circ := \bar{g}(\theta) = (Ha,\, H'^{-1}b,\, H\Sigma H') \in \Theta_4$.*

This shows that the diffusion process (16) is closed under linear invertible transformations, i.e. that g preserves Θ_4.

5.2 Invariance

It is simple to verify that $eig_i(N(K))$ are invariant with respect to this group of transformations. Note in fact that

$$N(K^\circ) = H'^{-1}N(K)H'$$

and that the eigenvalues are invariant with respect to pre-multiplication by H'^{-1} and post- multiplication by H', $eig_i(H'^{-1}NH') = eig_i(N)$. Hence also the random variable V_{l-r} in (16), which is a function of $eig_i(N)$, is invariant under G.

5.3 MC design reduction

One can hence apply Theorem 3 to the present case. Let again ψ be defined as $\psi(\theta) := (J, \kappa, \Phi)(\theta)$. Here all the definition in Section 4.1 apply, substituting Σ in place of Ω, s in place of r and $p - r$ in place of p. Let also $F := \{\psi(\theta), \theta \in \Theta_4\}$.

Theorem 4 partitions Θ_4 through the maximal invariant function $\psi(\theta)$ into sets of the form $\Theta_{4,\psi^\circ} := \{\theta \in \Theta_4 : \psi(\theta) = \psi^\circ\}$. One can simulate just one process $\mathbb{P}_\theta$ choosing one $\theta = \varphi(\psi^\circ)$ for each Θ_{ψ°, because the distribution of V_{l-r} is constant over Θ_{4,ψ° by Theorem 1. In this way all points θ that lie in Θ_{4,ψ° are represented by a single MC experiment. The next corollary describes the representative DGP in Θ_{4,ψ°.

Corollary 10 *In order to draw from the distribution $\mathbb{P}_{\varphi(\psi)}$ for a fixed value of $\psi = (J, \kappa, \Phi)$, one can simulate a diffusion process $K(t)$ as in (16) with $a = (J' - I_s : \kappa' : 0)'$, $b = (I_s : 0)'$, $\Sigma = \mathrm{diag}(\Phi_{s \times s}, I_{p-r-s})$. The eigenvalues in J are less or equal to 1 in modulus by the assumption $|eig(I_r + \beta'\alpha)| < 1$.*

6 VAR(q) design

In this section we discuss how the results in the previous sections can be extended to VAR(q) processes for $q \geq 2$. For VAR(q) one wishes to obtain parametrizations which directly control the stable characteristic roots of $|A(z)| = 0$.

The approach we take is to consider the companion form of the VAR, and then modify the techniques introduced in the previous sections in order to (possibly) reduce

the MC design dimension and, at the same time, obtain designs where one can directly control the stable roots.

The choice of the companion form is intuitive, although a moment reflection suggests that this increases the dimensionality of the problem. In fact the companion matrix is $(n+p) \times (n+p)$ where $n := p(q-1)$. For increasing q, the elements of the companion form increase as $p^2 q^2$, whereas the AR parameters $A_1, ..., A_q$ increase as $p^2 q$.

One hence may wish to consider other techniques that reduce the MC design dimension without inflating the dimension of the problem to the companion form; this is however beyond the scope of the present paper. We here restrict attention to the type of techniques used for VAR(1), when applied to the companion form. We obtain results similar in spirit to the ones of Section 4.

We obtain a MC dimension reduction when comparing the companion form parametriation, indicated as $\widetilde{\theta} \in \widetilde{\Theta}$, with the one obtained by invariance, $\widetilde{\psi} \in \widetilde{F}_1$. In general there is not, however, a positive MC reduction when comparing $\widetilde{\psi} \in \widetilde{F}_1$ with the direct parametration $\theta \in \Theta_1$ introduced in Section 3.1.

This section is organized as follows. In Subsection 6.1 we first review a companion form representation and state GRT in terms of it. This allows to define the enlarged parametrization $\widetilde{\theta} \in \widetilde{\Theta}$ in Subsection 6.2. We next present extensions of the invariance of RRR in Subsection 6.3; finally in Subsection 6.4 we present results for MC design reduction.

6.1 Companion form

In this subsection we present a state-space formulation for an I(1) VAR(q). Many state-space formulations exist for an I(1) VAR(q). We here choose a particular one which is simple to work with, see (18) below.

Consider the process $\widetilde{X}_t := (\Delta X_t' : \Delta X_{t-1}' : ... : \Delta X_{t-q+2}' : X_{t-q+1}')' = (U_t' : X_{t-q+1}')'$; the state vector $\widetilde{X}_t$ is $(n+p) \times 1$, and satisfies the companion form:

$$
\begin{pmatrix} \Delta X_t \\ \vdots \\ \Delta X_{t-q+2} \\ X_{t-q+1} \end{pmatrix} = \begin{pmatrix} \Psi_1 & ... & \Psi_{q-1} & \alpha\beta' \\ I_p & & & \\ & \ddots & & \\ & & I_p & I_p \end{pmatrix} \begin{pmatrix} \Delta X_{t-1} \\ \vdots \\ \Delta X_{t-q+1} \\ X_{t-q} \end{pmatrix} + \begin{pmatrix} \mu \\ \end{pmatrix} D_t + \begin{pmatrix} \varepsilon_t \\ \end{pmatrix} \text{ i.e.}
$$

$$
\widetilde{X}_t = \widetilde{A}\widetilde{X}_{t-1} + \widetilde{\mu}D_t + \widetilde{\varepsilon}_t
$$

The matrix $\widetilde{A}$ is called the companion matrix. Again here we partition the deterministic components as $\mu D_t = \alpha\rho_1' D_{1t} + \mu_2 D_{2t}$. Let $U := (e_1 \otimes I_p)$, where e_1 is $(q-1) \times 1$ with 1 in the first entry and 0 elsewhere; with this notation $\widetilde{\varepsilon}_t = U\varepsilon_t$, $\widetilde{\mu}_2 D_{2t} = U\mu_2 D_{2t}$. Let also $\widetilde{\varepsilon}_t^* := U(\mu_2 D_{2t} + \varepsilon_t)$. The EC form of the companion equation is

$$
\begin{pmatrix} \Delta^2 X_t \\ \vdots \\ \Delta^2 X_{t-q+2} \\ \Delta X_{t-q+1} \end{pmatrix} = \begin{pmatrix} \Psi_1 - I_p & ... & \Psi_{q-1} & \alpha \\ I_p & -I_p & & \\ & \ddots & -I_p & \\ & & I_p & 0 \end{pmatrix} \begin{pmatrix} I_n & \\ \beta' & \rho_1' \end{pmatrix} \begin{pmatrix} \Delta X_{t-1} \\ \vdots \\ \Delta X_{t-q+1} \\ X_{t-q} \\ D_{1t} \end{pmatrix} + \widetilde{\varepsilon}_t^*
$$

$$
\Delta\widetilde{X}_t = \widetilde{\alpha} \begin{pmatrix} \widetilde{\beta}' & \widetilde{\rho}_1' \end{pmatrix} \begin{pmatrix} \widetilde{X}_{t-1} \\ D_{1t} \end{pmatrix} + \widetilde{\mu}_2 D_{2t} + \widetilde{\varepsilon}_t \tag{17}
$$

where

$$\widetilde{\alpha} := \begin{pmatrix} \Psi_1 - I_p & \dots & \Psi_{q-1} & \alpha \\ I_p & -I_p & & \\ & \ddots & -I_p & \\ & & I_p & \end{pmatrix}, \qquad \widetilde{\beta} := \begin{pmatrix} I_n & \\ & \beta \end{pmatrix} \qquad \widetilde{\rho}_1' := \begin{pmatrix} 0 \\ \rho_1' \end{pmatrix} \qquad (18)$$

6.2 An enlarged class of probability measures

In this subsection we discuss the extended parameterization associated with the companion form (17). It is simple to verify that the conditions (a), (b), (c) for X_t can be restated in terms of the companion form coefficients $\widetilde{\alpha}$ and $\widetilde{\beta}$ as follows:[6]

(a1) $\widetilde{A}$ has all (non-zero) eigenvalues within the unit disc, $|eig(\widetilde{A})| < 1$, or $eig(\widetilde{A}) = 1$, where $eig(\cdot)$ indicates a generic eigenvalues of the argument matrix;

(b1) $I_{pq} - \widetilde{A}$ has reduced rank $n + r$, so that it allows representation $I_{pq} - \widetilde{A} = -\widetilde{\alpha}\widetilde{\beta}'$ for $\widetilde{\alpha}$, $\widetilde{\beta}$ full column rank $n + p \times n + r$ matrices;

(c1) $\widetilde{\beta}'\widetilde{\alpha}$ has full rank (or equivalently $\widetilde{\alpha}_\perp'\widetilde{\beta}_\perp$ has full rank).

Concerning (a1) we note that $eig(\widetilde{A}) = eig(\widetilde{A}-I)+1$, where $eig(\widetilde{A}-I) = eig(\widetilde{\alpha}\widetilde{\beta}')$ are either 0 or equal to $eig(\widetilde{\beta}'\widetilde{\alpha})$. Hence (a1) and (c1) can be reformulated also as follows:

(a2) $|eig(I_{n+r} + \widetilde{\beta}'\widetilde{\alpha})| < 1$ or $eig(I_{n+r} + \widetilde{\beta}'\widetilde{\alpha}) = 1$;

(c2) $eig(I_{n+r} + \widetilde{\beta}'\widetilde{\alpha}) \neq 1$.

We condense the set of conditions (a) (b) (c) into the following two conditions on the companion matrix:

(b3) $I_{pq} - \widetilde{A} = -\widetilde{\alpha}\widetilde{\beta}'$ for $\widetilde{\alpha}$, $\widetilde{\beta}$ full column rank $n + p \times n + r$ matrices;

(a3) (c3) $|eig(I_{n+r} + \widetilde{\beta}'\widetilde{\alpha})| < 1$

We decompose $\widetilde{\Omega} = \text{diag}(\Omega, 0)$ as $\widetilde{\Omega} = \widetilde{\Upsilon}\,\text{diag}(I_p, 0)\widetilde{\Upsilon}'$, where

$$\widetilde{\Upsilon} := \left(\widetilde{\Upsilon}_1 : \widetilde{\Upsilon}_2\right) := \left(U\Omega^{1/2} : U_\perp\right), \qquad \widetilde{\mu}_2 = U\mu_2. \qquad (19)$$

Recall here that $U := (e_1 \otimes I_p)$. We take $\widetilde{\theta} := (\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\rho}_1, \widetilde{\mu}_2, \widetilde{\Upsilon})$ satisfying (a3) (b3) (c3) with $\widetilde{\Upsilon}$ of full column rank $n+p$; this defines $\widetilde{\Theta}$. Note that the first p columns $\widetilde{\Upsilon}_1$ of $\widetilde{\Upsilon}$ are associated with non-zero covariances, and the remaining n columns $\widetilde{\Upsilon}_2$ premultiply the singular part of $\widetilde{\varepsilon}_t$. We label $\mathbb{P}_{\widetilde{\theta}}$ as the Gaussian measure on $\widetilde{X}$ induced by $\widetilde{\varepsilon}_t$ in (17), and call the associated set of probability measures $\widetilde{\mathcal{P}} := \{\mathbb{P}_{\widetilde{\theta}}, \widetilde{\theta} \in \widetilde{\Theta}\}$.

Define also $k(\theta)$ as the function that maps $(\alpha, \beta, \rho_1, \Psi, \mu_2, \Omega) \in \Theta_1$ into $\widetilde{\theta} \in \widetilde{\Theta}$, where $\widetilde{\alpha}$, $\widetilde{\beta}$, $\widetilde{\rho}_1$ are defined in (18) and $\widetilde{\mu}_2$, $\widetilde{\Upsilon}$ are defined in (19). Define the set $\widetilde{\Theta}_1 := \{\widetilde{\theta} : \widetilde{\theta} = k(\theta), \theta \in \Theta_1\}$, which is a proper subset of $\widetilde{\Theta}$, $\widetilde{\Theta}_1 \subset \widetilde{\Theta}$, $\mathcal{P}_1 \subset \widetilde{\mathcal{P}}$. In other words the $\widetilde{\Theta}$ parametrization includes also processes that cannot be represented

---

[6]This formulation is well known, see e.g. Hansen (2005), Lemma A.1 and A.2.

as a VAR(q), which are characterized by the pattern of 0 and I matrices in the lower part of (17) and (18).

As for $\theta \in \Theta$, the $\widetilde{\theta} \in \widetilde{\Theta}$ parametrization is not identified, because each $\mathbb{P}_{\widetilde{\theta}}$ depends on $\widetilde{\alpha}$, $\widetilde{\beta}$ only through their product $\widetilde{\alpha}\widetilde{\beta}'$; moreover in the $\widetilde{\theta}$ parametrization $\mathbb{P}_{\widetilde{\theta}}$ is invariant with respect to the choice of $\widetilde{\Upsilon}_2$ provided $\widetilde{\Upsilon}$ is of full rank. This defines equivalence classes and we assume that parameters values are treated as identical if they belong to the same class.

The dimension $d_{\widetilde{\Theta}}$ of the parameterization $\widetilde{\theta} \in \widetilde{\Theta}$ is calculated as follows. $\widetilde{\alpha}$, $\widetilde{\beta}$ contain $(2(n+p)-n-r)(n+r) = (n+2p-r)(n+r)$ identified coordinates. The additional parameters in $\widetilde{\rho}_1$ are $(n+r)m_1$, and $\widetilde{\mu}_2$ has $(n+p)m_2$ entries; $\widetilde{\Upsilon}_1$ contains $(n+p)p$ elements, where we note that one can e.g. choose $\widetilde{\Upsilon}_2 = \widetilde{\Upsilon}_{1\perp}$. Hence

$$d_{\widetilde{\Theta}} = (n+2p-r+m_1)(n+r) + (n+p)(m_2+p). \tag{20}$$

Note that this dimension is (much) bigger than the corresponding dimension d_{Θ_1}, where $d_{\Theta_1} = (2p-r)r + p(p+1)/2 + rm_1 + pm_2$, see Subsection 4.2. In this sense the dimension of Θ_1 is 'inflated' to the one of $\widetilde{\Theta}$ before obtaining a MC reduction by means of invariance.

In the next subsection we consider extensions of the invariance properties of RRR related to the companion form (17).

6.3   Extensions of RRR invariance

In this subsection we give extensions of the invariance properties of RRR; these extensions are used to define a different class of linear transformations to be applied to the state vector in (17) to which the relevant RRR is invariant.

Let Z_{it} be $n_i \times 1$ vectors of generic variables, $i = 0, 1, 2$, and partition Z_{2t} as $Z_{2t} := (Z'_{21,t} : Z'_{22,t})' := (L'_t : K'_t)'$, where $Z_{2j,t}$ is $n_{2j} \times 1$, $j = 1, 2$. Define also a different set of variables vectors $Z^{\dagger}_{it}$ as follows: $Z^{\dagger}_{it} := (Z'_{it} : Z'_{21,t})'$, $i = 0, 1$ and $Z^{\dagger}_{2t} := Z_{22,t}$. Let $s_1 := \min(n_0, n_1)$, $s_2 := s_1 + n_{21}$; denote by $M^*_{ij} := M_{ij.K}$ the moment matrices of the Z_{it} variables corrected for $Z_{2j,t} := K_t$.

The following theorem describes the connection between the RRR involving the set of Z_{it} variables and the one involving the $Z^{\dagger}_{it}$ variables.

Theorem 11 *Let $\widehat{\Lambda} := \{\widehat{\lambda}_i\}_{i=1}^{s_1}$ be the eigenvalues in $RRR(Z_{0t}, Z_{1t}; Z_{2t})$, $\widehat{\Lambda}^{\dagger} := \{\widehat{\lambda}^{\dagger}_i\}_{i=1}^{s_2}$ be the eigenvalues in $RRR(Z^{\dagger}_{0t}, Z^{\dagger}_{1t}; Z^{\dagger}_{2t})$ and $\mathcal{U} := \{1\}_{i=1}^{n_{21}}$ a set of n_{21} elements all equal to 1. If M^*_{LL} has full rank, then $\widehat{\Lambda}^{\dagger} = \widehat{\Lambda} \cup \mathcal{U}$.*

We observe that the set $\mathcal{U}$ of unit eigenvalues is due to the identity of the subset of variables $L_t := Z_{21}$ present both in $Z^{\dagger}_{0t}$ and $Z^{\dagger}_{1t}$. Theorem 11 is used in the rest of the paper setting Z_{it} equal to the values defined in (4) and taking $Z_{21,t} := U_{t-1}$, $Z_{22,t} = D_{2t}$. We observe that M^*_{LL} has full rank provided $T > n$.

This establishes the equivalence, except for the set $\mathcal{U}$ of eigenvalues equal to 1, of $RRR(Z_{0t}, Z_{1t}; Z_{2t})$ and $RRR(Z^{\dagger}_{0t}, Z^{\dagger}_{1t}; Z^{\dagger}_{2t})$. With abuse of language we say that the two RRR have the same eigenvalues.

We next note that Theorem 11 implies that one can resort to a RRR based on the companion form in place of the original $RRR(Z_{0t}, Z_{1t}; Z_{2t})$.

Corollary 12 *Let Z_{it} be defined as in (4) and $T > n$; then $RRR(Z_{0t}, Z_{1t}; Z_{2t})$ has the same eigenvalues as $RRR(\widetilde{Z}_{0t}, \widetilde{Z}_{1t}; D_{2t})$ where $\widetilde{Z}_{0t} := \Delta \widetilde{X}_t$, $\widetilde{Z}_{1t} := (\widetilde{X}'_{t-1} : D'_{1t})'$ and $\widetilde{X}_t$ is defined in (17).*

Corollary 12 shows that one can substitute $\text{RRR}(Z_{0t}, Z_{1t}; Z_{2t})$ with $\text{RRR}(\widetilde{Z}_{0t}, \widetilde{Z}_{1t}; D_{2t})$. Applying Theorem 3 to $\text{RRR}(\widetilde{Z}_{0t}, \widetilde{Z}_{1t}; D_{2t})$ one obtains the following extended version of the invariance results for RRR.

Theorem 13 *Let $T > n$ and consider the eigenvalues $\widetilde{\lambda}_i$ of $\text{RRR}(\widetilde{Z}_{0t}, \widetilde{Z}_{1t}; D_{2t})$ where $\widetilde{Z}_{0t} := \Delta\widetilde{X}_t$, $\widetilde{Z}_{1t} := (\widetilde{X}'_{1t-1} : D'_{1t})'$. The eigenvalues $\widetilde{\lambda}_i$ are invariant with respect to the following joint transformation of the variables Z_{it}, $i = 0, 1$:*

$$\widetilde{Z}_{0t} \mapsto H_0 \widetilde{Z}_{0t} \qquad \widetilde{Z}_{0t} \mapsto H_1 \widetilde{Z}_{1t}, \tag{21}$$

where H_i, $i = 0, 1$ are square invertible matrices.

In particular we consider the transformation $\widetilde{X}_t^{\circ} := H\widetilde{X}_t$ and the associated transformations (21) with $H_0 = H$, $H_1 = \text{diag}(H, I_{m_1})$. We note that the square matrix H is $(n + p) \times (n + p)$, and hence of bigger dimension than the one in Subsection 3.2 considered for a VAR(1).

6.4 MC design

In this section we give the main results concerning VAR(q) processes. We first note that the proof in Theorem 2 can be applied to the $\widetilde{\theta}$ parametrization in order to show that $g(\widetilde{X}_t) := H\widetilde{X}_t$ induces the transformation $\bar{g}(\widetilde{\theta}) := (H\widetilde{\alpha}, H'^{-1}\widetilde{\beta}, \widetilde{\rho}_1, H\widetilde{\mu}_2, H\widetilde{\Upsilon})$ on the parameters, and that this defines a group $\bar{G}$.

Let $\Sigma := \widetilde{\Upsilon}\widetilde{\Upsilon}'$; we next define the following function $\psi(\widetilde{\theta})$ that is invariant with respect to the transformation $\bar{g} \in \bar{G}$ on the parameters:

$$\widetilde{\psi}(\widetilde{\theta}) := (\widetilde{J}, \widetilde{\kappa}, \widetilde{\gamma}, \widetilde{\zeta}, \widetilde{\phi})(\widetilde{\theta}),$$

where:

1. $(\widetilde{J}, \widetilde{R})(\widetilde{\theta})$ is the real Jordan pair from the real Jordan decomposition of $C := I_{n+r} + \left(\widetilde{\beta}'\Sigma\widetilde{\beta}\right)^{-\frac{1}{2}} \widetilde{\beta}'\widetilde{\alpha} \left(\widetilde{\beta}'\Sigma\widetilde{\beta}\right)^{\frac{1}{2}}$, of dimension $n+r$; $\widetilde{J}$ is a $n+r \times n+r$ real Jordan matrix and $\widetilde{R}$ is the nonsingular transformation that satisfy $C = \widetilde{R}\widetilde{J}\widetilde{R}^{-1}$; $\widetilde{J}$ contains (at most) $n + r$ distinct eigenvalues, less than 1 in modulus;

2. $\widetilde{\kappa}(\widetilde{\theta}) := \text{dg}(\widetilde{\kappa}^*)^{-1} \widetilde{\kappa}^*$, where $(\widetilde{Q}, \widetilde{\kappa}^*)(\theta)$, is the QR decomposition of

$$\left(\widetilde{\beta}'_{\perp}\Sigma^{-1}\widetilde{\beta}_{\perp}\right)^{-1/2} \widetilde{\beta}'_{\perp}\Sigma^{-1}\widetilde{\alpha} \left(\widetilde{\beta}'\Sigma\widetilde{\beta}\right)^{\frac{1}{2}}.$$

   $\widetilde{\kappa}$ is a $(p - r) \times p$ upper triangular matrix, with ones on the main diagonal. It contains $p(p - r) - (p - r)(p - r + 1)/2$ elements in the upper half;

3. $\widetilde{\gamma}(\widetilde{\theta}) := \widetilde{a}^{-1/2}\widetilde{R}^{-1} \left(\widetilde{\beta}'\Sigma\widetilde{\beta}\right)^{-\frac{1}{2}} \widetilde{\rho}'_1$, where $\widetilde{a}$ is the first element on the main diagonal of $\text{diag}\left(\widetilde{R}^{-1}, I\right) UU' \text{diag}\left(\widetilde{R}^{-1\prime}, I\right)$; $\widetilde{\gamma}$ is a $n + r \times m_1$ matrix, with $(n + r)m_1$ elements;

4. $\widetilde{\zeta}(\widetilde{\theta}) := H\widetilde{\mu}_2$, where $H = H_2H_1$ and

$$H_1 := \left(\widetilde{\beta} \left(\widetilde{\beta}'\Sigma\widetilde{\beta}\right)^{-1/2} : \Sigma^{-1}\widetilde{\beta}_{\perp} \left(\widetilde{\beta}'_{\perp}\Sigma^{-1}\widetilde{\beta}_{\perp}\right)^{-1/2}\right)',$$

$$H_2 := \text{diag} \left(\widetilde{a}^{-1/2}\widetilde{R}^{-1} : \text{dg}(\widetilde{\kappa}^*)^{-1}\widetilde{Q}'\right)$$

   $\widetilde{\zeta}$ is $(n + p) \times m_2$.

20

	$r=0$	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
$p-r$				$d_{\widetilde{\Theta}} - d_{\widetilde{F}_1}$								$d_{\Theta_1} - d_{\widetilde{F}_1}$				
1	3	15	44	90	153	233	330	444	0	-4	-6	-6	-4	0	6	14
2	8	30	69	125	198	288	395	519	-5	-8	-9	-8	-5	0	7	16
3	21	53	102	168	251	351	468	602	-9	-11	-11	-9	-5	1	9	19
4	38	80	139	215	308	418	545	689	-16	-17	-16	-13	-8	-1	8	19
5	63	115	184	270	373	493	630	784	-22	-22	-20	-16	-10	-2	8	20
6	92	154	233	329	442	572	719	883	-31	-30	-27	-22	-15	-6	5	18
7	129	201	290	396	519	659	816	990	-39	-37	-33	-27	-19	-9	3	17

Table 2: Left panel: $d_{\widetilde{\Theta}} - d_{\widetilde{F}_1}$. Right panel: $d_{\Theta_1} - d_{\widetilde{F}_1}$. Case $q = 2$, no deterministics.

5. $\widetilde{\phi}'(\theta)$ is the upper triangular $p \times (n+p)$ matrix obtained from the QR decomposition of $\widetilde{\Upsilon}'_1 \widetilde{\beta} \left(\widetilde{\beta}' \Sigma \widetilde{\beta} \right)^{-1/2} \widetilde{R}'^{-1} \widetilde{a}^{-1/2}$; ϕ contains $np+p(p+1)/2-1$ free elements, where the last one is restricted by the first element on the main diagonal of $\widetilde{\phi}\widetilde{\phi}'$ to unity;

Let $\widetilde{F}$ indicate the parameter space of $\widetilde{\psi}$ when $\widetilde{\theta}$ varies in $\widetilde{\Theta}$, and $\widetilde{F}_1$ the one when $\widetilde{\theta}$ varies in $\widetilde{\Theta}_1$. We next define the map $\widetilde{\varphi} : \widetilde{F} \longmapsto \widetilde{\Theta}$, $\widetilde{\varphi}(\widetilde{\psi}) := (\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\rho}_1, \widetilde{\mu}_2, \widetilde{\Upsilon})(\widetilde{\psi})$, that maps $\widetilde{\psi}$ back into $\widetilde{\theta}$ as follows:

$$\widetilde{\alpha}(\widetilde{\psi}) = \left(\begin{array}{c} \widetilde{J} - I \\ \widetilde{\kappa} \end{array} \right), \qquad \widetilde{\beta}(\widetilde{\psi}) = \left(\begin{array}{c} I_{n+r} \\ 0_{(p-r)\times(n+r)} \end{array} \right),$$

$$\widetilde{\rho}_1(\widetilde{\psi}) := \widetilde{\gamma}, \qquad \widetilde{\mu}_2(\widetilde{\psi}) = \widetilde{\zeta}, \qquad \widetilde{\Upsilon}(\widetilde{\psi}) = \left(\widetilde{\phi} : \widetilde{\phi}_\perp \right). \tag{22}$$

We also use the notation $\widetilde{f}(\widetilde{\theta}) := \widetilde{\varphi} \circ \widetilde{\psi} \left(\widetilde{\theta} \right)$. We here state the same results for Theorem 2 for the $\widetilde{\theta}$ parametrization.

Theorem 14 *The function $\widetilde{\psi}(\widetilde{\theta})$ is invariant with respect to the action of $\bar{g} \in \bar{G}$. The point $\widetilde{f}(\widetilde{\theta}) := \widetilde{\varphi}(\widetilde{\psi}(\widetilde{\theta}))$ is on the same orbit as $\widetilde{\theta}$, i.e. $\widetilde{f}(\widetilde{\theta}) = \bar{g}\left(\widetilde{\theta} \right)$ for some $\bar{g} \in \bar{G}$. Moreover $\widetilde{\psi}$ is maximally invariant, i.e. $\widetilde{\psi}(\widetilde{\theta}^\circ) = \widetilde{\psi}(\widetilde{\theta})$ for $\widetilde{\theta}^\circ, \widetilde{\theta} \in \widetilde{\Theta}$ implies $\widetilde{\theta}^\circ = \bar{g}(\widetilde{\theta})$ for some $\bar{g} \in \bar{G}$.*

We note that Theorem 14 can be used to partition $\widetilde{\Theta}$ by the maximal invariant function $\widetilde{\psi}(\widetilde{\theta})$. However, we are interested in the subset $\widetilde{\Theta}_1$ of $\widetilde{\Theta}$ that contains the VAR(q) processes. QUIQUI

We observe that for $\widetilde{\theta} \in \widetilde{\Theta}_1$ the maximal invariant function $\widetilde{\psi}(\widetilde{\theta})$ has a special structure. In particular we make the following remarks.

1. $\widetilde{J}$ derives from the real Jordan decomposition of

$$C := I_{n+r} + \left(\widetilde{\beta}' \Sigma \widetilde{\beta} \right)^{-\frac{1}{2}} \widetilde{\beta}' \widetilde{\alpha} \left(\widetilde{\beta}' \Sigma \widetilde{\beta} \right)^{\frac{1}{2}} =$$

$$= \left(\begin{array}{cccccc} \Omega^{-1/2} \Psi_1 \Omega^{1/2} & \dots & & & \Omega^{-1/2} \Psi_{q-1} & \Omega^{-1/2} \alpha (\beta'\beta)^{1/2} \\ \Omega^{1/2} & 0 & & & & \\ & I & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & I & 0 & \\ & & & & (\beta'\beta)^{-1/2} \beta' & I_r \end{array} \right)$$

whose eigenvalues are less than 1 in modulus by assumption (a2) (c2) or (a3) (c3). One needs $2\lfloor (n+r)/2 \rfloor + 2(n+r) - 1$ coordinates to represent $\widetilde{J}$.

2. From the companion form of a VAR(q), $\widetilde{\Upsilon} = \mathrm{diag}\left(\Omega^{1/2}, I_n\right)$, and $\widetilde{\beta}'_\perp = (0 : \beta'_\perp : 0)$, so that $\widetilde{\beta}'_\perp \Sigma^{-1} = \widetilde{\beta}'_\perp$ and $\left(\widetilde{\beta}'_\perp \Sigma^{-1} \widetilde{\beta}_\perp\right)^{-1/2} \widetilde{\beta}'_\perp \Sigma^{-1} \widetilde{\alpha} \left(\widetilde{\beta}' \Sigma \widetilde{\beta}\right)^{\frac{1}{2}} = (0 : (\beta'_\perp \beta_\perp)^{-1/2} \beta'_\perp : 0)$. In this case, one has $\widetilde{\kappa} = (0 : \breve{\kappa} : 0)$, where $\breve{\kappa} := \mathrm{dg}(\kappa^*)^{-1} \kappa^*$ and κ^* is the upper triangular matrix in the QR decomposition of $(\beta'_\perp \beta_\perp)^{-1/2} \beta'_\perp$. Note that $\breve{\kappa}$ requires only $(p-r)(p+r-1)/2$ coordinates.

3. The coefficients to the deterministic components $\widetilde{\gamma}$ and $\widetilde{\zeta}$ require $(n+r)m_1$ coordinates and $(n+p)m_2$ coordinates respectively. This number is much bigger than the original dimensions of ρ_1 and μ_2, respectively of dimension $r \times m_1$ and $p \times m_2$.

4. Concerning $\widetilde{\phi}$, we observe that $\widetilde{\Upsilon}'_1 \widetilde{\beta} \left(\widetilde{\beta}' \Sigma \widetilde{\beta}\right)^{-1/2} = (I_p : 0)$; however $\widetilde{\phi}'$ is the triangular matrix in the QR decomposition of $(I_p : 0)\widetilde{R}^{-1}\widetilde{a}^{-1/2}$, which depends via $\widetilde{R}^{-1}$ on C. Thus there is no obvious reduction in the number of coordinates needed to represent $\widetilde{\phi}$ with respect to the case $\widetilde{\theta} \in \widetilde{\Theta}$.

5. The number of dimensions needed to represent $\widetilde{\psi} \in \widetilde{F}_1$ is thus

$$d_{\widetilde{F}_1} = 2\left\lfloor \frac{n+r}{2} \right\rfloor + 2(n+r) - 2 + \frac{1}{2}(p-r)(p+r-1)+$$
$$+ np + \frac{1}{2}p(p+1) + (n+r)m_1 + (n+p)m_2.$$

This can be compared with $d_{\widetilde{\Theta}}$ from (20) as well as with d_{Θ_1} from (2). Table 2 reports $d_{\widetilde{\Theta}} - d_{\widetilde{F}_1}$ and $d_{\Theta_1} - d_{\widetilde{F}_1}$ for $q = 2$ and $m_1 = m_2 = 0$. It can be seen that $d_{\widetilde{\Theta}} - d_{\widetilde{F}_1} > 0$ while $d_{\Theta_1} - d_{\widetilde{F}_1}$ can also be negative, i.e. there are more coordinates in $\widetilde{F}_1$ than in Θ_1.

6. Further restrictions can be achieved exploiting the structure of $\widetilde{J}$, using the same principles as in Subsection 4.3. Consider in particular a scalar Jordan block $J_{n_i}(\lambda_i) = \lambda_1 I_{n_i}$ and let ϕ_i be the corresponding block of ϕ, with rank decomposition $\xi \eta'$. One can then show that ϕ_i can be substituted by $(\kappa'_i : 0)'$, where κ_i is the upper triangular matrix in the QR decomposition of $(\xi'\xi)^{1/2}\eta'$, see Subsection 4.3.

7. Even regardless of the (possible) MC design dimension reduction, the $\widetilde{\psi}$ parametrization is better suited to control the stationary roots of the VAR(q) system through $\widetilde{J}$; this is not possible using directly the parametrization in terms of $\Psi_1, \dots \Psi_{q-1}, \alpha, \beta, \Omega$.

7 Conclusions

In this paper we have considered reductions in design dimensions that can be achieved thanks to invariance properties of reduced rank regression and the property of I(1) VAR processes to be closed under linear transformations. These MC design reductions can be considerable. The results are applied to the indexing of the limit distribution of LR CI rank test under local alternatives and to VAR of order higher than 1.

References

Anderson T. W. (1984) *An introduction to multivariate statistical analysis*, second edition. John Wiley, New York.

Bauer D. and M. Wagner (2005) A Canonical Form for Unit Root Processes in the State Space Framework, Mimeo

Cavaliere G., L. Fanelli and P. Paruolo (2005) On cointegration-rank determination under rank constraints, Quaderno di Facoltà 2001/17 Università dell'Insubria, Varese Italy; revised July 2005.

Hansen P. R. (2005). Granger's Representation Theorem: A Closed-Form Expression for I(1) Processes. *Econometrics Journal* 8, 23-38.

Horn R. A. and C. R. Johnson (1985) *Matrix analysis*, Cambridge, Cambridge University Press.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors, *Journal of Economics Dynamics and Control*, 12, 231-254.

Johansen, S. (1996). *Likelihood based inference in cointegrated vector autoregressive models*, Oxford: Oxford University Press.

Johansen S. (2002) A small sample correction for the test of cointegration rank in the vector autoregressive model, *Econometrica* 70, 1929-1961.

Gonzalo J. (1994) Five alternative methods of estimating long-run equilibrium relationships, *Journal of Econometrics* 60, 203-233.

Lehmann E. L. (1986) Testing statistical hypotheses. Second edition. Springer, New York.

Lütkepohl H., P. Saikkonen and C. Trenkler (2001) Maximum eigenvalue versus trace tests for the cointegrating rank of a VAR process, *Econometrics Journal* 4, 287–310.

Marzocchi W., Mulargia F., Paruolo P. (1992) The correlation of geomagnetic reversals and mean sea level in the last 150 m. y., *Earth Planetary Science Letters* 111, p. 383-393.

Nielsen B. (2004) On the distribution of likelihood ratio test statistics for cointegration rank, *Econometric Reviews* 23, 1–23.

Saikkonen, P. and H. Lütkepohl (2000). Testing for the cointegrating rank of a VAR process with an intercept. *Econometric Theory* 16, 373–406.

Toda, H. Y. (1994). Finite sample properties of likelihood ratio tests for cointegrating ranks when linear trends are present. *Review of Economics and Statistics* 76, 66–79.

Toda, H. Y. (1995). Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions. *Econometric Theory* 11, 1015–32.

8 Appendix

Proof. of Theorem 2. Pre-multiplying $A(L)X_t = \mu D_t + \varepsilon_t$ by H one obtains $HA(L)H^{-1}HX_t = H\mu D_t + H\varepsilon_t$ which defines a VAR for $X_t^\circ := HX_t$ with AR polynomial $A^\circ(L) := HA(L)H^{-1}$, deterministic coefficients $\mu^\circ := H\mu$, and Gaussian noise $\varepsilon_t^\circ := H\varepsilon_t \sim N\left(0, \Omega^\circ\right)$, $\Omega^\circ := H\Omega H'$.

We need to verify that the measure P_{θ° of X_t° is in $\mathcal{P}_1$. First we observe that (a) is verified because the $\left|A^\circ(z)\right| = |H|\,|A(z)|\,|H^{-1}| = |A(z)|$ and hence the AR characteristic polynomials $\left|A^\circ(z)\right|$ and $|A(z)|$ have the same zeroes. (b) is also satisfied; in fact $\mathrm{rk}\left(A^\circ(1)\right) = \mathrm{rk}\left(HA(1)H^{-1}\right) = \mathrm{rk}\left(A(1)\right) = r$ because H is of full rank. Hence one can write $A^\circ(1) = -\alpha^\circ \beta^{\circ\prime}$ with $\alpha^\circ = H\alpha$ and $\beta^\circ = H'^{-1}\beta$. Note that $\alpha_\perp^\circ = H'^{-1}\alpha_\perp$, $\beta_\perp^\circ = H\beta_\perp$.

Next consider (c); one finds $\dot{A}^\circ(1) = H\dot{A}(1)H^{-1}$, so that

$$\alpha_\perp^{\circ\prime}\dot{A}^\circ(1)\beta_\perp^\circ = \alpha_\perp' H^{-1} H\dot{A}(1)H^{-1}H\beta_\perp = \alpha_\perp'\dot{A}(1)\beta_\perp$$

and obviously $\alpha_\perp^{\circ\prime}\dot{A}^\circ(1)\beta_\perp^\circ$ is of full rank $p-r$ iff $\alpha_\perp'\dot{A}(1)\beta_\perp$ is; this shows that (c) is satisfied.

Finally consider ρ_1°, μ_2° and Ω°; one has $\rho_1' = \rho_1^{\circ\prime}$. One also sees that $\Omega^\circ := H\Omega H'$ is symmetric positive definite because H and Ω are full rank. This completes the proof. ∎

Proof. of Theorem 3. First we show that the definition of S_{ij} is unaffected by multiplication of Z_{2t} by H_2, square and nonsingular. In fact

$$S_{ij} := M_{ij} - M_{i2}M_{22}^{-1}M_{2j} = M_{ij} - M_{i2}H_2'\left(H_2 M_{22}H_2'\right)^{-1}H_2 M_{2j}.$$

Next observe that S_{ij} is unaffected by substituting Z_{0t} with $Z_{0t}^\circ := Z_{0t} + H_{02}Z_{2t}$ and Z_{1t} with $Z_{1t}^\circ := Z_{1t} + H_{12}Z_{2t}$. In fact

$$\begin{aligned}
S_{ij}^\circ := M_{ij}^\circ - M_{i2}^\circ M_{22}^{-1}M_{2j}^\circ &= M_{ij} + H_{i2}M_{2j} + M_{i2}H_{j2}' + H_{i2}M_{22}H_{j2}' \\
&\quad - \left(M_{i2} + H_{i2}M_{22}\right)M_{22}^{-1}\left(M_{j2} + H_{j2}M_{22}\right)' \\
&= M_{ij} - M_{i2}M_{22}^{-1}M_{2j} = S_{ij}.
\end{aligned}$$

Finally the eigenvalues $\widehat{\Lambda}$ are invariant with respect to substitution of Z_{it} with $H_i Z_{it}$, $i = 0, 1$; in fact because H_i are invertible, one has

$$0 = \left|\widehat{\lambda}S_{11} - S_{10}S_{00}^{-1}S_{01}\right| = |H_1|\left|\widehat{\lambda}S_{11} - S_{10}S_{00}^{-1}S_{01}\right||H_1'| =$$

$$= \left|\widehat{\lambda}H_1 S_{11}H_1' - H_1'S_{10}S_{00}^{-1}S_{01}H_1'\right| = \left|\widehat{\lambda}H_1 S_{11}H_1' - H_1 S_{10}H_0'\left(H_0 S_{00}H_0'\right)^{-1}H_0 S_{01}H_1'\right|.$$

This proves the first statement of the theorem.

The second part is proved as follows. Let V indicate the generalized eigenvectors associated with $\mathrm{RRR}(Z_{0t}, Z_{1t}; Z_{2t})$ that satisfy $V'S_{11}V = I_p$, $V'S_{10}S_{00}^{-1}S_{01}V = \widehat{\Lambda}$. If a function $f(S_{00}, S_{10}, S_{11})$ is invariant with respect to (7), then one can choose $H_1 = V'$, $H_0' = S_{00}^{-1}S_{01}V$, $H_2 = I$ to find, by the assumed invariance of f, that

$$f(S_{00}, S_{10}, S_{11}) = f(H_0 S_{00}H_0', H_1 S_{10}H_0', H_1 S_{11}H_1') = f(\widehat{\Lambda}, \widehat{\Lambda}, I_p),$$

which is a function of $\widehat{\Lambda}$. This completes the proof. ∎

For ease of exposition, before the proof of Theorem 4 we first prove the following Lemma.

Lemma 15 $f(\theta) := \varphi \circ \psi(\theta)$ *lies on the same orbit as* θ, *i.e.* $f(\theta) := \varphi(\psi(\theta)) = \bar{g}(\theta)$ *for some* $\bar{g} \in \bar{G}$.

Proof. In order to show that $f(\theta) := \varphi(\psi(\theta)) = \bar{g}(\theta)$, we show how to construct a transformation $g(X)$ that has $\bar{g}(\theta) = f(\theta)$. Let $g(X)$ be represented by HX_t, with H square and nonsingular. We build H in (multiplicative) stages, i.e. we consider $H := H_2 H_1$. Define

$$H_1 := \left(\beta \left(\beta' \Omega \beta \right)^{-1/2} : \Omega^{-1} \beta_\perp \left(\beta'_\perp \Omega^{-1} \beta_\perp \right)^{-1/2} \right)'$$

as in (9) and $X_t^{(1)} := H_1 X_t$. $X_t^{(1)}$ follows a VAR(1) $\Delta X_t^{(1)} = \alpha^{(1)} \beta^{*(1)\prime} (X_{t-1}^{(1)\prime} : D_{1t}')' + \mu_2^{(1)} D_{2t} + \varepsilon_t^{(1)}$ with $\alpha^{(1)} = H_1 \alpha$, $\beta^{(1)} := \left(I_r : 0_{r \times (p-r)} : \rho_1' \right)'$, $\mu_2^{(1)} := H_1 \mu_2$, $\varepsilon_t^{(1)} = H_1 \varepsilon_t \sim N(0, I_p)$. Here we have used twice the fact that by construction $H_1^{-1} = \Omega H_1'$.

Next partition $\alpha^{(1)}$ into the first r and the last $p - r$ rows,

$$\alpha_1^{(1)} := \left(\beta' \Omega \beta \right)^{-1/2} \beta' \alpha \left(\beta' \Omega \beta \right)^{1/2},$$
$$\alpha_2^{(1)} := \left(\beta'_\perp \Omega^{-1} \beta_\perp \right)^{-1/2} \beta'_\perp \Omega^{-1} \alpha \left(\beta' \Omega \beta \right)^{1/2}.$$

Note that $\alpha_1^{(1)}$ is square and of full rank because of condition (c). $\alpha_1^{(1)}$ is associated with the stationary dynamics because $\beta^{*(1)\prime} X_t^{(1)}$ selects the first r component in $X_t^{(1)}$, and these are stationary by GRT. $\alpha_2^{(1)}$ are the loadings of the remaining variables on $\beta^{*(1)\prime} X_t^{(1)}$. We here consider a joint transformation of $\alpha_1^{(1)}, \alpha_2^{(1)}$.

For $\alpha_1^{(1)}$ we consider the real Jordan decomposition of $\alpha_1^{(1)} + I_r$, see e.g. Horn and Johnson (1985) p. 126, Theorem 3.1.11:

$$\alpha_1^{(1)} + I_r = RJR^{-1},$$

where J is the Jordan matrix described in the text and R is a real square invertible matrix. Note that the above decomposition can be written $\alpha_2^{(1)} = R(J - I)R^{-1}$.

For $\alpha_2^{(1)}$ note that by assumption $j := \text{rk}(\beta'_\perp \Omega^{-1} \alpha) = \text{rk}\left(\alpha_2^{(1)} \right)$, so that one can rank-decompose $\alpha_2^{(1)}$ into $\alpha_2^{(1)} = \xi \eta'$, say, for ξ and η full column rank $p - r \times j$ matrices. We consider the QR decomposition of $(\xi' \xi)^{1/2} \eta'$,

$$(\xi' \xi)^{1/2} \eta' = Q \kappa$$

where Q is $j \times j$ and orthogonal and κ is $j \times r$, upper triangular with positive diagonal elements; this choice is unique see e.g. Horn and Johnson (1985) Theorem 2.6.1.

Denote by a the first element on the main diagonal of $R^{-1} R^{-1\prime}$; this element is positive because the first row in R^{-1} is nonzero due to the non-singularity of R^{-1}. We next consider the second transformation

$$H_2 := \text{diag}\left(a^{-1/2} R^{-1}, H_3 \right), \qquad H_3 := \left(\xi (\xi' \xi)^{-1/2} Q : \xi_\perp (\xi'_\perp \xi_\perp)^{-1/2} \right)$$

see (10) and $W_t := X_t^{(2)} := H_2 X_t^{(1)}$; note that H_3 is orthogonal and $H_2^{-1} = \text{diag}\left(R a^{1/2}, H_3 \right)$. $X_t^{(2)}$ follows a VAR(1) $\Delta X_t^{(2)} = \alpha^{(2)} \beta^{*(2)\prime} (X_{t-1}^{(2)\prime} : D_{1t}')' + \mu_2^{(2)} D_{2t} + \varepsilon_t^{(2)}$ with

$$\alpha^{(2)} = ((J - I)' : \kappa' : 0)'$$

$\beta^{*(2)\prime} = (I_r : 0 : a^{-1/2} R^{-1} \rho_1')$, $\mu_2^{(1)} := H \mu_2$, $\varepsilon_t^{(2)} = H_2 \varepsilon_t^{(1)} \sim N(0, H_2 H_2')$, $H_2 H_2' = \text{diag}(a^{-1} R^{-1} R^{-1\prime}, I_{p-r}) =: \text{diag}(\Phi, I_{p-r})$, $\gamma := a^{-1/2} R^{-1} \rho_1'$. In the above when $j = r$

25

we assume that $\xi = I_r$ and $\xi_\perp$ is dropped from the transformation. We have hence show that $\bar{g}(\theta) = f(\theta)$; this completes the proof. ∎

Proof. of Theorem 4. We first verify that $\psi(\bar{g}(\theta)) = \psi(\theta)$, i.e. ψ is $\bar{g} \in \bar{G}$ invariant. Recall that when $g(X)$ has matrix representation HX_t, then $\bar{g}$ maps θ in $\theta^\circ := \bar{g}(\theta) = (H\alpha, H'^{-1}\beta, \rho_1, H^{-1}\mu_2, H\Omega H')$. Note that $\beta_\perp^\circ = H\beta_\perp$.

We next calculate $\psi(\bar{g}(\theta))$; (J°, R°) are the real Jordan pair of

$$\left(\beta^{\circ\prime}\Omega^\circ\beta^\circ\right)^{-1/2}\beta^{\circ\prime}\alpha^\circ\left(\beta^{\circ\prime}\Omega^\circ\beta^\circ\right)^{1/2} + I$$

where

$$\left(\beta^{\circ\prime}\Omega^\circ\beta^\circ\right)^{-1/2}\beta^{\circ\prime}\alpha^\circ\left(\beta^{\circ\prime}\Omega^\circ\beta^\circ\right)^{1/2} = \left(\beta'H^{-1}H\Omega HH'^{-1}\beta\right)^{-1/2}\beta'H^{-1}H\alpha\left(\beta'H^{-1}H\Omega HH'^{-1}\beta\right)^{1/2}$$
$$= (\beta'\Omega\beta)^{-1/2}\beta'\alpha(\beta'\Omega\beta)^{1/2},$$

and hence $\left(J^\circ, R^\circ\right) = (J, R)$. $\left(Q^\circ, \kappa^\circ\right)$ are functions of $\left(\beta_\perp^{\circ\prime}\Omega^{\circ-1}\beta_\perp^\circ\right)^{-1/2}\beta_\perp^{\circ\prime}\Omega^{\circ-1}\alpha^\circ$, which is invariant, because it satisfies

$$\left(\beta_\perp^{\circ\prime}\Omega^{\circ-1}\beta_\perp^\circ\right)^{-1/2}\beta_\perp^{\circ\prime}\Omega^{\circ-1}\alpha^\circ\left(\beta^{\circ\prime}\Omega^\circ\beta^\circ\right)^{1/2} = \left(\beta_\perp'\Omega^{-1}\beta_\perp\right)^{-1/2}\beta_\perp'\Omega^{-1}\alpha(\beta'\Omega\beta)^{1/2}.$$

Hence $\left(Q^\circ, \kappa^\circ\right) = (Q, \kappa)$. Note that a° and hence $\Phi^\circ := a^\circ R^{\circ-1}R^{\circ-1\prime}$ are also invariant, being a function of R°, which is invariant. Moreover γ° is a function of R° and $\beta^{\circ\prime}\Omega^\circ\beta^\circ$, which are invariant. Finally $\zeta^\circ := H_2^\circ H_1^\circ \mu_2^\circ$; by (10) we note that H_2° is invariant, $H_2^\circ = H_2$, being a function of $\left(J^\circ, R^\circ\right)$ and $\left(Q^\circ, \kappa^\circ\right)$. From (9) one finds

$$H_1^\circ := \left(\beta(\beta'\Omega\beta)^{-1/2} : \Omega^{-1}\beta_\perp\left(\beta_\perp'\Omega^{-1}\beta_\perp\right)^{-1/2}\right)' H^{-1}$$

and hence $\zeta^\circ = H_2 H_1 H^{-1}H\mu_2 = H\mu_2 =: \zeta$, i.e. ζ° is also invariant. This shows that ψ is invariant.

In order to show that ψ is maximally invariant, we need to show that $\psi(\theta^\circ) = \psi(\theta)$ implies $\theta^\circ = \bar{g}(\theta)$ for some $\bar{g} \in \bar{G}$. Now $\psi(\theta^\circ) = \psi(\theta)$ implies $f(\theta^\circ) := \varphi(\psi(\theta^\circ)) = \varphi(\psi(\theta)) =: f(\theta)$, which are hence identical and on the same orbit as θ° and θ by Lemma 15; hence there exists a $\bar{g} \in \bar{G}$ such that $\theta^\circ = \bar{g}(\theta)$. This completes the proof. ∎

Proof. of Corollary 5. Simply note that W_t in (12) represents $\mathbb{P}_{\varphi(\psi)}$. ∎

Proof. of Corollary 6. Note that $p^2/2 = (p-r)^2/2 + r^2/2 + (p-r)r$. One hence has

$$\begin{aligned} d_{\Theta_1} - d_F &= (2p-r)r + p(p+1)/2 - (2\lfloor r/2\rfloor + 2r - 1 + r(r+1) - 1) \\ &= (3p-2r)r + (p-r)^2/2 + r^2/2 + p/2 - \left(2\lfloor r/2\rfloor + 3r - 2 + r^2\right) \\ &= 3(p-r-1)r + (p-r)^2/2 + \left(r^2/2 - \lfloor r/2\rfloor\right) + (p/2 - \lfloor r/2\rfloor) + 2. \end{aligned}$$

the last expression shows that $d_{\Theta_1} - d_F > 0$, because it is the sum of non-negative terms, where some of these are strictly positive. ∎

Proof. of Theorem 7. We first prove that $f_2(\theta)$ is on the same orbit of θ for $\theta \in \Theta_2$. In order to do so, we construct a suitable g transformation on W_t given in (12) given by $Y_t := HW_t$ with $H = \text{diag}(H_3, H_4, I_{p-j-r})$, $H_3 := \text{diag}(s_1\phi_{11}^{-1/2}, ..., s_r\phi_{rr}^{-1/2})$, $H_4 := \text{diag}(s_1, ..., s_j)$ Note that H_3 commutes with J, $H_3 J = J H_3$, because both are diagonal matrices. Hence $Y_{1t} := H_3 W_{1t} = J(H_3 W_{1t-1} + H_3\gamma D_{1t}) + H_3\zeta_1 D_{2t} + H_3\upsilon_{1t}$. The second block of variables reads $Y_{2t} := H_4 W_{2t} = H_4\kappa H_3^{-1}(H_3 W_{1t-1} + H_3\gamma D_{1t}) + \upsilon_{2t}$. The third block of variables in unaffected. Note that $H_4\kappa H_3^{-1}$ is again an upper-triangular matrix with all positive coefficients on the main diagonal. Moreover $\gamma^* = H_3\gamma$, $\zeta^* = H\zeta$.

Y_t hence follows a VAR(1) process with the same J as W_t and a κ matrix with the same structure as W_t; the covariance of the errors $\mathrm{diag}(\Phi, I_{p-r})$ is now replaced by $\mathrm{diag}(H_3 \Phi H_3,\ H_4 H_4,\ I_{p-r-j})$, because the H_i matrices are diagonal. Note that $H_4 H_4 = I_j$ and that

$$
H_3 \Phi H_3 = \begin{pmatrix}
1 & s_1 \phi_{12} s_2 & \ldots & & s_1 \phi_{1r} s_r \\
s_2 \phi_{21} s_1 & & & & \\
\vdots & & & & s_{r-1} \phi_{r-1,r} s_r \\
s_r \phi_{r1} s_1 & \ldots & s_r \phi_{r,r-1} s_{r-1} & & 1
\end{pmatrix}
$$

where $s_1 \phi_{1l} s_l = \phi_{1l} s_l \geq 0$ because $s_1 := 1$. Hence $H_3 \Phi H_3$ is a correlation matrix (symmetric and positive definite) with non-negative correlations in the first row and column. This shows that there is a $\bar{g}_2 \in \bar{G}$ such that $f_2(\theta) = \bar{g}_2(f(\theta))$. By Lemma 15, $f(\theta)$ is on the same orbit as θ, i.e. $f(\theta) = \bar{g}(\theta)$, so that $f_2(\theta) = \bar{g}_2 \circ \bar{g}(\theta) =: \bar{g}_3(\theta)$, i.e. $f_2(\theta)$ and θ are on the same orbit.

In order to show that ψ_2 is maximally invariant on Θ_2, we need to show that $\psi_2(\theta^\circ) = \psi_2(\theta)$ implies $\theta^\circ = \bar{g}(\theta)$ for some $\bar{g} \in \bar{G}$, $\theta, \theta^\circ \in \Theta_2$. Now $\psi_2(\theta^\circ) = \psi_2(\theta)$ implies $f_2(\theta^\circ) := \varphi(\psi_2(\theta^\circ)) = \varphi(\psi_2(\theta)) =: f_2(\theta)$, which are hence identical and on the same orbit as θ° and θ by the discussion above; hence there exists a $\bar{g} \in \bar{G}$ such that $\theta^\circ = \bar{g}(\theta)$. This completes the proof. ∎

We next give a lemma that is later used the proof of Theorem 8.

Lemma 16 *Let Φ be a $r \times r$ positive definite real symmetric matrix; then there exists a real upper triangular $r \times r$ matrix V with positive elements on the main diagonal that satisfies $V \Phi V' = I_r$.*

Proof. Define D as a square matrix of order r with all zero elements except for the ones on the minor diagonal, all equal to 1. Note that $DD = I$, and that for any conformable A, DAD has the same elements of A arranged in reverse order, both column- and row-wise. Let $Q := D \Phi D = PP'$ where P is the Choleski factor, a lower triangular matrix with positive diagonal elements. From the Choleski decomposition of Q one has $P^{-1} Q P^{-1\prime} = I$, where P^{-1} is lower triangular with positive diagonal elements. Pre- and post- multiplying by D one finds $DP^{-1}QP^{-1\prime}D = DD = I$. Substituting the definition Q one concludes $V \Phi V' = I$, where $V := DP^{-1}D$ is upper triangular, with positive elements on the diagonal. This completes the proof. ∎

Proof. of Theorem 8. The proof is the same as Theorem 7, substituting H_3 with V as given in Lemma 16 and H_4 with the identity I_j. This completes the proof. ∎

Proof. of Theorem 9. The same proof of Theorem 2 goes through. ∎

Proof. of Corollary 10. This is immediate from Theorem 2. ∎

Proof. of Theorem 11. Let $M_{ij}^\dagger$, $S_{ij}^\dagger$ indicate M_{ij}, S_{ij} in Theorem 3 for $\mathrm{RRR}(Z_{0t}^\dagger, Z_{1t}^\dagger; Z_{2t}^\dagger)$. For notational convenience we let $L_t := Z_{21,t}$, $K_t := Z_{22,t}$, and let $M_{ij}^* := M_{ij.K}$. We find that

$$
S_{11}^\dagger = \begin{pmatrix} M_{11}^* & M_{1L}^* \\ M_{L1}^* & M_{LL}^* \end{pmatrix}, \qquad
S_{10}^\dagger = \begin{pmatrix} M_{10}^* & M_{1L}^* \\ M_{L0}^* & M_{LL}^* \end{pmatrix}, \qquad
S_{00}^\dagger = \begin{pmatrix} M_{00}^* & M_{0L}^* \\ M_{L0}^* & M_{LL}^* \end{pmatrix},
$$

$$
S_{10}^\dagger S_{00}^{\dagger -1} S_{10}^\dagger = \begin{pmatrix} M_{10.L}^* M_{00.L}^{*-1} M_{01.L}^* + M_{1L}^* M_{LL}^{*-1} M_{L1}^* & M_{1L}^* \\ M_{L1}^* & M_{LL}^* \end{pmatrix}
$$

$$
= \begin{pmatrix} S_{10} S_{00}^{-1} S_{01} + M_{1L}^* M_{LL}^{*-1} M_{L1}^* & M_{1L}^* \\ M_{L0}^* & M_{LL}^* \end{pmatrix}.
$$

Where we have used the fact that $M^*_{ij.L} = S_{ij}$ by the Lowell-Frisch-Waugh theorem. The eigenvalue problem $\left| \widehat{\lambda}^\dagger S^\dagger_{11} - S^\dagger_{10} S^{\dagger -1}_{00} S^\dagger_{10} \right| = 0$ thus reads

$$\left| \widehat{\lambda}^\dagger \begin{pmatrix} M^*_{11} & M^*_{1L} \\ M^*_{L1} & M^*_{LL} \end{pmatrix} - \begin{pmatrix} M^*_{1L} M^{*-1}_{LL} M^*_{L1} + S_{10} S^{-1}_{00} S_{01} & M^*_{1L} \\ M^*_{L1} & M^*_{LL} \end{pmatrix} \right| = 0.$$

The rank of M^*_{LL} is assumed full. Using standard properties of determinants, and setting $Q := M^*_{1L} M^{*-1}_{LL} M^*_{L1} + S_{10} S^{-1}_{00} S_{01}$, one finds

$$0 = \left| \widehat{\lambda}^\dagger M^*_{LL} - M^*_{LL} \right| \left| \widehat{\lambda}^\dagger M^*_{11} - Q - \left(\left(\widehat{\lambda}^\dagger - 1 \right) M^*_{1L} \right) \left(\left(\widehat{\lambda}^\dagger - 1 \right) M^*_{LL} \right)^{-1} \left(\left(\widehat{\lambda}^\dagger - 1 \right) M^*_{L1} \right) \right|$$

$$= \left| \widehat{\lambda}^\dagger M^*_{LL} - M^*_{LL} \right| \left| \widehat{\lambda}^\dagger M^*_{11} - Q - \left(\widehat{\lambda}^\dagger - 1 \right) M^*_{1L} M^{*-1}_{LL} M^*_{L1} \right|$$

$$= \left| \widehat{\lambda}^\dagger M^*_{LL} - M^*_{LL} \right| \left| \widehat{\lambda}^\dagger S_{11} - S_{10} S^{-1}_{00} S_{01} \right|. \tag{23}$$

The first factor gives solutions $\widehat{\lambda}^\dagger = 1$, with multiplicity given by n_{21}, the dimension of M^*_{LL}. The last factor in (23) is $\mathrm{RRR}(Z_{0t}, Z_{1t}; Z_{2t})$, see (6), which gives equality of the $\widehat{\lambda}^\dagger_i$ eigenvalues different from 1 with the $\widehat{\lambda}_i$ eigenvalues. $\blacksquare$

Proof. of Corollary 12. Let $Z_{21,t} := U_{t-1}$, $Z^\dagger_{2t} := Z_{22,t} = D_{2t}$, and recall that $\widetilde{X}_t = (U'_t : X'_{t-q+1})'$. Recall that $\mathrm{RRR}(Z_{0t}, Z_{1t}; Z_{2t})$ and $\mathrm{RRR}(Z^\dagger_{0t}, Z^\dagger_{1t}; Z^\dagger_{2t})$ have the same eigenvalues, (except for the unit eigenvalues in modulo $\mathcal{U}$). We wish to show that $\mathrm{RRR}(Z^\dagger_{0t}, Z^\dagger_{1t}; Z^\dagger_{2t})$ and $\mathrm{RRR}(\widetilde{Z}_{0t}, \widetilde{Z}_{1t}; D_{2t})$ have the same eigenvalues by showing that $Z^\dagger_{it}$ is a linear invertible transformation of $\widetilde{Z}_{it}$, $i = 0, 1$.

In fact $Z^\dagger_{0t} := (Z'_{0t} : U'_{t-1})' = (\Delta X'_t : U'_{t-1})'$ and $\widetilde{Z}_{0t} := \Delta \widetilde{X}_t = (\Delta U'_t : \Delta X_{t-q+1})$; similarly $Z^\dagger_{1t} := (Z'_{1t} : U'_{t-1})' = (X'_{t-1} : D'_{1t} : U'_{t-1})'$ and $\widetilde{Z}_{1t} := (\widetilde{X}'_{1t-1} : D'_{1t})' = (U'_{t-1} : X'_{t-q} : D'_{1t})'$. This completes the proof. $\blacksquare$

Proof. of Theorem 13. Combine Theorem 3 and Corollary 12. $\blacksquare$

Proof. of Theorem 14. Recall $\widetilde{\varepsilon}^*_t := \widetilde{\Upsilon}_1 (\mu_2 D_{2t} + \varepsilon_t) + \widetilde{\Upsilon}_2 \cdot 0$. We find that $\widetilde{f}(\widetilde{\theta}) = \bar{g}(\widetilde{\theta})$ by proving that there exist a transformation $\widetilde{X}^\circ_t := H \widetilde{X}_t = g\left(\widetilde{X}_t \right)$ that implies $\widetilde{f}(\widetilde{\theta}) = \bar{g}(\widetilde{\theta})$. g is obtained in two multiplicative steps, $H = H_2 H_1$. Consider first (17) and the transformation $\widetilde{X}^{(1)}_t := H_1 \widetilde{X}_t$ with

$$H_1 := \left(\widetilde{\beta} \left(\widetilde{\beta}' \Sigma \widetilde{\beta} \right)^{-1/2} : \Sigma^{-1} \widetilde{\beta}_\perp \left(\widetilde{\beta}'_\perp \Sigma^{-1} \widetilde{\beta}_\perp \right)^{-1/2} \right)',$$

$H^{-1}_1 = \Sigma H'_1$. One finds

$$\Delta \widetilde{X}^{(1)}_t := H_1 \Delta \widetilde{X}_t = H_1 \widetilde{\alpha} \begin{pmatrix} \widetilde{\beta}' & \widetilde{\rho}'_1 \end{pmatrix} \begin{pmatrix} H^{-1}_1 H_1 & \\ & I \end{pmatrix} \begin{pmatrix} \widetilde{X}_{t-1} \\ D_{1t} \end{pmatrix} + H_1 \widetilde{\varepsilon}^*_t,$$

$$= \widetilde{\alpha}^{(1)} \begin{pmatrix} I_{n+r} & 0 & \left(\widetilde{\beta}' \Sigma \widetilde{\beta} \right)^{-1/2} \widetilde{\rho}'_1 \end{pmatrix} \begin{pmatrix} \widetilde{X}^{(1)}_{t-1} \\ D_{1t} \end{pmatrix} + H_1 \widetilde{\varepsilon}^*_t$$

where

$$\widetilde{\alpha}^{(1)} := \begin{pmatrix} \widetilde{\alpha}^{(1)}_1 \\ \widetilde{\alpha}^{(1)}_2 \end{pmatrix} = \begin{pmatrix} \left(\widetilde{\beta}' \Sigma \widetilde{\beta} \right)^{-1/2} \widetilde{\beta}' \widetilde{\alpha} \\ \left(\widetilde{\beta}'_\perp \Sigma^{-1} \widetilde{\beta}_\perp \right)^{-1/2} \widetilde{\beta}'_\perp \Sigma^{-1} \widetilde{\alpha} \end{pmatrix} \left(\widetilde{\beta}' \Sigma \widetilde{\beta} \right)^{1/2}.$$

Next consider the real Jordan decomposition of $\widetilde{\alpha}^{(1)}_1 + I_{n+r} = \widetilde{R} \widetilde{J} \widetilde{R}^{-1}$ and the QR decomposition of $\widetilde{\alpha}^{(1)}_2 = \widetilde{Q} \widetilde{\kappa}^*$; the second transformation $\widetilde{X}^{(2)}_t := H_2 \widetilde{X}^{(1)}_t$ with $H_2 :=$

$\mathrm{diag}\left(\widetilde{a}^{-1/2}\widetilde{R}^{-1} : \mathrm{dg}(\widetilde{\kappa}^*)^{-1}\widetilde{Q}'\right)$; one finds

$$\Delta\widetilde{X}_t^{(2)} = \begin{pmatrix} \widetilde{J} - I \\ \mathrm{dg}(\widetilde{\kappa}^*)^{-1}\widetilde{Q}'\widetilde{\alpha}_2^{(1)} \end{pmatrix} \begin{pmatrix} I_{n+r} & 0 & \widetilde{a}^{-1/2}\widetilde{R}^{-1}\left(\widetilde{\beta}'\Sigma\widetilde{\beta}\right)^{-1/2}\widetilde{\rho}_1' \end{pmatrix} \begin{pmatrix} \widetilde{X}_{t-1}^{(1)} \\ D_{1t} \end{pmatrix} + H\widetilde{\varepsilon}_t,$$

$$= \begin{pmatrix} \widetilde{J} - I \\ \widetilde{\kappa} \end{pmatrix} \begin{pmatrix} I_{n+r} & 0 & \widetilde{\gamma} \end{pmatrix} \begin{pmatrix} \widetilde{X}_{t-1}^{(1)} \\ D_{1t} \end{pmatrix} + H\widetilde{\varepsilon}_t.$$

This proves that $\widetilde{f}(\theta) = \bar{g}(\widetilde{\theta})$ for $\bar{g} \in \bar{G}$. We next show that $\widetilde{\psi}$ is invariant. This is proved exactly in the same way as in the proof of Theorem 4 by showing that $\widetilde{\alpha}_1^{(1)} := \left(\widetilde{\alpha}_1^{(1)\prime} : \widetilde{\alpha}_2^{(1)\prime}\right)'$ is invariant, and that $\widetilde{J}, \widetilde{R}, \widetilde{Q}, \widetilde{\kappa}, \widetilde{a}$ are functions of $\widetilde{\alpha}_1^{(1)}$. $\widetilde{\gamma}$ is a function of $\widetilde{\alpha}_1^{(1)}$ and $\widetilde{\beta}'\Sigma\widetilde{\beta}$ which is itself invariant. $\widetilde{\zeta}$ is proved to be invariant just as in Theorem 4. Finally $\widetilde{\phi}$ is a function of $\widetilde{\Upsilon}_1'\widetilde{\beta}, \widetilde{\beta}'\Sigma\widetilde{\beta}, \widetilde{R}, \widetilde{a}$, which are all invariant. From $\widetilde{f}(\widetilde{\theta}) = \bar{g}(\widetilde{\theta})$ for $\bar{g} \in \bar{G}$ one also concludes that $\widetilde{\psi}$ is maximal invariant. $\blacksquare$