

Wegge, Leon

**Working Paper**

## ARMAX(p,r,q) parameter identifiability without coprimeness

Working Paper, No. 12-17

**Provided in Cooperation with:**

University of California Davis, Department of Economics

*Suggested Citation:* Wegge, Leon (2012) : ARMAX(p,r,q) parameter identifiability without coprimeness, Working Paper, No. 12-17, University of California, Department of Economics, Davis, CA

This Version is available at:

<https://hdl.handle.net/10419/79674>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# UC DAVIS

---

## DEPARTMENT OF ECONOMICS Working Paper Series

ARMAX(p,r,q) Parameter Identifiability Without Coprimeness

Leon Wegge  
U.C. Davis

July 25, 2012

Paper # 12-17

Department of Economics  
One Shields Avenue  
Davis, CA 95616  
(530)752-0741

[http://www.econ.ucdavis.edu/working\\_search.cfm](http://www.econ.ucdavis.edu/working_search.cfm)

## ARMAX(P,R,Q) PARAMETER IDENTIFIABILITY WITHOUT COPRIMENESS

BY LEON L. WEGGE\* ,

*University of California, Davis*

In the  $ARMAX(p, r, q)$  model given  $(p, r, q)$ , the presence of multiple parameters is often ignored. In a *coprime* model Hannan has shown that the unrestricted reduced form (URF) parameter is identifiable under the simple condition that the end parameter matrix  $(\Pi_p, \Phi_r, \Psi_q)$  has full row rank. In applications it has been found convenient to assume, without test, that the model is coprime. But in *stable miniphase* models, coprime or noncoprime, the incidence of multiple parameters is very real whenever the tail end transfer (impulse) coefficient matrix  $\mathfrak{S}_{URF}$  has less than full row rank. This matrix contains the transfer matrices at lags  $r$  and  $q$  and beyond. If the timeseries process is significantly anchored in its past,  $\mathfrak{S}_{URF}$  has full row rank and the URF parameter is unique. This is testable in large samples. The rank of  $\hat{\mathfrak{S}}_{URF}$  assists in deciding uniqueness of the URF parameter, in quantifying the common factor that generates its multiplicity and in identifying a restricted reduced form (RRF) model.

---

\*Preliminary draft read at the Australian Econometric Society Meetings and at the Australian National University (1986). Recent drafts presented to the Bob Shumway Conference at Davis (2007) and to a Seminar at the Technische Universität Wien (2008). I am grateful for comments made by the late E.J.Hannan, by Hugo Kuersteiner, Bob Shumway and Manfred Deistler that have led to many clarifications in substance and style. All remaining misconceptions and errors are mine.

*AMS 2000 subject classifications:* Primary 37M10, 91G70

*Keywords and phrases:* equivalence, identifiability, multivariate Sylvester, coprimeness

**1. Introduction.** With autocorrelated errors represented as a moving average, Hannan E. (1969, 1971) derived the very simple result that in  $m$ -variate *coprime* models the reduced form  $ARMA(p, q)$  parameter is identifiable if and only if the end matrix  $\Pi_p$  in the autoregression and  $\Psi_q$  in the moving average jointly have rank  $m$ . The vast literature analyzing coprimeness and extending Hannan's core result is summarized in Hannan and Deistler (1988). Coprimeness implies assuming *a priori* that any left common factor (LCF) is "polynomial with a polynomial inverse". A coprimeness test would require rank tests over an infinite domain and is possible only if the parameter is given. As a LCF is a ratio of polynomial matrices, a polynomial LCF with polynomial inverse must be exceptional. Typically, either a LCF or its inverse or both are nonpolynomial and as shown below, if  $\Pi_p$  is nonsingular, any  $LCF \neq I_m$  must be an infinite power series.

Deistler M. (1976) asked if the assumption of coprimeness is a natural one and doubted that it is for the econometrician (statistician) inspired by economic (subject matter) theory. Clearly it is desirable to analyze parameter identifiability without having to assume that the model is coprime. Glover and Willems (1974) already obtained necessary and sufficient identifiability conditions for minimal state-space models and Deistler and Schrader (1979) derived sufficient identifiability conditions for a class of  $ARX(p, r, \cdot)$  models, both without assuming coprimeness.

In this paper the  $ARMAX(p, r, q)$  model, coprime or noncoprime alike, is considered under standard stability and MA invertibility hypotheses. Its essential result is that the unrestricted reduced form (URF) parameter is identifiable if and only if a  $p$ -stacked tail end transfer coefficient submatrix has full row rank  $pm$ . The intuition is that an identifiable process must have full impact extending to and beyond period  $q$ . If the rank is less than  $pm$  alternative parameters exist and these are derived.

The  $ARMA(p, q)$  model is introduced in Section 2 with its assumptions and notations. Section 3 describes the properties that equivalent models have in *Lemma 1* and *Lemma 2*. It concludes with the equivalence existence *Theorem 1* and with *Corollary 1* for the special model with commuting parameter. In Section 4 LCF properties are summarized and the conditions under which they are polynomial or nonpolynomial are obtained. In Section 5 the equivalence condition of *Theorem 1* is contrasted to results of the coprime literature. In Section 6 the  $ARMA(p, q)$  model is converted into a state-space model. The Glover and Willems (1974) identifiability condition is found to be identical to that of *Theorem 1* when  $q = p$ , and is sufficient but not necessary when  $q \neq p$ . The technical details are placed in an Appendix. A singular multivariate Sylvester matrix is shown to be a necessary

condition for the existence of a polynomial or a nonpolynomial LCF in Section 7. Extensions of the results to the URF, to the restricted reduced form (RRF) and to the structural  $ARMAX(p, r, q)$  model follow in *Theorem 2* and *Theorem 3* of Sections 8 and 9. The recommendation that a rank test of the stacked tail end transfer coefficient submatrix be made part of the parameter estimation procedure, however imperfect, is a natural implication of the analysis.

**2. ARMA( $p, q$ ) Assumptions and Notation.** The simplest time-series model in which the essential result of the paper holds is the unrestricted reduced form *URF ARMA*( $p, q$ ) model. For known positive integers  $p$  and  $q$  this is the stochastic process  $(y_t)_{t \in \mathcal{T}}$  of endogenous variables in  $\mathbb{R}^m$  satisfying

$$(1) \quad \sum_{i=0}^p \Pi_i y_{t-i} = \sum_{j=0}^q \Psi_j \eta_{t-j}, \quad t \in \mathcal{T}, \quad \mathcal{T} = \{1, \dots, T\},$$

where  $\Pi_0 = \Psi_0 = I_m$ ,  $\Pi_i \in \mathbb{R}^{m \times m}$  is the  $i$ -th AR parameter matrix and  $\Psi_j \in \mathbb{R}^{m \times m}$  the  $j$ -th MA parameter matrix. Both  $\Pi_i$  and  $\Psi_j$  are bounded and  $(\eta_t)_{t \in \mathcal{T}}$  are independent error variables on an underlying probability space with state space  $\mathbb{R}^m$ , mean  $E(\eta_t) = 0$  and positive definite covariance  $\Omega = E(\eta_t \eta_t')$ . The index is the set of natural numbers.

Let  $\Pi(z) := \sum_{i=0}^p \Pi_i z^i$ ,  $\Psi(z) := \sum_{j=0}^q \Psi_j z^j$  with  $z$  a complex scalar. The matrix polynomial  $z$ -transforms  $(\Pi(z), \Psi(z))$  will be called the model and

$$(\Pi^a(z), \Psi^a(z)) = \left( \sum_{i=0}^{p^a} \Pi_i^a z^i, \sum_{j=0}^{q^a} \Psi_j^a z^j \right), \quad \Pi_0^a = \Psi_0^a = I_m,$$

is an *alternative* model. Among the stability and miniphase assumption, see (e.g., Hannan and Deistler, 1988), the model assumptions are summarized under the following.

ASSUMPTION 1. *The  $(y_t)_{t \in \mathcal{T}}$  process (1) is causal and for  $|z| \leq 1$ , the set of polynomials*

$$(2) \quad (I_m, \Pi(z), \Pi^a(z), \Psi(z), \Psi^a(z), \Pi(z)^{-1}, \Pi^a(z)^{-1}, \Psi(z)^{-1}, \Psi^a(z)^{-1})$$

*is a ring of converging power series, with addition component-wise and multiplication given by the Cauchy product.*

Under Assumption 1,  $\Pi(z)$  is invertible as a formal matrix power series in the closed unit disk and the  $(y_t)_{t \in \mathcal{T}}$  process is unique given  $(\eta_t)_{t \in \mathcal{T}}$ .

The *final form* or *transfer* model is  $T_\eta(z) := \Pi(z)^{-1}\Psi(z)$ . Under Assumption 1 it is a converging series with unique transfer coefficient  $T_{\eta,i}$  at lag  $i = 1, \dots$ . The vector  $(T_{\eta,1}, T_{\eta,2}, \dots)$  constitutes the external identifiable observational information of the model.

The model  $(\Pi(z), \Psi(z))$  and the alternative  $(\Pi^a(z), \Psi^a(z))$  are *observationally equivalent* if they imply the same final form power series  $T_\eta(z) = \Pi(z)^{-1}\Psi(z) = \Pi^a(z)^{-1}\Psi^a(z)$ . It follows that the model and the alternative are observationally equivalent if and only if there exists a power series  $R(z) = \sum_{i=0}^{\infty} R_i z^i$ ,  $R_i \in \mathbb{R}^{m \times m}$  such that  $(\Pi^a(z), \Psi^a(z)) = R(z)(\Pi(z), \Psi(z))$  with leading term  $R_0 = I_m$ . By construction  $R(z) = \Pi^a(z)\Pi(z)^{-1}$  is a rational polynomial. The standard algebra and coprime literature deal mainly with polynomial  $R(z)$  and here all rational  $R(z)$  with converging power series are considered.

Observationally equivalent models  $(\Pi(z), \Psi(z))$  and  $(\Pi^a(z), \Psi^a(z))$  are said to be *degree equivalent* if  $\Pi^a(z)$  has degree  $p^a \leq p$  and  $\Psi^a(z)$  has degree  $q^a \leq q$ . Degree equivalence is the condition  $0 = \left( \Pi_{p+i}^a, \Psi_{q+i}^a \right) = \sum_{j=1}^{\infty} R_j (\Pi_{p+i-j}, \Psi_{q+i-j})$ ,  $\forall i, 0 < i$ . Models are said to be *equivalent* if they are observationally and degree equivalent. The matrix  $R(z)$  will be called a left common factor (LCF) if both conditions are satisfied.

*Identifiability* conditions imply that equivalent models  $(\Pi^a(z), \Psi^a(z)) \neq (\Pi(z), \Psi(z))$  do not exist. The contribution of this paper is the derivation of necessary and sufficient conditions under which equivalent models satisfying Assumption 1, do or do not exist.

#### NOTATION 1. *Block Toeplitz matrices*

In what follows, when matrices are block matrices with either a  $m \times m$  transfer coefficient matrix  $T_{\eta,i}$ , or an AR parameter matrix  $\Pi_i$ , or an MA parameter matrix  $\Psi_i$  as element, the noun block in naming a block row or block column is dropped. In time series analysis many matrices are Toeplitz matrices with the  $(i, j)$ -th element depending only on  $j - i$ . Such matrices can be defined by the 4 corner elements. As used below,  $(T_\eta)_{-k+i; -k+j}^{i;j}$  is the matrix having  $T_{\eta,i}$ ,  $T_{\eta,j}$ ,  $T_{\eta,-k+i}$ ,  $T_{\eta,-k+j}$  in its corners. Its top row  $(T_\eta)_{i;j} = (T_{\eta,i}, \dots, T_{\eta,j})$  has  $T_{\eta,i}$  as first and  $T_{\eta,j}$  as last element. Similarly its bottom row  $(T_\eta)_{-k+i; -k+j}$  has  $T_{\eta,-k+i}$  as first and  $T_{\eta,-k+j}$  as last element. In the same way define  $\Pi_{-k+i; -k+j}^{i;j}$  as the matrix having the AR autoregressive parameter matrices  $\Pi_i$  as elements and in  $\Psi_{-k+i; -k+j}^{i;j}$  the elements are the MA parameter matrices  $\Psi_i$ . The superscripts (subscripts) name the corner elements of the top (bottom) row. In a row subscripts are rising and in a column they are falling. A matrix with a negative subscript is a null matrix. If the matrix is a single vector only the bottom or the left hand indices are

reported. For example, if  $R(z)$  has degree  $j$  and  $R_{1;j} = (R_1, \dots, R_j)$ , the condition that  $(\Pi^a(z), \Psi^a(z))$  and  $(\Pi(z), \Psi(z))$  are equivalent is that

$$(3) \quad (\Pi^a(z), \Psi^a(z)) = R(z)(\Pi(z), \Psi(z)), \quad 0 = R_{1;j}(\Pi_{p-j+1;p}^{p;p+j-1}, \Psi_{q-j+1;q}^{q;q+j-1}),$$

as formulated for the  $ARMA(p, q)$ -model in Deistler, M. (1983).

With  $\rho(Z)$  denoting the rank of  $Z$ , the central result of the paper is this.

### 3. ARMA(p,q) Equivalence Condition.

LEMMA 1. *The observationally equivalent URF model  $(\Pi^a(z), \Psi^a(z)) = R(z)(\Pi(z), \Psi(z))$ , with  $R(z) = \sum_{i=0}^{\infty} R_i z^i$ ,  $R_0 = I_m$ , a converging power series, is degree equivalent if and only if  $R_{1;\infty}$  and  $(\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p)$  satisfy the equations*

$$(4) \quad 0 = (\Pi_1^a - \Pi_1, \Pi_2^a - \Pi_2, \dots, \Pi_p^a - \Pi_p) \mathfrak{S}_{URF,\eta},$$

$$(5) \quad R_{1;p} = (\Pi_1^a - \Pi_1, \Pi_2^a - \Pi_2, \dots, \Pi_p^a - \Pi_p) (\Pi_{1-p;0}^{0;p-1})^{-1}$$

and for  $t = (1, 2, \dots)$

$$(6) \quad ((R_{p+1}; R_{2p}), \dots, (R_{tp+1}; R_{(t+1)p}), \dots) = R_{1;p} (G, \dots, G^t, \dots)$$

where  $G := -\Pi_{1;p}^{p;2p-1} \Pi_{1-p;0}^{0;p-1})^{-1}$  and

$$\mathfrak{S}_{URF,\eta} := (T_{\eta})_{q-p+1;q+pm-p}^{q;q+pm-1} = \begin{pmatrix} T_{\eta,q} & T_{\eta,q+1} & \dots & T_{\eta,q+pm-1} \\ T_{\eta,q-1} & T_{\eta,q} & \dots & T_{\eta,q+pm-2} \\ \vdots & \vdots & \dots & \vdots \\ T_{\eta,q-p+1} & T_{\eta,q-p+2} & \dots & T_{\eta,q+pm-p} \end{pmatrix}.$$

The URF ARMA(p, q) parameter  $((\Pi_i)_{i=1,\dots,p}, (\Psi_i)_{i=1,\dots,q}, \Omega)$  is identifiable if and only if

$$(7) \quad \rho(\mathfrak{S}_{URF,\eta}) = pm.$$

*Proof.* Observational equivalence means  $\Psi^a(z) = \Pi^a(z)\Pi(z)^{-1}\Psi(z)$ . Subtracting  $\Psi(z)$  from both sides of the equation obtain

$$\Psi^a(z) - \Psi(z) = \Pi^a(z)\Pi(z)^{-1}\Psi(z) - \Psi(z) = (\Pi^a(z) - \Pi(z))T_{\eta}(z).$$

After substituting  $\Pi_i^a = \Pi_i = 0$ ,  $p < i$ ,  $\Psi_j^a = \Psi_j = 0$ ,  $q < j$ , and  $T_{\eta,k} = 0$ ,  $k < 0$ , the coefficients of  $z^{q+j}$ ,  $0 < j$  on both sides are

$$\Psi_{q+j}^a - \Psi_{q+j} = (\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p) (T'_{\eta,q+j-1}, \dots, T'_{\eta,q+j-p})' = 0$$

and  $\Psi^a(z)$  degree equivalence over  $\Psi_{q+j}^a = 0$ ,  $j = 1, \dots, J$ , requires

$$0 = (\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p) (T_\eta)_{q-p+1; q+J-p}^{q; q+J-1}$$

Since  $\Pi(z)T_\eta(z) = \Psi(z)$  implies  $(T_\eta)_{q+t-p+1}^{q+t} = \mathbb{J}_p(T_\eta)_{q+t-p}^{q+t-1}$ ,  $t \geq 1$ , where

$$pm \times pm \mathbb{J}_p := \begin{pmatrix} -\Pi_1 & \dots & -\Pi_{p-1} & -\Pi_p \\ I_m & \dots & 0 & 0 \\ \vdots & \dots & \vdots & \vdots \\ 0 & \dots & I_m & 0 \end{pmatrix},$$

the Caley-Hamilton theorem implies that at most  $pm$  conditions  $\Psi_{q+j}^a = 0$ ,  $j = 1, \dots, pm$ , are independent and  $\rho \left( (T_\eta)_{q-p+1; q+J-p}^{q; q+J-1} \right) = \rho(\mathfrak{S}_{URF, \eta})$  for  $pm \leq J$ . Therefore, (4) is the  $\Psi^a(z)$  degree equivalence condition over all  $\Psi_{q+j}^a = 0$ ,  $j = 1, \dots$ , and  $\rho(\mathfrak{S}_{URF, \eta}) = pm$  is the condition (7) that the AR and thus the MA and positive definite  $\Omega$  parameter is unique.

From  $\Pi^a(z) = R(z)\Pi(z)$  for  $i \in \{1, \dots, p\}$ , the  $\Pi_i^a$  coefficient satisfies  $\Pi_i^a - \Pi_i = R_{1;p}\Pi_{i-p}^{i-1}$ , implying  $(\Pi^a - \Pi)_{1;p} = R_{1;p}\Pi_{1-p;0}^{0;p-1}$  and (5). For  $i \in \{p+1, \dots\}$ ,  $\Pi^a(z) = R(z)\Pi(z)$  implies  $0 = R_{1;p}\Pi_{i-p}^{i-1} + R_{p+1;i}\Pi_0^{i-(p+1)}$  and

$$0 = (\Pi^a - \Pi)_{p+1;i} = R_{1;p}\Pi_{1;i-p}^{p;i-1} + R_{p+1;i}\Pi_{p+1-i;0}^{0;i-(p+1)}.$$

Setting  $i = tp$ , (6) is seen from the partition  $\Pi_{1;tp-p}^{p;tp-1} = (\Pi_{1;p}^{p;2p-1}, \mathcal{O})$ , with  $\mathcal{O}$  the  $pm \times (t-1)p$  null matrix, multiplying into the inverse of  $\Pi_{1+p-tp;0}^{0;tp-p-1}$ . This latter matrix is bidiagonal, having  $\Pi_{1-p;0}^{0;p-1}$  in the diagonal,  $\Pi_{1;p}^{p;2p-1}$  in the upper diagonal and the top row of its inverse is the geometric sequence  $(\Pi_{1-p;0}^{0;p-1})^{-1} (I_{pm}, G, \dots, G^{t-1})$ , with  $G := -\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1}$ .  $\square$

The degree equivalence condition of observationally equivalent models is composed of the  $\Psi^a(z)$  degree equivalence condition (4), the  $\Pi^a(z)$  determining role of  $R(z)$  in condition (5) and the  $\Pi^a(z)$  degree equivalence condition (6). The latter condition reveals how the LCF  $R(z)$  is either a polynomial of degree at most  $p$  or  $R_i$ ,  $p < i$ , is a matrix geometric converging power series. For any given  $(\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p)$  solving (4), conditions (5) and (6) can be solved uniquely and therefore condition (4) alone is the necessary and sufficient degree equivalence condition in observationally equivalent models.

LEMMA 2. *The  $\Psi^a(z)$  degree equivalence condition (4) implies the URF model  $(\Pi^a(z), \Psi^a(z)) = R(z)(\Pi(z), \Psi(z))$  is an equivalent model with*

$$(8) \quad R(z) = I_m + \left( \sum_{i=1}^p z^i \sum_{j=1}^i R_j \Pi_{i-j} \right) \Pi(z)^{-1}.$$

*Proof.* Given  $\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p$  from (4), observe that the conjectured relation (8) satisfies

$$\begin{aligned} R(z) &= I_m + R_{1;p} \Pi_{1-p;0}^{0;p-1} (zI_m, z^2I_m, \dots, z^pI_m)' \Pi(z)^{-1} \\ &= I_m + ((\Pi_1^a - \Pi_1)z + (\Pi_2^a - \Pi_2)z^2 + \dots + (\Pi_p^a - \Pi_p)z^p) \Pi(z)^{-1} \\ &= I_m + (\Pi^a(z) - \Pi(z)) \Pi(z)^{-1} \end{aligned}$$

and therefore,

$$\begin{aligned} (\Pi^a(z), \Psi^a(z)) &= (I_m + (\Pi^a(z) - \Pi(z)) \Pi(z)^{-1}) (\Pi(z), \Psi(z)) \\ &= \left( \Pi^a(z), \Psi(z) + \sum_{j=1}^{\infty} \sum_{i=1}^p (\Pi_i^a - \Pi_i) z^i T_{\eta, j-i} z^{j-i} \right) \\ &= (\Pi^a(z), \Psi(z) + \sum_{j=1}^q \sum_{i=1}^p ((\Pi_i^a - \Pi_i) z^i T_{\eta, j-i} z^{j-i})), \end{aligned}$$

after substituting from (4) that  $\sum_{i=1}^p (\Pi_i^a - \Pi_i) z^i T_{\eta, j-i} z^{j-i} = 0$ , for  $q+1 \leq j$ . The alternative model is thus observationally equivalent.  $\square$

Lemma 1 shows that observationally equivalent models are degree equivalent under condition (4) and Lemma 2 shows that degree equivalent models are observationally equivalent under condition (4). This may be summarized in an existence theorem.

**THEOREM 1 (Equivalent Model Existence).**  *$(\Pi^a(z), \Psi^a(z))$  equivalent to the unrestricted reduced form ARMA( $p, q$ ) model  $(\Pi(z), \Psi(z))$  exists if and only if condition (4) holds with  $(\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p) \neq 0$ , i.e., if and only if  $\rho(\mathfrak{S}_{URF, \eta}) < pm$ .*

In (4) the AR parameter differences between equivalent models are described by  $m$  vectors. Each vector is in the null space of  $\mathfrak{S}_{URF, \eta}$ . Some vectors may be null, but not all if there is any difference between the models. In the top row of  $\mathfrak{S}_{URF, \eta}$  stand the white noise transfer coefficient matrices at lags  $q, q+1, \dots$ . Going down the rows in each column the lag is reduced by one period. If  $\mathfrak{S}_{URF, \eta}$  has rank  $pm$  the alternative model is equal to the given model. The effect of the random error process  $(\eta_t)_{t \in T}$  on the evolution of the time series process  $(y_t)_{t \in \mathcal{T}}$  at lags  $q$  and/or beyond has a significant role in identifying the parameter. If the effect at lags  $q$  and beyond is null the parameter would not be unique. The contribution of the transfer coefficient matrix  $T_{\eta, q}$  is important. On the other hand, when  $p > 1$ , it is neither necessary nor sufficient for identifiability that  $T_{\eta, q}$  is nonsingular. In the extreme

case that  $T_{\eta,q} = 0$ , the parameter still may be identifiable. For example in the bivariate  $ARMA(2, 2)$  model  $T_{\eta,1} = \Psi_1 - \Pi_1$ ,  $T_{\eta,2} = \Psi_2 - \Pi_2 - \Pi_1 T_{\eta,1}$  and when  $T_{\eta,2} = 0$  the tail end transfer submatrix

$$\mathfrak{S}_{URF,\eta} = (T_{\eta})_{1,4}^{2,5} = \begin{pmatrix} 0 & -\Pi_2 T_{\eta,1} & \Pi_1 \Pi_2 T_{\eta,1} & (\Pi_2^2 - \Pi_1^2 \Pi_2) T_{\eta,1} \\ T_{\eta,1} & 0 & -\Pi_2 T_{\eta,1} & \Pi_1 \Pi_2 T_{\eta,1} \end{pmatrix}$$

can have full row rank 4 and the parameter identifiable.

With consistent estimators  $\hat{T}_{\eta,i}$ ,  $i = 1, \dots, q + pm - 1$ , of the white noise transfer coefficient matrices, the identifiability condition (7) is testable through a rank test. If the rank condition fails estimators of the dimension of the null space are available as a guide for the researcher in considering restricted reduced form (RRF) models.

REMARK 1. The equivalence condition (4) is valid for any population, coprime or not coprime, of size  $T \geq (m+1)p + q$ , where  $mp + q$  observations are needed to define transfer coefficients lagged up to  $mp + q - 1$  periods, in addition to the  $p + 1$  initial observations needed in (1). This number is what Tigelaar H. (1983) showed to be the second-order minimum informative sample size in coprime models.

REMARK 2. Deistler, M. (1983), in this notation, derives the equation

$$(T_{\eta,q+1}, T_{\eta,q+2}, \dots, T_{\eta,q+mp}) = -(\Pi_1, \Pi_2, \dots, \Pi_p) \mathfrak{S}_{URF,\eta} \quad (D.10)$$

between transfer coefficients and AR coefficients. If  $(D.10)^a$  is the equation  $(D.10)$  with the AR parameter vector replaced by  $\Pi_{1;p}^a$  of an equivalent model, this paper's equivalence condition (4) is the equation  $(D.10) - (D.10)^a$ . Under Assumption 1, it is the unique equivalence condition in any model, coprime or not coprime. If the AR and MA transforms commute there is a simpler criterion.

COROLLARY 1. *When  $\Pi(z)$  and  $\Psi(z)$  commute, the URF  $ARMA(p, q)$  models  $(\Pi(z), \Psi(z))$  and  $(\Pi^a(z), \Psi^a(z))$  are equivalent if and only if*

$$(9) \quad 0 = (\Pi_1^a - \Pi_1, \Pi_2^a - \Pi_2, \dots, \Pi_p^a - \Pi_p) (T_{\eta})_{q-p+1;q}^{q;q+p-1}$$

*and the parameter is identifiable if and only if  $\rho\left((T_{\eta})_{q-p+1;q}^{q;q+p-1}\right) = pm$ .*

*Proof.* With commuting  $z$ -transforms,  $\Pi(z)^{-1}\Psi(z) = \Psi(z)\Pi(z)^{-1} = T_{\eta}(z)$  imply  $T_{\eta}(z)\Pi(z) = \Psi(z)$  and  $T_{\eta,q+t} = -(T_{\eta,q+t-p}\Pi_p + \dots + T_{\eta,q+t-1}\Pi_1)$ .

For  $t \geq 1$  we have

$$(10) \begin{pmatrix} T_{\eta, q+t+p-1} \\ \vdots \\ T_{\eta, q+t} \end{pmatrix} = - \begin{pmatrix} T_{\eta, q+t-1} & \cdots & T_{\eta, q+t+p-2} \\ \vdots & \cdots & \vdots \\ T_{\eta, q+t-p} & \cdots & T_{\eta, q+t-1} \end{pmatrix} \begin{pmatrix} \Pi_p \\ \vdots \\ \Pi_1 \end{pmatrix},$$

implying  $\rho(\mathfrak{S}_{URF, \eta}) = \rho\left((T_\eta)_{q-p+1; q}^{q; q+p-1}\right)$ .  $\square$

Corollary 1 implies that there does not exist a common root between two scalar polynomials of degrees  $p$  and  $q$  if and only if  $(T_\eta)_{q-p+1; q}^{q; q+p-1}$  is non-singular. Below this condition is seen to be identical to the condition that the  $m$ -variate Sylvester matrix is nonsingular and a singular multivariate Sylvester matrix is shown to be necessary for an equivalent model to exist.

**4. Polynomial and Nonpolynomial LCF.** Under the equivalence existence condition  $\rho(\mathfrak{S}_{URF, \eta}) < pm$ , there exists  $R(z) = \sum_{i=0}^{\infty} R_i z^i \neq I_m$  satisfying  $0 = R_{1;p} \Pi_{1-p;0}^{0;p-1} \mathfrak{S}_{URF, \eta}$ ,  $R_{1;p} \neq 0$ , with these properties.

**Property 1.**  $\rho(R_{1;p}) = pm - \rho(\mathfrak{S}_{URF, \eta})$  and  $R_{p+1;\infty}$  is linear in  $R_{1;p}$ , as follows from (5) and (6) of Lemma 1.

**Property 2.** There exists a continuous family of LCF  $\bar{R}(z) = I_m + \sum_{i=1}^{\infty} F R_i z^i$ , with  $m \times m$   $F$  an arbitrary matrix, generating corresponding alternative equivalent models, as follows from first degree homogeneity of  $R_{1;\infty}$  in  $\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p$  through (4), (5) and (6) of Lemma 1.

**Property 3.** The polynomial  $\sum_{i=0}^p R_i z^i$  is a LCF if and only if

$$(11) \quad 0 = R_{1;p} \left( \Pi_{1-p;0}^{0;p-1} \mathfrak{S}_{URF, \eta}, \Pi_{1;p}^{p;2p-1} \right) \\ = R_{1;p} \left( \Psi_{q-p+1; q+pm-p}^{q; q+pm-1} - \Pi_{1;p}^{p;2p-1} (T_\eta)_{q-2p+1; q-2p+pm}^{q-p; q-p+pm-1}, \Pi_{1;p}^{p;2p-1} \right).$$

This follows from equation (6). To derive the second equality in (11), the relation  $\Psi_k = T_{\eta, k} + \sum_{i=1}^p \Pi_i T_{\eta, k-i}$ ,  $k \in \{0, 1, \dots, \}$  is used.

The distinction between polynomial and nonpolynomial LCF hinges on whether the geometric power sequence in (6) is finite or not finite.

**DEFINITION 1.** *The  $mp \times mp$  matrix  $N$  is nilpotent of degree  $g$  if  $N^g = 0$ .  $N$  nilpotent implies  $|N| = 0$ ,  $tr N = 0$ , its characteristic polynomial is  $\lambda^{mp}$  and  $g \leq mp$ . The matrix  $N$  is not nilpotent if all its positive powers are not null.*

Conditions (4), (5), (6) and  $G = -\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1}$  being a nilpotent matrix imply:

COROLLARY 2. *The polynomial  $\sum_{i=0}^{gp} R_i z^i$ ,  $2 \leq g \leq mp$ , is a LCF if and only if*

$$\begin{aligned} i) \quad & 0 = R_{1;p} \Pi_{1-p;0}^{0;p-1} \mathfrak{S}_{URF,\eta}, \\ ii) \quad & R_{tp+1;(t+1)p} = R_{1;p} (-\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1})^t, t = 1, \dots, g-1, \\ iii) \quad & 0 = (\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1})^g. \end{aligned}$$

Condition i) repeats that the  $p$  leading terms in a LCF have a primary role. An equivalent model  $(\Pi^a(z), \Psi^a(z))$  differs from  $(\Pi(z), \Psi(z))$  if and only if  $\Pi^a(z) - \Pi(z) \neq 0$  and this difference is described in (5) through at most  $p$  matrices of  $R_{1;p}$ . Condition ii) defines the secondary role of  $R_{tp+1;(t+1)p}$ . It must contribute to  $\Pi_{tp+1;(t+1)p}^a$  the matrix  $R_{tp+1;(t+1)p} \Pi_{1-p;0}^{0;p-1}$  to cancel out the contribution  $R_{(t-1)p+1;tp} \Pi_{1;p}^{p;2p-1}$  made to it by its left neighbour. This role requires that  $R_{tp+1;(t+1)p}$  follow the matrix-variate geometric series with ratio  $G$ . Finally, condition iii) sets  $R_i = 0$ ,  $gp < i$  and completes the required  $\Pi^a(z)$  degree equivalence. Corollary 2 opens new insight in the LCF structure.

**Property 4.** The power sequence  $\sum_{i=0}^{\infty} R_i z^i$  is a *nonpolynomial* LCF if and only if

$$\begin{aligned} i) \quad & 0 = R_{1;p} \Pi_{1-p;0}^{0;p-1} \mathfrak{S}_{URF,\eta}, \\ ii) \quad & R_{tp+1;(t+1)p} = R_{1;p} (-\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1})^t, t = 1, \dots, \\ iii) \quad & 0 \neq (\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1})^t, t = 1, \dots, \end{aligned}$$

where sufficient conditions for non-nilpotency *iii)* are

$$(12) \quad \left| \Pi_{1;p}^{p;2p-1} \right| \neq 0 \text{ or } \text{tr}(\Pi_{1;p}^{p;2p-1}) \neq 0.$$

In summary, the polynomial  $\sum_{i=0}^d R_i z^i$  of degree  $d$  is an LCF if and only if it satisfies  $R_{1;p} \Pi_{1-p;0}^{0;p-1} \mathfrak{S}_{URF,\eta} = 0$  with  $R_{1;p} \neq 0$  and from (6) its degree  $d$

- i) is less than or equal to  $p$  if and only if  $R_{1;p} \Pi_{1;p}^{p;2p-1} = 0$
- ii) satisfies  $p+1 \leq d \leq gp$ , if and only if  $\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1}$  is nilpotent of degree  $g$ ,  $2 \leq g$ , i.e.,  $\Pi(z)^{-1}$  is polynomial
- iii) converges to infinity as a nonpolynomial geometric power series if and only if  $\Pi_{1;p}^{p;2p-1} (\Pi_{1-p;0}^{0;p-1})^{-1}$  is not nilpotent, i.e.,  $\Pi(z)^{-1}$  is not polynomial.

REMARK 3. In constructing examples of equivalent URF  $ARMA(p, q)$  models with  $p > 1$  and for a variety of  $R_{1;p}$ , use  $\Psi_k = \sum_{i=0}^p \Pi_i T_{\eta, k-i}$  to rewrite the condition (4) in the form

$$(13) \quad 0 = (R_1, \dots, R_i, \dots, R_p) \Pi_{1-p;0}^{0;p-1} \mathfrak{S}_{URF,\eta}$$

where the  $i$ -th row of  $\Pi_{1-p;0}^{0;p-1} \mathfrak{S}_{URF,\eta}$ ,  $i = 1, \dots, p$ , is the matrix

$$\left( \left( \begin{array}{c} \Psi_{q-i+1} - \sum_{j=p-i+1}^p \\ \Pi_j T_{\eta, q-i+1-j} \end{array} \right) \cdots \left( \begin{array}{c} \Psi_{q-i+pm} - \sum_{j=p-i+1}^p \\ \Pi_j T_{\eta, q-i+pm-j} \end{array} \right) \right).$$

**5. Identifiability in coprime models.** It is convenient to assume that *coprime* models are a subset in the ring of convergent power series listed at (2).

ASSUMPTION 2. *The URF  $(\Pi(z), \Psi(z))$  model and all equivalent alternative URF models  $(\Pi^a(z), \Psi^a(z))$  satisfy Assumption 1 and for all  $z \in \mathcal{C}$ , the complex scalar domain,  $\rho(\Pi(z), \Psi(z)) = \rho(\Pi^a(z), \Psi^a(z)) = m$ .*

The algebraic theory surveyed by McDuffee C. (1933) and applied in Kailath T. (1980), Hannan and Deistler (1988), provides the analysis. Polynomial matrices are studied there in a coprime population. In this population, if coprime  $(\Pi(z), \Psi(z))$  and coprime  $(\Pi^a(z), \Psi^a(z))$  are equivalent satisfying (3), both  $R(z)$  and  $R(z)^{-1}$ , are polynomial and thus unimodular. The untested assumption of coprimeness together with the received polynomial algebraic theory removes from consideration nonpolynomial and nonunimodular LCF  $R(z)$  in the equivalence relation (3). Both are incompatible with the 'dual' coprimeness assumption of the model population.

Hannan E. (1969) originated the application and derived the result that the parameter of the coprime  $ARMA(p, q)$  model is identifiable if and only if  $\rho(\Pi_p, \Psi_q) = m$ . If  $R_1(\Pi_p, \Psi_q) = 0$  with  $R_1 \neq 0$ , the first degree polynomial  $I_m + R_1 z$  is a LCF since condition (4)

$$\begin{aligned} 0 &= (R_1, 0, \dots, 0) \Pi_{1-p;0}^{0;p-1} (T_\eta)_{q-p+1; q+pm-p}^{q; q+pm-1} \\ &= (R_1, 0, \dots, 0) \begin{pmatrix} \Psi_q - \Pi_p T_{\eta, q-p} & \Psi_{q+1} - \Pi_p T_{\eta, q-p+1} & \cdots \\ \sum_{i=0}^{p-2} \Pi_i T_{\eta, q-1-i} & \sum_{i=0}^{p-2} \Pi_i T_{\eta, q-i} & \cdots \\ \cdots & \cdots & \cdots \\ T_{\eta, q-p+1} & T_{\eta, q-p+2} & \cdots \end{pmatrix} \end{aligned}$$

is satisfied in the coprime population, as Hannan E. (1969) showed.

Deistler M. (1976, pg.37) derived that the degree of a polynomial LCF is at most  $mp$  in coprime populations. Correspondingly, Deistler, M. (1983) conceived of the polynomial LCF existence condition, here applied to the URF coprime  $ARMA(p, q)$  model, as the condition (3) when  $j = mp$ .

Deistler's condition for the polynomial  $\sum_{i=0}^{mp} R_i z^i$ ,  $R_{mp} \neq 0$  with finite integer  $m \geq 1$ , to be a LCF is that the array  $R_{1;mp}$  satisfies

$$0 = R_{1;mp}(\Pi_{-mp+p+1;p}^{p;p+mp-1}, \Psi_{-mp+q+1;q}^{q;q+mp-1}) = (R_{1;p}, R_{p+1;mp}) \begin{pmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{pmatrix}.$$

To see how Deistler's condition in coprime populations compares to the polynomial LCF condition derived in the previous section, consider the partition

$$\left( \begin{array}{c|c} Z_1 & Z_2 \\ \hline Z_3 & Z_4 \end{array} \right) = \left( \begin{array}{cccc|cccc} \Delta & 0 & \dots & 0 & 0 & \Xi_1 & 0 & \dots & 0 & 0 \\ \Lambda & \Delta & \dots & 0 & 0 & \Xi_2 & \Xi_1 & \dots & 0 & 0 \\ 0 & \Lambda & \dots & 0 & 0 & \Xi_3 & \Xi_2 & \dots & 0 & 0 \\ \vdots & \vdots \\ 0 & 0 & \dots & \Lambda & \Delta & \Xi_m & \Xi_{m-1} & \dots & \Xi_2 & \Xi_1 \end{array} \right)$$

$$\Delta := \Pi_{1;p}^{p;2p-1}, \quad \Lambda := \Pi_{1-p;0}^{0;p-1} \quad \text{and} \quad \Xi_i := \Psi_{q-ip+1;q-(i-1)p}^{q-(i-1)p;q-(i-2)p-1}, \quad i = 1, \dots, m.$$

Deistler's condition  $0 = R_{(m-1)p+1;mp}(\Pi_{1;p}^{p;2p-1}, \Psi_{q-p+1;q}^{q;q+p-1})$  implies Hannan's condition  $0 = R_{mp}(\Pi_p, \Psi_q) = 0$ ,  $R_{mp} \neq 0$  and  $I_m + zR_{mp}$  is a LCF.

With  $Z_3$  bi-diagonal and  $G = -\Pi_{1;p}^{p;2p-1}(\Pi_{1-p;0}^{0;p-1})^{-1}$ , the top row of  $Z_3^{-1}$  is

$$(\Pi_{1-p;0}^{0;p-1})^{-1} (I_{mp}, G, \dots, G^{m-2}) = (P_{1-p;0}^{0;p-1}, P_{1;p}^{p;2p-1}, \dots, P_{(m-3)p+1;(m-2)p}^{(m-2)p;(m-1)p-1}),$$

where  $P_i$ ,  $i = 0, \dots$ , are the matrices in the power series  $\sum_{i=0}^{\infty} P_i z^i = \Pi(z)^{-1}$ . Deistler's condition with  $q \leq mp - 1$  is equivalent to

$$\begin{aligned} R_{p+1;mp} &= -R_{1;p} Z_1 Z_3^{-1}, \text{ implying } R_{(j-1)p+1;jp} = R_{1;p} G^{j-1}, \quad j = 1, \dots, m, \\ 0 &= R_{1;p} (Z_2 - Z_1 Z_3^{-1} Z_4) \\ &= R_{1;p} \left( G^{m-1} \Delta, G^j \sum_{i=1}^{m-j} G^{i-1} \Xi_i, \quad j = 0, \dots, m-1 \right). \end{aligned}$$

After substitution of  $G^j = \Pi_{1-p;0}^{0;p-1} P_{(j-1)p+1;jp}^{jp;(j+1)p-1}$  verify that

$$G^j \sum_{i=1}^{m-j} G^{i-1} \Xi_i = \Pi_{1-p;0}^{0;p-1} \left( \sum_{i=j}^{m-1} P_{(i-1)p+1;ip}^{ip;(i+1)p-1} \Psi_{q-(i-j)p+1;q-(i-j)p}^{q-(i-j)p;q-(i-j)p-1} \right).$$

Finally using  $T_{\eta, q+jp+\ell} = \sum_{i=0}^{mp-1} P_{jp+\ell+i} \Psi_{q-i}$  when  $\ell = 0, p-1$  or  $1-p$ , Deistler's condition is found to be the requirement that  $R_{1;p}$  satisfy

$$\begin{aligned} 0 &= R_{1;p} \left( -G^m, \Pi_{1-p;0}^{0;p-1} (T_{\eta})_{q+(j-1)p+1;q+jp}^{q+jp;q+(j+1)p-1}, \quad j = 0, \dots, m-1 \right) \\ &= R_{1;p} \left( -G^m, \Pi_{1-p;0}^{0;p-1} (T_{\eta})_{q-p+1;q+mp-p}^{q;q+mp-1} \right). \end{aligned}$$

Deistler's condition contains the polynomiality condition in the first term.  $R_{1;p}G^m = 0$  could be satisfied by  $R_{1;p}\Pi_{1;p}^{p;2p-1} = 0$ , implying the LCF is of degree  $\leq p$ , or by  $G^m = 0$  when the model's  $\Pi_{1;p}^{p;2p-1}(\Pi_{1-p;0}^{0;p-1})^{-1}$  is nilpotent of degree  $m$  and the LCF has degree between  $p$  and  $mp$ , which was the desired result. When  $R_{1;p}G^m \neq 0$  Deistler's condition implies a LCF of degree  $mp$  does not exist. The second condition is the general equivalence existence condition (4) which must be satisfied, even when the first condition holds under nilpotency of  $G$ . It confirms the primary role of  $R_{1;p}$  in the LCF, which is not apparent in his condition. If nonpolynomial  $R(z)$  are admissible as LCF, the polynomiality condition can be ignored and Deistler's condition is equivalent to condition (4). From Property 2 it is obvious that the search for unimodular  $R(z)$  is not advanced by imposing a polynomiality condition. Every polynomial and nonpolynomial LCF  $R(z)$  generates a family of LCF, containing both unimodular and nonunimodular members, with the unimodular members occupying the set of measure zero where  $|R(z)|$  is independent of  $z$ . There are no unimodular LCF existence theorems to be discovered since they exist whenever a LCF  $R(z)$  exists. Similarly, tests for unimodular LCF are not separable from tests for any LCF and these are tests of condition (4).

Hannan E. (1971) and Deistler, M. (1983) consider restricted reduced form (RRF) models by building column degree information into the parameter matrices, if the highest lag of some variable in some equation is known to be less than  $p$  or  $q$ , respectively. Since column degree information defines additional AR and MA exclusion restrictions, these can be added to the conditions (4) that restrict the differences in the AR parameter to lie in the null-space of  $\mathfrak{S}_{\eta,URF}$ . Under this option, column degree restrictions belong either to the minimum necessary restrictions for identifiability or are testable *over – identifying* restrictions, when the URF parameter is identifiable without them.

REMARK 4. Under Assumption 1 the observational equivalence condition  $(\Pi^a(z), \Psi^a(z)) = R(z)(\Pi(z), \Psi(z))$  is identical to the condition that

$$(14) \quad (\Pi(z), \Psi(z))U(z) = (R(z)^{-1}, 0)$$

with

$$2m \times 2m \quad U(z) := \begin{pmatrix} \Pi^a(z)^{-1} & -\Pi(z)^{-1}\Psi(z)\Pi^a(z) \\ 0 & \Pi^a(z) \end{pmatrix}.$$

Observe that  $U(z)$  is unimodular and (14) formally satisfies the construction in Kailath T. (1980, pg.377). If the alternative also satisfies degree equivalence,  $R(z)$  is a LCF and  $(\Pi^a(z), \Psi^a(z))$  an equivalent model. The equations

(14) always have the solution  $R(z) = I_m$ , implying  $\Pi^a(z) = \Pi(z)$ . Euclidean methods such as Hermite form constructions are available if both  $R(z)$  and its inverse are polynomial. To calculate equivalent solutions with  $R(z) \neq I_m$ , simple linear Euclidian pivotal methods do not generate equivalent models unless guided by the unknown  $\Pi^a(z)$ .

**6. Identifiability in minimal state-space models.** The label minimality as used in the state space literature is quite distinct from the coprime assumption. The results of this paper and those contained in the Proposition of Glover and Willems (1974) are easily compared since both are stated in terms of the transfer coefficients. The concept of minimality as defined in the Linear Systems literature is best illustrated after converting the  $ARMA(p, q)$  model into state-space form. Whereas there is more than one way to accomplish this, consider the following construction.

With  $M := \text{Max}(p, q)$ , let  $z_{t+1} := E[(y'_{t+M}, \dots, y'_{t+1})' | t]$ ,  $z_t \in \mathbb{R}^{mM}$ , be the vector of conditional expectations of  $M$  future values of  $y_t$  in period  $t$ . Given  $\Pi_{p+i} = \Psi_{q+i} = 0$ ,  $0 < i$ , model (1) implies

$$E\left[\left(\sum_{i=0}^M \Pi_i y_{t+M-i} - \sum_{i=0}^M \Psi_i \eta_{t+M-i}\right) | t-1\right] = 0 \quad \text{and}$$

$$E(y_{t+i} | t) = E(y_{t+i} | t-1) + T_{\eta, i} \eta_t.$$

Its state-space prediction error form is the AR and observation equation

$$(15) \quad z_{t+1} = \mathbb{J}_M z_t + (T_\eta)_1^M \eta_t, \quad y_t = C z_t + \eta_t,$$

where  $m \times mM$   $C := (0, \dots, I_m)$  and

$$\mathbb{J}_M := \begin{pmatrix} -\Pi_1 & \dots & -\Pi_{M-1} & -\Pi_M \\ I_m & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & I_m & 0 \end{pmatrix}.$$

Solving the AR equation, the reachability condition requires that

$$mM = \rho\left((T_\eta)_1^M, \mathbb{J}_M (T_\eta)_1^M, \dots, \mathbb{J}_M^{T-1} (T_\eta)_1^M\right) = \rho\left((T_\eta)_{1;T}^{M;T+M-1}\right)$$

$$(16) \quad = \rho\left(\begin{pmatrix} \Psi_{p+1;q}^{q;2q-p-1} & 0 \\ (T_\eta)_{1;q-p}^{p;q-1} & (T_\eta)_{q-p+1;T}^{q;T+p-1} \end{pmatrix}\right) \quad \text{when } p < q = M, mM \leq T,$$

$$(17) \quad = \rho\left((T_\eta)_{1;T}^{p;T+p-1}\right) \quad \text{when } q \leq p = M, mM \leq T,$$

where (16) follows by replacing the element  $T_{\eta,i}$  in the top  $q - p$  rows of  $(T_{\eta})_{1;T}^{q;T+q-1}$  by the vector product  $\Pi_{0;p}(T_{\eta})_{i-p}^i = \Psi_i$ , beginning with the top row and ending with the  $(q - p) - th$  row. With  $T_{\eta,i}$  the top element of  $(T_{\eta})_{i-p}^i$  and the other elements standing below it in the same column, the rank equality follows. With  $(T_{\eta})_{p+1;q}^{q;2q-p-1}$  a lower recursive matrix with  $\Psi_q$  on the diagonal, (15) is reachable if and only if  $\rho((T_{\eta})_{q-p+1;qm}^{q;qm+p-1}) = pm$  and  $\Psi_q$  is nonsingular. Clearly  $\rho(C', \mathbb{J}'_M C', \dots, \mathbb{J}'_M{}^{mM-1} C') = mM$  so that observability holds and (15) is in minimal form if (16) and (17) hold.

The Glover-Willems Proposition holds under minimality. Their analysis uses the property that alternative state-space parameters  $\delta^a$  are linked to the given assumed parameter  $\delta$  through the transformation  $J_M(\delta^a) = FJ_M(\delta)F^{-1}$ ,  $(T_{\eta}^a)_{1;1}^{M;M} = F(T_{\eta})_{1;1}^{M;M}$ ,  $C^a = CF^{-1}$  with a nonsingular matrix  $F$  of order  $mM$ . The Glover-Willems identifiability conditions are full column rank conditions on the matrix of derivatives of the alternative state-space parameters with respect to the transformation matrix  $F$  and with respect to the  $ARMA(p, q)$  parameter, at  $F = I_{mM}$ . In the Appendix it is shown that these full column rank conditions are satisfied so that the minimality conditions and the Glover-Willems URF parameter identifiability conditions are identical. For easy comparison the latter and the identifiability condition (7) of this paper are listed here.

$p, q$	Minimality and Glover-Willems Identifiability Condition	Condition (7) without Minimality
$p < q$	$\rho((T_{\eta})_{q-p+1;qm}^{q;qm+p-1}) = pm,  \Psi_q  \neq 0$	$\rho((T_{\eta})_{q-p+1;q+pm-p}^{q;q+pm-1}) = pm$
$q \leq p$	$\rho((T_{\eta})_{1;pm}^{p;pm+p-1}) = pm$	$\rho((T_{\eta})_{q-p+1;q+pm-p}^{q;q+pm-1}) = pm$

When  $p < q$  the Glover-Willems condition on the tail end transfer coefficient matrix and condition (7) are identical, since  $\rho((T_{\eta})_{q-p+1;qm}^{q;qm+p-1}) = \rho((T_{\eta})_{q-p+1;q+pm-p}^{q;q+pm-1})$  when  $p < q$ , but the condition that the matrix  $\Psi_q$  is nonsingular, as required for reachability, is not required for condition (7). When  $q = p$  the Glover-Willems condition and condition (7) are identical. When  $q < p$  the condition (7) requires that

$$\begin{aligned} pm &= \rho((T_{\eta})_{q-p+1;q+pm-p}^{q;q+pm-1}) = \rho((T_{\eta})_{q-p+1;p+pm-p}^{q;p+pm-1}) \\ &= \rho \left( \begin{array}{cc} (T_{\eta})_{q-p+1;0}^{q;p-1} & (T_{\eta})_{1;pm}^{p;pm+p-1} \end{array} \right), \end{aligned}$$

implying Glover-Willems is sufficient but not necessary for condition (7). The matrix  $(T_{\eta})_{q-p+1;0}^{q;p-1}$  contributes  $p - q$ ,  $q < p$ , extra columns to condition

(7), each column containing the identity matrix  $T_{\eta,0} = I_m$  in a different row, in addition to the columns that determine the reachability condition.

In summary, the Glover-Willems condition is a sufficient identifiability condition for the URF  $ARMA(p, q)$  parameter and identical to it if  $p = q$ . When  $p \neq q$  the condition under minimality differs in some detail from the conditions without minimality. In particular, condition (7) does not require a nonsingular  $\Psi_q$  when  $p < q$  and an extra matrix  $(T_\eta)_{q-p+1;0}^{q;p-1}$  contributes to the required rank when  $q < p$ .

**7. Matricvariate Sylvester matrix and Identifiability.** The identifiability results and the common factor properties are implications of the tail end transfer coefficient matrix to have or not to have full row rank. This result can be more informative about the model when it is stated in terms of the AR  $\Pi_i$  and MA  $\Psi_i$  parameter matrices. Ultimately the failure of the rank condition is rooted in a feature of the parameter matrices and there are many causes. Sylvester calculated that two scalar polynomials have no common root (factor) if and only if, in today's language, the Sylvester matrix is nonsingular [Sylvester 1840], where the elements of the Sylvester matrix are the coefficients of the two scalar polynomials. If in the Sylvester matrix the scalar coefficients are replaced by the AR and the MA parameter matrices, it is shown now that the  $ARMA(p, q)$  parameter is identifiable if the matricvariate Sylvester matrix is nonsingular and a necessary, but not sufficient, condition for the existence of a common factor is that the matricvariate Sylvester matrix is singular. If the  $(\Pi(z), \Psi(z))$  transforms commute this condition is also sufficient. Sylvester's scalar common root existence condition is thereby generalized to the matrix common factor existence condition.

**DEFINITION 2 (Multivariate Sylvester Matrix).** The  $m$ -variate Sylvester matrix in partitioned form is the  $(p+q)m \times (q+p)m$  matrix

$$(18) \quad S(p, q) := (\Pi_{1-q;0}^{p;p+q-1}, \Psi_{1-p;0}^{q;q+p-1}) = \left( \frac{\Pi_{1;q}^{p;p+q-1} \mid \Psi_{q-p+1;q}^{q;q+p-1}}{\Pi_{1-q;0}^{0;q-1} \mid \Psi_{1-p;0}^{q-p;q-1}} \right)$$

$$= \left( \begin{array}{ccc|ccc} \Pi_p & \cdots & \Pi_{p+q-1} & \Psi_q & \cdots & \Psi_{q+p-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \Pi_1 & \cdots & \Pi_q & \Psi_{q-p+1} & \cdots & \Psi_q \\ \hline \Pi_0 & \cdots & \Pi_{q-1} & \Psi_{q-p} & \cdots & \Psi_{q-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \Pi_{1-q} & \cdots & \Pi_0 & \Psi_{1-p} & \cdots & \Psi_0 \end{array} \right).$$

COROLLARY 3 (Sylvester Identifiability). *The URF ARMA( $p, q$ ) parameter is identifiable if the matrix-variate Sylvester matrix (18) is nonsingular. When  $(\Pi(z), \Psi(z))$  commute a nonsingular matrix-variate Sylvester matrix is also a necessary identifiability condition.*

*Proof.*  $\Pi_{1-q;0}^{0;q-1}$  a unitary upper recursive matrix of order  $qm$  implies

$$(19) \quad \rho(S(p, q)) = qm + \rho(\Psi_{q-p+1;q}^{q;q+p-1} - \Pi_{1;q}^{p;p+q-1} (\Pi_{1-q;0}^{0;q-1})^{-1} \Psi_{1-p;0}^{q-p;q-1}).$$

Equating the coefficients of  $z^i$ ,  $i = k, \dots, k-p-q+1$ , in  $\Pi(z)T_\eta(z) = \Psi(z)$ , one obtains the equations

$$\Pi_{-p-q+1;0}^{0;p+q-1} (T_\eta)_{-p-q+k+1}^k = \Psi_{-p-q+k+1}^k.$$

These equations for  $k = q, \dots, p+q-1$ , constitute the system

$$(20) \quad \Pi_{-p-q+1;0}^{0;p+q-1} (T_\eta)_{1-p;0}^{q;p+q-1} = \Psi_{1-p;0}^{q;p+q-1}.$$

Define the partitions

$$\Pi_{-p-q+1;0}^{0;p+q-1} = \begin{pmatrix} \Pi_{1-p;0}^{0;p-1} & \Pi_{1;q}^{p;p+q-1} \\ 0 & \Pi_{1-q;0}^{0;q-1} \end{pmatrix}, \quad \Psi_{1-p;0}^{q;p+q-1} = \begin{pmatrix} \Psi_{q-p+1;q}^{q;p+q-1} \\ \Psi_{1-p;0}^{q-p;q-1} \end{pmatrix}.$$

From (20) the solution of  $(T_\eta)_{1-p;0}^{q;p+q-1}$ , conformably partitioned, is

$$\begin{aligned} & \begin{pmatrix} (T_\eta)_{q-p+1;q}^{q;p+q-1} \\ (T_\eta)_{1-p;0}^{q-p;q-1} \end{pmatrix} = \begin{pmatrix} \Pi_{1-p;0}^{0;p-1} & \Pi_{1;q}^{p;p+q-1} \\ 0 & \Pi_{1-q;0}^{0;q+1} \end{pmatrix}^{-1} \begin{pmatrix} \Psi_{q-p+1;q}^{q;p+q-1} \\ \Psi_{1-p;0}^{q-p;q-1} \end{pmatrix} \\ & = \begin{pmatrix} (\Pi_{1-p;0}^{0;p-1})^{-1} & -(\Pi_{1-p;0}^{0;p-1})^{-1} \Pi_{1;q}^{p;p+q-1} (\Pi_{1-q;0}^{0;q-1})^{-1} \\ 0 & (\Pi_{1-q;0}^{0;q-1})^{-1} \end{pmatrix} \begin{pmatrix} \Psi_{q-p+1;q}^{q;p+q-1} \\ \Psi_{1-p;0}^{q-p;q-1} \end{pmatrix} \end{aligned}$$

and

$$(21) \quad \Pi_{1-p;0}^{0;p-1} (T_\eta)_{q-p+1;q}^{q;p+q-1} = (\Psi_{q-p+1;q}^{q;p+q-1} - \Pi_{1;q}^{p;p+q-1} (\Pi_{1-q;0}^{0;q-1})^{-1} \Psi_{1-p;0}^{q-p;q-1}).$$

With nonsingular  $\Pi_{1-p;0}^{0;p-1}$  of order  $pm$ , (19) and (21) imply

$$(22) \quad \rho(S(p, q)) = qm + \rho((T_\eta)_{q-p+1;q}^{q;p+q-1})$$

and  $S(p, q)$  is nonsingular if and only if  $\rho((T_\eta)_{q-p+1;q}^{q;p+q-1}) = pm$ . If  $(T_\eta)_{q-p+1;q}^{q;p+q-1}$  has rank  $pm$ , the tail end transfer coefficient matrix  $\mathfrak{S}_{URF,\eta}$  also has rank  $pm$ , the former being a submatrix of the latter, and the identifiability condition

(7) of Lemma 1 is satisfied if  $S(p, q)$  is nonsingular. Corollary 1 then implies that, if  $\Pi(z)$  and  $\Psi(z)$  commute, a nonsingular  $m$ -variate Sylvester matrix is necessary and sufficient for identifiability.  $\square$

Corollary 3 shows that any of the many roots of the  $m$ -variate Sylvester matrix may be the *root* cause for the tail end transfer coefficient matrix to have less than full row rank. From its top row it is seen that  $\rho(\Pi_p, \Psi_q) < m$  implies that the  $m$ -variate Sylvester matrix is singular and a common factor exists. Most other roots are hidden and not discernable algebraically, especially those implying nonpolynomial LCF. However, not every cause of singularity of  $S(p, q)$  implies that a common factor exists.

REMARK 5 (Identifiable, yet singular  $m$ -variate Sylvester matrix). It is possible that the  $m$ -variate Sylvester matrix is singular but the given parameter is identifiable. For example in the  $m$ -variate  $ARMA(1, 2)$  model with the parameter matrices  $(\Pi_1, \Psi_1, \Psi_2)$ , the  $3m \otimes 3m$  Sylvester matrix

$$S(1, 2) = (\Pi_{-1;0}^{1;2}, \Psi_0^2) = \begin{pmatrix} \Pi_1 & 0 & \Psi_2 \\ I_m & \Pi_1 & \Psi_1 \\ 0 & I_m & I_m \end{pmatrix}$$

has rank  $2m + \rho(T_{\eta,2})$  and the determinant  $|S(1, 2)| = |\Psi_2 - \Pi_1\Psi_1 + \Pi_1^2| = |T_{\eta,2}|$ , as seen from  $T_{\eta,2}(z) = \Pi(z)^{-1}\Psi(z)$ . The identifiability condition (7) is

$$\rho((T_{\eta})_{2;m+1}) = \rho \left( T_{\eta,2}, \quad (-\Pi_1)T_{\eta,2}, \quad \dots, \quad (-\Pi_1)^{m-1}T_{\eta,2} \right) = m.$$

$|S(1, 2)| = 0$  if  $R_1(\Pi_1, \Psi_2) = 0$  with  $R_1 \neq 0$  implying  $I_m + zR_1$  is a LCF. It could also be zero with  $|\Pi_1| \neq 0$  and with  $R_1\Pi_1^{i-1}T_{\eta,2} = 0$ ,  $i = 1, \dots, m$ , implying  $I_m + zR_1(I_m + z\Pi_1)^{-1}$  is a LCF. A bivariate example is

$$\Pi_1 = \begin{pmatrix} \pi_{11} & \pi_{12} \\ 0 & \pi_{22} \end{pmatrix}, \quad T_{\eta,2} = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 0 & b \\ 0 & c \end{pmatrix},$$

where  $a$ ,  $b$  and  $c$  are real numbers. A third case is that  $|S(1, 2)| = 0$  with  $|\Pi_1| \neq 0$  and  $\rho((T_{\eta})_{2;m+1}) = m$  implying  $R_1 = 0$  and an identifiable parameter. Changing the position of  $a$  to the (2,2) position in  $T_{\eta,2}$  above with  $\pi_{12} \neq 0$  is an example. This last case is not possible when parameter matrices commute, since  $\Pi_1 T_{\eta,2} = T_{\eta,2} \Pi_1$  implies  $\rho((T_{\eta})_{2;m+1}) \leq \rho(T_{\eta,2})$ .

In the remaining pages the URF  $ARMA(p, q)$  equivalence conditions are extended to the URF, the RRF and the structural  $ARMAX(p, r, q)$  models.

**8. The URF ARMAX(p,r,q).** For known positive integers  $p, r$  and  $q$  the unrestricted reduced form  $ARMAX(p, r, q)$  historical timeseries process is of the form

$$(23) \quad \sum_{i=0}^p \Pi_i y_{t-i} = \sum_{i=0}^r \Phi_i x_{t-i} + \sum_{i=0}^q \Psi_i \eta_{t-i} \quad t \in \mathcal{T}.$$

The parameter matrix  $\Phi_i \in R^{m \times k}$  and  $(x_t)_{t \in \mathcal{T}}$  with  $x_t \in R^k$  are bounded exogenous variables independent of the error variables  $(\eta_t)_{t \in \mathcal{T}}$ . The exogenous variable  $z$ -transform  $\Phi(z) := \sum_{i=0}^r \Phi_i z^i$  is added to the list of transforms satisfying Assumption 1. With  $z_t = ((x_t)', \dots, (x_{t-r})')'_{t \in \mathcal{T}}$ , the exogenous data matrix  $(r+1)k \times T$   $Z'_T = (z_1, \dots, z_T)/\sqrt{T}$  has full row rank over the sample. Variables such as a constant or time trend will have to be added to (23) as separate exogenous variables and are not explicitly presented here. Given these assumptions, the exogenous variable transfer (final form) power series  $T_x(z) := \Pi(z)^{-1}\Phi(z)$  is identifiable.

The model  $(\Pi(z), \Phi(z), \Psi(z))$  and the alternative  $(\Pi^a(z), \Phi^a(z), \Psi^a(z))$  are equivalent if and only if there exists a converging power series  $R(z) = \sum_{i=0}^{\infty} R_i z^i$  such that

$$(\Pi^a(z), \Phi^a(z), \Psi^a(z)) = R(z)(\Pi(z), \Phi(z), \Psi(z))$$

and for  $\forall i, 0 < i$ ,

$$(\Pi_{p+i}^a, \Phi_{r+i}^a, \Psi_{q+i}^a) = \sum_{j=1}^{\infty} R_j (\Pi_{p+i-j}, \Phi_{r+i-j}, \Psi_{q+i-j}) = 0.$$

The extension of Theorem 1 to the URF  $ARMAX(p, r, q)$  model is this.

**THEOREM 2 (Identifiability).** *The URF ARMAX(p, r, q) parameter*

$$\Theta := ((\Pi_i)_{i=(1, \dots, p)}, (\Phi_i)_{i=(0, \dots, r)}, (\Psi_i)_{i=(1, \dots, q)}, \Omega),$$

*the exogenous matrix  $Z'_T$  having rank  $(r+1)k$ , is identifiable if and only if*

$$(24) \quad \rho(\mathfrak{S}_{URF}) = pm, \quad \mathfrak{S}_{URF} := (\mathfrak{S}_{URF,x}, \mathfrak{S}_{URF,\eta}),$$

*with*

$$\mathfrak{S}_{URF,x} := (T_x)_{r-p+1; r+mp-p}^{r; r+mp-1}, \quad \mathfrak{S}_{URF,\eta} := (T_\eta)_{q-p+1; q+mp-p}^{q; q+mp-1}.$$

*Proof.* The exogenous variable contribution to equivalence follows verbatim from (4) and Lemma 1, in its proof replacing  $\Psi(z)$  by  $\Phi(z)$ ,  $q$  by  $r$  and  $T_\eta(z)$  by  $T_x(z)$ . Equivalence through the exogenous variable process requires that for  $r + pm - 1 \leq T$  we have

$$0 = \left( \Pi_1^a - \Pi_1, \quad \Pi_2^a - \Pi_2, \quad \dots, \Pi_p^a - \Pi_p \right) (T_x)_{r-p+1;T-p+1}^{r;T}.$$

(24) is the joint contribution of both the exogenous and the random error process to identifiability of the URF parameter.  $\square$

REMARK 6. Deistler and Schrader (1979) were the first to consider the  $ARX(p, r, \cdot)$  model, i.e., the model with correlated errors not necessarily following a  $MA(q)$  process. Without assuming coprimeness they derived a set of sufficient identifiability conditions. The condition  $\rho(\mathfrak{S}_{URF,x}) = pm$  is both a necessary and sufficient condition for their  $ARX(p, r, \cdot)$  model. Condition (24) is useful in testing hypotheses about the order  $q$  for a given  $r$  or the order  $r$  for a given  $q$ , given  $p$ .

COROLLARY 4 (Sylvester identifiability). *The URF ARMAX( $p, r, q$ ) parameter with the exogenous data matrix  $Z_T'$  having rank  $(r + 1)k$ , is identifiable if the conjoined Sylvester matrix*

$$S(p, r, q) := \left( \Pi_{1-N;0}^{p;p+N-1}, \Phi_{-p+1+r-N;r-N}^{r;r+p-1}, \Psi_{q-p+1-N;q-N}^{q;q+p-1} \right)$$

has rank  $m(p + N)$ , where  $N = \text{Max}(r, q)$ . When the parameter matrices commute this rank condition is also necessary.

*Proof.*

$$S(p, r, q) = \begin{pmatrix} \left( \Pi_{1;N}^{p;p+N-1} & \Phi_{r-p+1;r}^{r;r+p-1} & \Psi_{q-p+1;q}^{q;q+p-1} \right) \\ \Pi_{1-N;0}^{0;N-1} & \Phi_{r-p+1-N;r-N}^{r-p;r-1} & \Psi_{q-p+1-N;q-N}^{q-p;q-1} \end{pmatrix}$$

implies

$$\rho(S(p, r, q)) = mN + \rho \left( \Phi_{r-p+1;r}^{r;r+p-1} - \Pi_{1;N}^{p;p+N-1} (\Pi_{1-N;0}^{0;N-1})^{-1} \Phi_{r-p+1-N;r-N}^{r-p;r-1}, \right. \\ \left. \Psi_{q-p+1;q}^{q;q+p-1} - \Pi_{1;N}^{p;p+N-1} (\Pi_{1-N;0}^{0;N-1})^{-1} \Psi_{q-p+1-N;q-N}^{q-p;q-1} \right).$$

Writing the partitions  $\Pi_{1;N}^{p;p+N-1} = \begin{pmatrix} \Pi_{1;q}^{p;p+q-1} & \Pi_{q+1;N}^{p+q;p+N-1} \end{pmatrix}$ ,

$$\Pi_{1-N;0}^{0;N-1} = \begin{pmatrix} \Pi_{1-q;0}^{0;q-1} & \Pi_{1;N-q}^{q;N-1} \\ 0 & \Pi_{-N+q+1;0}^{0;N-q-1} \end{pmatrix}, \quad \Psi_{q-p+1-N;q-N}^{q-p;q-1} = \begin{pmatrix} \Psi_{-p+1;0}^{q-p;q-1} \\ 0 \end{pmatrix},$$

where  $\Pi_{k;k+j-i}^{i;j}$  is an empty matrix if  $j < i$ . When  $q \leq r$  it is seen that

$$\begin{aligned} & \Psi_{q-p+1;q}^{q;q+p-1} - \Pi_{1;N}^{p;p+N-1} (\Pi_{1-N;0}^{0;N-1})^{-1} \Psi_{q-p+1-N;q-N}^{q-p;q-1} \\ &= \Psi_{q-p+1;q}^{q;q+p-1} - \Pi_{1;q}^{p;p+q-1} (\Pi_{1-q;0}^{0;q-1})^{-1} \Psi_{1-p;0}^{q-p;q-1} = \Pi_{1-p;0}^{0;p-1} (T_\eta)_{q-p+1;q}^{q;q+p-1}, \end{aligned}$$

with the last equality shown at (21). Similarly it follows that

$$\Phi_{r-p+1;r}^{r;r+p-1} - \Pi_{1;N}^{p;p+N-1} (\Pi_{1-N;0}^{0;N-1})^{-1} \Phi_{r-p+1-N;r-N}^{r-p;r-1} = \Pi_{1-p;0}^{0;p-1} (T_x)_{r-p+1;r}^{r;r+p-1}.$$

With  $\Pi_{1-p;0}^{0;p-1}$  a unitary upper recursive matrix of order  $pm$ , we have

$$\rho(S(p, r, q)) = mN + \rho((T_x)_{r-p+1;r}^{r;r+p-1}, (T_\eta)_{q-p+1;q}^{q;q+p-1})$$

and  $S(p, r, q)$  having full row rank  $m(p + N)$  implies

$$\rho((T_x)_{r-p+1;r}^{r;r+p-1}, (T_\eta)_{q-p+1;q}^{q;q+p-1}) = mp.$$

The case  $r \leq q$  is shown in the same way.  $\square$

REMARK 7. *Restricted Reduced Form (RRF) Identifiability*

If  $\rho(\mathfrak{S}_{URF}) < pm$  the URF parameter is not identifiable, let  $\theta$  be the vector of the RF parameter  $\Theta$ . With a list  $r_R(\theta) = 0$  of RF parameter restrictions, such as column degree restrictions that specify the highest lag of some variable in some equation to be less than  $p, r, q$ , respectively, or of RF restrictions implied by the scalar component model structure of Tiao and Tsay [1989], the RRF parameter, satisfying  $r_R(\theta) = 0$ , may be identifiable. From observational equivalence we have

$$(\Phi_1^a - \Phi_1, \dots, \Phi_r^a - \Phi_r) = (\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p) (T_x)_{1-p;r-p}^{0;r-1},$$

$$(\Psi_1^a - \Psi_1, \dots, \Psi_q^a - \Psi_q) = (\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p) (T_\eta)_{1-p,q-p}^{0,q-1}$$

and these relations imply that linear RF parameter restrictions are linear equations in  $(\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p)$ . If  $\mathfrak{S}_{RRF_i}$  is the matrix  $\mathfrak{S}_{URF}$ , augmented by the coefficients of  $pm - \rho(\mathfrak{S}_{URF})$  restrictions involving the  $i$ -th RF equation, not counting  $(\Phi_0^a - \Phi_0)_i$ -restrictions, the  $i$ -th RRF equation parameter  $((\Pi_j^a)_{j=1,\dots,p}, (\Phi_j^a)_{j=0,\dots,r}, (\Psi_j^a)_{j=1,\dots,q})_i$  is equivalent when

$$(25) \quad 0 = (\Pi_1^a - \Pi_1, \dots, \Pi_{k_1}^a - \Pi_{k_1}, \dots, \Pi_p^a - \Pi_p)_i \mathfrak{S}_{RRF_i}$$

and identifiable if and only if  $\rho(\mathfrak{S}_{RRF_i}) = pm$ .

For example, if  $((\Pi_{k_1}^a - \Pi_{k_1})_{ij_1}, (\Phi_{k_2}^a - \Phi_{k_2})_{ij_2}, (\Psi_{k_3}^a - \Psi_{k_3})_{ij_3}) = 0$  are the restrictions holding in the population with  $pm - \rho(\mathfrak{S}_{URF}) = 3$ , the  $i$ -th row of AR differences  $(\Pi_1^a - \Pi_1, \dots, \Pi_p^a - \Pi_p)_i = 0$ , if and only if  $\rho(\mathfrak{S}_{RRF_i}) = pm$ , where

$$\mathfrak{S}_{RRF_i} := \begin{pmatrix} (\mathfrak{S}_{URF,x} \ \mathfrak{S}_{URF,\eta})_1 & 0 & (T_{x,k_2-1})^{j_2} & (T_{\eta,k_3-1})^{j_3} \\ \vdots & \vdots & \vdots & \vdots \\ (\mathfrak{S}_{URF,x} \ \mathfrak{S}_{URF,\eta})_{k_1} & (I_m)^{j_1} & (T_{x,k_2-k_1})^{j_2} & (T_{\eta,k_3-k_1})^{j_3} \\ \vdots & \vdots & \vdots & \vdots \\ (\mathfrak{S}_{URF,x} \ \mathfrak{S}_{URF,\eta})_p & 0 & (T_{x,k_2-p})^{j_2} & (T_{\eta,k_3-p})^{j_3} \end{pmatrix},$$

with  $Z_i$  the  $i$ -th row,  $Z^j$  the  $j$ -th column and  $Z_{ij}$  the  $(i, j)$ -th element of the matrix  $Z$ .

The RRF parameter is identifiable if and only if  $\rho(\mathfrak{S}_{RRF_i}) = pm$ ,  $i = 1, \dots, m$ , each row having a list of  $pm - \rho(\mathfrak{S}_{URF})$  RF coefficient restrictions. These have the effect of augmenting the  $\mathfrak{S}_{URF}$  matrix by unit columns for AR restrictions or by columns of the  $((T_x)_{1-p;r-p}^{0,r-1}, (T_\eta)_{1-p;q-p}^{0,q-1})$  matrices for EX and MA restrictions. These columns are not part of the matrix  $\mathfrak{S}_{URF}$ .

**9. The structural form ARMAX(p,r,q).** The structural form (SF)  $ARMAX(p, r, q)$  model is the timeseries process

$$(26) \quad \sum_{i=0}^p B_i y_{t-i} + \sum_{i=0}^r \Gamma_i x_{t-i} = \sum_{i=0}^q \Delta_i \epsilon_{t-i}, \quad t \in \mathcal{T},$$

with  $B_0$  nonsingular,  $\Delta_0 = I_m$ ,  $E(\epsilon_t) = 0$ ,  $E(\epsilon_t \epsilon_{t'}') = \delta_{tt'} \Sigma$ . Its RF is the model (23) by setting  $(\Pi_i = B_0^{-1} B_i)_{i=0, \dots, p}$ ,

$$(27) \quad (\Phi_i = -B_0^{-1} \Gamma_i)_{i=0, \dots, r}, \quad (\Psi_i = B_0^{-1} \Delta_i B_0)_{i=1, \dots, q}, \quad \eta_t = B_0^{-1} \epsilon_t,$$

implying  $\Omega = B_0^{-1} \Sigma B_0'^{-1}$  given the RF covariance  $E(\eta_t \eta_t') = \Omega$ .

The identifiability analysis of the SF  $ARMAX(p, 0, 0)$  parameter given an identifiable RF parameter is the work of the Cowles commission [Koopmans ed., 1950]. Their analysis is transferable to the SF  $ARMAX(p, r, q)$  parameter provided either the RF identifiability conditions of Theorem 2 or the RRF identifiability conditions of Remark 7, are satisfied.

In defining parameter coefficient restrictions it is necessary to write the URF and the unrestricted structural form (USF) parameter as a vector of vectors.

NOTATION 2. *Vector operations*

Let  $\kappa = \text{vec}K$  denote the row vector of the row of rows of the matrix  $K$ . The  $\text{vec}$  operator satisfies  $\text{vec}UVW = \text{vec}V(U' \otimes W)$ . Also  $\kappa^* = \text{vec}K^*$  denotes the row of rows of  $K^*$ , where  $K^*$  is the symmetric matrix  $K$  after omitting the elements above its diagonal. We have  $\kappa^* = \kappa \frac{\partial \kappa^*}{\partial \kappa'}$ , setting  $\frac{\partial (K^*)_{ij}}{\partial (K^*)_{ji}} = \frac{\partial (K^*)_{ij}}{\partial (K^*)_{ij}} = \frac{1}{2}$  and  $\kappa = \kappa^* \frac{\partial \kappa}{\partial \kappa'^*}$ , setting  $\frac{\partial (K_{ij})}{\partial (K^*)_{ij}} = \frac{\partial (K_{ji})}{\partial (K^*)_{ij}} = 1, i \geq j$ . For nonsingular  $K$  one obtains  $\frac{\partial (\text{vec}K^{-1})'}{\partial \text{vec}K} = -(K^{-1} \otimes (K')^{-1})$  and  $UVW$  symmetric implies  $\text{vec}V(U' \otimes W) = \text{vec}V'(W \otimes U')$ .

DEFINITION 3. *Parameter vectors*

The URF parameter  $\theta$  has the  $pm^2 + (r+1)km + qm^2 + m(m+1)/2$  components

$$(28) \quad \theta := (\pi, \phi, \psi, \omega^*) := ((\pi_i)_{i=1,\dots,p}, (\phi_i)_{i=0,\dots,r}, (\psi_i)_{i=1,\dots,q}, \omega^*).$$

The USF parameter  $\alpha$  has the  $(p+1)m^2 + (r+1)km + qm^2 + m(m+1)/2$  components

$$(29) \quad \alpha := (\beta, \gamma, \delta, \sigma^*) := ((\beta_i)_{i=0,\dots,p}, (\gamma_i)_{i=0,\dots,r}, (\delta_i)_{i=1,\dots,q}, \sigma^*).$$

The USF parameter  $\alpha(\theta)$  having the URF parameter  $\theta$  is the vector

$$(30) \quad \alpha(\theta) = \left( \beta_0(I_m \otimes \Pi_i)_{i=0,\dots,p}, -\beta_0(I_m \otimes \Phi_i)_{i=0,\dots,r}, \right. \\ \left. \beta_0(\Psi_i' B_0'^{-1} \otimes I_m)_{i=1,\dots,q}, \omega^* \frac{\partial \kappa}{\partial \kappa'^*} (B_0' \otimes B_0) \frac{\partial \kappa^*}{\partial \kappa'} \right)$$

since  $m \times m$  symmetric matrix  $K$  implies

$$\begin{aligned} \sigma^*(\theta) &= \text{vec}(B_0 \Omega B_0')^* \\ &= \text{vec}(B_0 \Omega B_0') \frac{\partial \kappa^*}{\partial \kappa'} = \omega(B_0' \otimes B_0) \frac{\partial \kappa^*}{\partial \kappa'} = \omega^* \frac{\partial \kappa}{\partial \kappa'^*} (B_0' \otimes B_0) \frac{\partial \kappa^*}{\partial \kappa'}. \end{aligned}$$

DEFINITION 4. *Regular restrictions*

Given the transfer matrices  $\mathfrak{S}_{URF}$  at (24), the column list of RF  $r_R(\theta) = 0$  and SF  $r_S(\alpha(\theta)) = 0$  parameter restrictions is regular if

- i)  $r_R(\theta)$  has  $\rho\left(\frac{\partial r_R(\theta)}{\partial \theta}\right) = m(pm - \rho(\mathfrak{S}_{URF}))$  and contains  $pm - \rho(\mathfrak{S}_{URF})$  restrictions linear in  $(\Pi_1, \dots, \Pi_p, \Phi_1, \dots, \Phi_r, \Psi_1, \dots, \Psi_q)_i$  such that the augmented matrix  $\mathfrak{S}_{RRF_i}$ , at (25) of Remark 7 has rank  $pm$ ,  $i = 1, \dots, m$ ,
- ii)  $r_S(\alpha(\theta))$  is differentiable with  $m^2 \leq \rho\left(\frac{\partial r_S}{\partial \alpha}\right) = \sharp r_S$ ,  $\rho\left(\frac{\partial r_S(\alpha(\theta))}{\partial \alpha(\theta)} \frac{\partial \alpha(\theta)'}{\partial \beta_0}\right) = m^2$  and these ranks are constant in the parameter space.

**THEOREM 3 (Structural Parameter Local Identifiability).** *With the exogenous datamatrix  $Z'_T$  having rank  $(r+1)k$ , under Assumption 1 and under a regular column list of  $RF$   $r_R(\theta) = 0$  and  $SF$   $r_S(\alpha) = 0$  parameter restrictions, the restricted  $SF$  parameter  $\alpha(\theta)$  having the  $RRF$   $\theta$  parameter, is locally identifiable if and only if*

$$(31) \quad \rho(\mathfrak{S}_{SF}) = m^2, \quad \text{where} \quad \mathfrak{S}_{SF} := \frac{\partial r_S(\alpha(\theta))}{\partial \alpha(\theta)} \frac{\partial \alpha(\theta)'}{\partial \beta_0}$$

is the matrix of derivatives of the  $SF$  restrictions  $r_S(\alpha(\theta)) = 0$  with respect to  $\beta_0$ . Written in the order of the structural parameter components

$$\begin{aligned} \mathfrak{S}_{SF} &= \sum_{i=0}^p \frac{\partial r_S}{\partial \beta_i} (I_m \otimes \Pi'_i) - \sum_{i=0}^r \frac{\partial r_S}{\partial \gamma_i} (I_m \otimes \Phi'_i) \\ &+ \sum_{i=1}^q \frac{\partial r_S}{\partial \delta_i} (I_m \otimes B_0'^{-1} \Psi'_i - B_0 \Psi_i B_0^{-1} \otimes B_0'^{-1}) + 2 \frac{\partial r_S}{\partial \sigma^*} \frac{\partial \kappa^*'}{\partial \kappa} (I_m \otimes B_0 \Omega) \\ &= \left[ \sum_{i=0}^p \frac{\partial r_S}{\partial \beta_i} (I_m \otimes B'_i) + \sum_{i=0}^r \frac{\partial r_S}{\partial \gamma_i} (I_m \otimes \Gamma'_i) \right. \\ &\left. + \sum_{i=1}^q \frac{\partial r_S}{\partial \delta_i} (I_m \otimes \Delta'_i - \Delta_i \otimes I_m) + 2 \frac{\partial r_S}{\partial \sigma^*} \frac{\partial \kappa^*'}{\partial \kappa} (I_m \otimes \Sigma) \right] [I_m \otimes B_0'^{-1}]. \end{aligned}$$

*Proof.* Under a regular list  $r_R(\theta)$  of restrictions, from Remark 7 and (25), the URF or RRF parameter  $\theta$  is identifiable.  $\mathfrak{S}_{SF}$  is the matrix of derivatives of the structural restrictions with respect to  $\beta_0$ , calculated from the links in (30) between the components of the structural parameter  $\alpha(\theta)$  to  $\beta_0$ , given the identifiable parameter  $\theta$ . In particular,  $\Delta_i = B_0 \Psi_i B_0^{-1}$  and  $\Sigma = B_0 \Omega B_0'^{-1}$  imply

$$\begin{aligned} \frac{\partial(\text{vec} \Delta_i)'}{\partial \beta_0} &= (I_m \otimes B_0'^{-1} \Psi'_i) \frac{\partial \beta_0'}{\partial \beta_0} + (B_0 \Psi_i \otimes I_m) \frac{\partial(\text{vec} B_0^{-1})'}{\partial \beta_0} \\ &= I_m \otimes B_0'^{-1} \Psi'_i - (B_0 \Psi_i \otimes I_m) (B_0^{-1} \otimes B_0'^{-1}) \\ \frac{\partial \sigma^*'}{\partial \beta_0} &= \frac{\partial}{\partial \beta_0} (\text{vec}(B_0 \Omega B_0') \frac{\partial \kappa^*'}{\partial \kappa})' = 2 \frac{\partial \kappa^*'}{\partial \kappa} (I_m \otimes B_0 \Omega) \frac{\partial \beta_0'}{\partial \beta_0}. \end{aligned}$$

Local uniqueness of  $\beta_0$  given  $\theta$ , thus of  $\alpha$ , is implied by the implicit function theorem.  $\square$

In the  $ARMAX(p, 0, 0)$  model the RF parameter is identifiable if the  $k \times T$  exogenous data matrix has full rank  $k$  and the errors have a positive definite covariance. Starting from an identifiable RF the contribution to the SF parameter identifiability by the first two terms of  $\mathfrak{S}_{SF}$  is analyzed by the Cowles Commission in Koopmans et. al. (1950) and Koopmans and Reiersol (1950). Global identifiability conclusions are obtained if the restrictions in those two first terms are linear. The contribution of general whole system and  $\Sigma$  restrictions to SF local identifiability is stated in Wegge L. (1965), with global identifiability results if the restrictions are recursive. MA parameter restrictions contribute the third term of  $\mathfrak{S}_{SF}$  to the rank identifiability criterion. With MA restrictions (31) is nonlinear in  $\beta_0$  and identifiability is local.

As in the work of the Cowles Commission it is held *useful* to distinguish RF from SF identifiability and to assert that, if the RF parameter  $\theta$  is identifiable, if the restrictions  $r_S(\alpha)$  are regular, i.e., independent and *adequate*, and if the rank at (31) is constant Fisher F. (1966), then the structural parameter is identifiable locally *generically*. The SF parameter restrictions  $r_S(\alpha) = 0$  are adequate if the list  $r_S$  contains at least  $m^2$  restrictions in total, with at least  $m$  components of the list  $r_S(\alpha) = 0$  involving at least  $m$  coefficients of the  $i - th$  row

$$((B_j)_{j=0,\dots,p}, (\Gamma_j)_{j=0,\dots,r}, (\Delta_j)_{j=1,\dots,q}, \Sigma^*)_i,$$

with at least one restriction specifying that some coefficient in row  $i$  has a known nonzero value,  $i = 1, 2, \dots, m$ , to avoid homogeneity of  $r_S(\alpha) = 0$  in the parameter of a single equation. The assertion above is found *practical* because it helps in distinguishing between minimum necessary restrictions and those unneeded "*over - identifying*" restrictions, that are testable.

**10. Conclusions.** Hannan E. (1969) discovered the RF parameter identifiability problem in the  $ARMA(p, q)$  model. He proved the simple result that in URF coprime models the parameter is identifiable if and only if the end parameter matrix  $(\Pi_p, \Psi_q)$  has full row rank. He also suggested that common factors in non-coprime models are removable. This suggestion is neither implementable nor always desirable. Information about the presence of nonunimodular common factors is at best available only if the parameter is known, not before its estimation. In the meantime, in most theoretical and applied statistical work the model, without testing, is assumed to be coprime.

Glover and Willems (1974) derived identifiability conditions that hold for coprime and noncoprime minimal state-space systems. Deistler M. (1976)

questioned the assumption of coprimeness and Deistler and Schrader (1979) analyzed identifiability in a special model without assuming coprimeness. This paper has been inspired by this literature and it contributes towards removing coprimeness from the list of assumptions, while retaining stability and minimum phase assumptions that are needed in the construction of transfer coefficient functions and common factor transforms. It advocates a testable procedure that replaces the coprimeness assumption.

Theorems 1 and 2 state the central result. It is shown that in stable models with invertible moving average there are no alternative URF equivalent parameters, if and only if a submatrix  $\mathfrak{S}_{URF}$  of transfer coefficients at lags  $q$ ,  $r$  and beyond, is of full row rank  $pm$ . The suggestion follows that the estimation procedure of the  $ARMAX(p, r, q)$  model incorporate a stepwise monitoring of the rank of that matrix. If the rank is  $\lambda$ ,  $pm - \lambda$  linear independent RF coefficient restrictions on each RF row may render the RRF parameter identifiable. In dealing with SF equations with an identifiable URF or RRF parameter, restrictions satisfying adequacy criteria of the Cowles Commission guide the researcher in distinguishing exactly identifying from over-identifying testable restrictions.

Under normality of the white noise error process, the RF conditional MLE  $\hat{\theta}$  is a recommended estimator. This is a locally converging sequence of generalized least squares estimators, as described in (e.g., Reinsel et. al., 1992) or an approximating sequence of least squares estimators advocated in Spliid H. (1983). A necessary and sufficient consistency condition is the identifiability condition  $\rho(\mathfrak{S}_{URF}) = pm$  or  $\rho(\mathfrak{S}_{RRF_i}) = pm, i = 1, \dots, m$ . This condition implies the nonsingularity of the information matrix Rothenberg T. (1973) and the local peakedness of the conditional likelihood. The Cragg and Donald (1996) rank test is applicable in testing the rank of  $\mathfrak{S}_{URF}$ . In applications with sample data, the calculated rank of any estimated matrix  $\hat{\mathfrak{S}}_{URF}$  is  $pm$  and the rank test is a test of the hypothesis that the calculated rank is less than  $pm$ . It is a test of nonidentifiability, so called in Koopmans and Hood [1953], assigning confidence levels to the calculated rank.

## References.

- Cragg, J.G. and S.G. Donald (1996), On the Asymptotic Properties of the LDU-Based Tests of the Rank of a Matrix, *Journal of the American Statistical Association*, 91(435), 1301-1309.
- Deistler M. (1976), The identifiability of Linear Econometric Models with autocorrelated errors, *International Economic Review*, 17, No 1, 26-46.
- Deistler M. (1978), The structural identifiability of Linear Models with autocorrelated errors in the case of cross-equation restrictions, *Journal of Econometrics* 8, 23-31.
- Deistler M. and J. Schrader (1979), Linear Models with Autocorrelated Errors, Identifiability

- bility in the Absence of Minimality Assumptions, *Econometrica*, Vol. 47, No. 2, March, 495-504.
- Deistler M. (1983), The properties of the parametrization of ARMAX systems and their relevance for structural estimation and dynamic specification, *Econometrica* 51, No. 4, 1187-1207.
- Fisher, F.M. (1966), *The Identification Problem in Econometrics*, McGraw Hill, New York.
- Glover K. and J.C.Willems (1974), Parametrizations of linear dynamic systems: canonical forms and identifiability, *IEEE Trans. Autom. Control* AC-19, 640-645.
- Hannan, E.J. (1969) The identification of vector mixed autoregressive-moving average systems, *Biometrika* 56, 223-225.
- Hannan, E.J. (1971) The identification problem for multiple equation systems with moving average errors, *Econometrica* 39, 751-765.
- Hannan, E.J. and M. Deistler (1988), *The Statistical Theory of Linear Systems*, John Wiley & Sons, New York.
- Hood, W.C. and T.C. Koopmans (eds.) (1953), *Studies in Econometric Method*, Cowles Foundation Monograph No. 14, Yale University Press.
- Kailath, T. (1980), *Linear Systems*, Prentice Hall Inc., N.J.
- Koopmans, T.C. and O. Reiersol (1950), The Identification of Structural Characteristics, *The Annals of Mathematical Statistics*, Vol. 21, No. 2, June, 165-181.
- Koopmans, T.C. ed. (1950) *Statistical Inference in Dynamic Economic Models*, Cowles Commission for Research in Economics, Monograph No. 10, John Wiley & Sons.
- McDuffee, C.C. (1933), *The Theory of Matrices*, Springer, Berlin.
- Reinsel, G.C., Basu S. and Yap, S.F. (1992), Maximum Likelihood Estimators in the Multivariate Autoregressive Moving Average Model from a Generalized Least Squares viewpoint, *Journal of Time Series Analysis*, Vol. 13, No 2, 133-145.
- Rothenberg, T.J. (1973), *Efficient estimation with apriori information*, Cowles Foundation Monograph 23, Yale University Press.
- Spliid, H. (1983), A Fast Estimation Method for the Vector Autoregressive Moving Average Model with Exogenous Variables, *Journal of the American Statistical Association*, Vol.78, No 384, 843-849.
- Sylvester, J. (1840) A Method of determining by mere Inspection the Derivatives from two Equations of any Degree, *The London and Edinburgh Philosophical Magazine and Journal of Science*, January-June, Vol. XVI, 132-135.
- Tiao, G. and Tsay, R. (1989), Model specification in multivariate timeseries (with discussions), *Journal of the Royal Statistical Society B*, 51, 157-213.
- Tigelaar, H. H. (1983) The informative sample size for dynamic multiple equation systems with moving average errors, *Statistica Neerlandica*, Vol. 37, Issue 4, 169-179.
- Wegge, L.L. (1965), Identifiability Criteria for a system of equations as a whole, *Australian Journal of Statistics*, 7, 67-77.

**11. Grover-Willems State-space Identifiability.** Further details about the conversion of the  $ARMA(p, q)$  model into a state-space model are given here and it is shown that for this state-space prediction error form model the minimality condition is the required Glover-Willems identifiability condition.

Let the state variables be the conditional expectations in period  $t$  of the future values of  $y_t$ . Simple rules are followed, information about past error variables is never lost and the expectation about a future value is equal to

the mean, i.e., we set  $E(\eta_{t-i}|t) = \eta_{t-i}$  and  $E(\eta_{t+1+i}|t) = 0$ ,  $0 \leq i$ .

From the causal transfer equations  $y_t = \sum_{i=0}^{\infty} T_{\eta,i} \eta_{t-i}$  obtain

$$E(y_{t+i}|t) = E(y_{t+i}|t-1) + T_{\eta,i} \eta_t, \quad 0 < i,$$

reflecting that  $\eta_t$  is revealed in period  $t$ . From the RF equations (1) obtain

$$\begin{aligned} E(y_{t+M}|t) &= -\sum_{i=1}^p \Pi_i E(y_{t+M-i}|t) + \sum_{i=0}^q \Psi_i E(\eta_{t+M-i}|t) \\ &= -\sum_{i=1}^p \Pi_i [E(y_{t+M-i}|t-1) + T_{\eta,M-i} \eta_t] + \Psi_q E(\eta_{t+M-q}|t) \\ &= -\sum_{i=1}^p \Pi_i [E(y_{t+M-i}|t-1) + T_{\eta,M} \eta_t], \end{aligned}$$

using  $\Psi_M = \sum_{j=0}^p \Pi_j T_{\eta,M-j}$ .

With  $M := \text{Max}(p, q)$ , with the state variables  $z_{t+1} := E(y_{t+1}^{t+M}|t)$  and with the state parameter  $\delta = (\pi, \psi)$ , where  $\pi = \text{vec}(\Pi_1, \dots, \Pi_p)$ ,  $\psi = (\psi_1, \dots, \psi_q)$  and  $\psi_j = \text{vec} \Psi_j$ ,  $\text{vec} Z$  being the row of rows of the matrix  $Z$ , the expectations are the equations

$$(32) \quad z_{t+1} = \mathbb{J}_M(\delta) z_t + T_{\eta}(\delta)_1^M \eta_t,$$

with

$$mM \times mM \quad \mathbb{J}_M(\delta) := \begin{pmatrix} -\Pi_1 & \dots & -\Pi_{M-1} & -\Pi_M \\ I_m & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & I_m & 0 \end{pmatrix}.$$

These and the observation equation  $y_t = Cz_t + \eta_t$ , with  $m \times mM$   $C := (0, \dots, 0, I_m)$ , are the state-space prediction error form equations of the  $ARMA(p, q)$  model.

As shown above when  $p < q$ , (32) is reachable if and only if  $\Psi_q$  is non-singular and  $\rho((T_{\eta})_{q-p+1;qm}^{q;qm+p-1}) = pm$ . When  $q \leq p$ , (32) is reachable if and only if  $\rho((T_{\eta})_{1;pm}^{p;pm+p-1}) = pm$ . The Glover-Willems identifiability theorem in minimal state-space models is this.

**PROPOSITION 1.** *Glover-Willems Proposition: [1974]. The parameter of the minimal state-space model (32) is identifiable if and only if  $\rho(Z(M)) =$*

$(mM)^2 + (p + q)m^2$  where

$$(33) \quad Z(M) := \begin{pmatrix} I_{mM} \otimes \mathbb{J}_M(\delta)' - \mathbb{J}_M(\delta) \otimes I_{mM} & \frac{\partial(\text{vec}\mathbb{J}_M(\delta))'}{\partial\delta} \\ I_{mM} \otimes (T_\eta(\delta)_1^M)' & \frac{\partial(\text{vec}T_\eta(\delta)_1^M)'}{\partial\delta} \\ -(0, \dots, I_m) \otimes I_{mM} & 0 \end{pmatrix}.$$

It will be shown that  $Z(M)$  of the  $ARMA(p, q)$  model has full column rank  $(mM)^2 + (p + q)m^2$ . Write (33) in partitioned form as

$$Z(M) = \begin{pmatrix} Z_1 & Z_2 & \frac{\partial j_M'}{\partial\pi} & 0 \\ I_{mM} \otimes (T_\eta(\delta)_1^M)' & \frac{\partial t_M'}{\partial\pi} & \frac{\partial t_M'}{\partial\psi} \\ 0 & -I_{Mm^2} & 0 & 0 \end{pmatrix},$$

where  $j_M = \text{vec}\mathbb{J}_M(\delta)$ ,  $t_M = \text{vec}T_\eta(\delta)_1^M$  and  $(Mm)^2 \times (Mm)^2$

$$(Z_1, Z_2) = \begin{pmatrix} A + B_1 & B_2 & \dots & B_{M-1} & B_M \\ C & A & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & A & 0 \\ 0 & 0 & \dots & C & A \end{pmatrix},$$

with  $A := I_m \otimes J_M(\delta)'$ ,  $B_i := \Pi_i \otimes I_{mM}$ ,  $C := -I_{Mm^2}$  and the last (block) column in  $(Z_1, Z_2)$  is  $Z_2 = (B_M', 0, \dots, 0, A)'$ .

If  $0_{r,s}$  is a  $r \times s$  null matrix, the matrix of partial derivatives

$$(34) \quad \frac{\partial j_M'}{\partial\pi} = \begin{pmatrix} D \\ 0_{(M-1)Mm^2, pm^2} \end{pmatrix}, \quad \text{with } D := I_m \otimes \begin{pmatrix} I_{mp} \\ 0_{m(M-p), mp} \end{pmatrix}.$$

Letting  $E := I_{mM} \otimes (T_\eta(\delta)_1^M)'$  after substitution in (34) one gets

$$Z(M) = \begin{pmatrix} A + B_1 & B_2 & \dots & B_{M-1} & B_M & D & 0 \\ C & A & \dots & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ddots & A & 0 & 0 & 0 \\ 0 & 0 & \dots & C & A & 0 & 0 \\ & & E & & & \frac{\partial t_M'}{\partial\pi} & \frac{\partial t_M'}{\partial\psi} \\ 0 & 0 & \dots & 0 & C & 0 & 0 \end{pmatrix},$$

Eliminating sequentially the last row, the upper recursive matrix having  $C$  in its diagonal and the second to last column one has

$$\rho(Z(M)) = (mM)^2 + pm^2 + \rho\left(\frac{\partial t_M'}{\partial\psi}\right).$$

From the relations  $T_{\eta,i} = \Psi_i - \sum_{k=1}^i \Pi_k T_{\eta,i-k}$  we have

$$\frac{\partial(\text{vec}T_{\eta,i})'}{\partial\psi_j} = \delta_{ij}I_{m^2} - \sum_{k=1}^{i-1} (\Pi_k \otimes I_m) \frac{\partial(\text{vec}T_{\eta,i-k})'}{\partial\psi_j}$$

where  $\delta_{ij}$  is the Kronecker delta. Since  $\frac{\partial(\text{vec}T_{\eta,i-k})'}{\partial\psi_j} = 0$  for  $i < j$ , the matrix  $\frac{\partial t'_M}{\partial\psi}$  has  $I_{m^2}$  as element when  $i = j$ , the null matrix when  $i < j$  and has rank  $qm^2$ . The Glover-Willems condition  $\rho(Z(M)) = (mM)^2 + (p+q)m^2$  is satisfied by the reduced form minimal state-space system (32).  $\square$

DEPARTMENT OF ECONOMICS  
PRINTEADE1,U1