Statistics

1981

ARMAX-Model Parameter Identification
without and with Latent Variables

by

Leon L.F. Wegge

Working Paper Series
No. 178

September 1981

Department of Economics
University, Davis of California
Davis, California

ARMAX-Model Parameter Identification without and with Latent Variables

Leon L.F. Wegge

University of Bonn and University of California, Davis

Summer 1981.

Research supported by the Sonderforschungsbereich 21 of the University of Bonn, on leave there from the University of California at Davis.

ARMAX-Model Parameter Identification without and with Latent Variables

by Leon L. Wegge 1/

There are three parts to this paper. In the first part we obtain identifiability conditions for ARMAX-Model parameters without latent variables when the external characteristics are a finite number of means and covariances. We show how the conditions obtained here are related to conditions discussed by Fisher [4], Hatanaka [7], Hannan [6], Deistler [1] and others in the literature on asymptotic identifiability. Our results are formally complete in the sense of characterizing the quantity and quality of extra prior information when Fisher's Conditions or Hannan's rank condition on the parameter matrices of highest order or if his minimality conditions do not hold. Having conditions for identifiability that require a finite number of first and second moments has obvious advantages in suggesting estimators that are natural composites of the identifying function and the estimating function for the external information.

In the second part of the paper latent variables are introduced and the results of the first part are re-stated when the external information are the subvectors of first moments and the principal submatrices of the intertemporal covariances associated with the observed variables. As a general rule, with latent variables extra prior information is needed for identifiability. In dynamic models however the extra prior information can be made up in many more ways than in the static models analyzed in [10]. Even more important is the conclusion that under observability conditions, as this is defined in linear system theory, no extra information on the parameters that determine the means of the variables, is needed.

1/ I wish to thank P. Schönfeld and N. Christopeit of the Sonderforschungsbereich 21 at Universität Bonn and M. Deistler for the many conversations during the formative stages of this paper. All remaining errors are mine.

In part three of the paper we consider the dynamic factor analysis model with fixed exogenous variables and illustrate how independence assumptions between measurement and state variable equations alone can establish identifiability of the model parameter. The example itself should be of interest also because it has a resemblance with the models considered in the literature on Kalman-filters. In this paper however the parameters of the model are not known and are to be estimated.

I. Identifiability Conditions for ARMAX(p,q)-Model Parameters

1. The Three Forms of the ARMAX(p,q)-Model

Below write $Z_T' = (z_t)$ for the matrix with T columns, having z_t as its t-th column, t=1,..,T. Let LZ_T' be the matrix with z_t placed in the (t+1)th column, t=0,..,T-1.

The ARMAX Model under consideration is introduced in three familiar forms, the structural form (SF), the reduced form (RF) and the final form (FF). These are the G+H equations in each period t=1,..,T, and are written in standard notation

(SF) $$B_0 Y_T' + B_1 LY_T' + \Gamma X_T' = U_T', \quad U_T' = (u_t) = \Big(\sum_{k=0}^{q} \Delta_k e_{t-k}\Big), \text{ with } e_t = 0 \text{ for } t \leq 0,$$

B_i is $(G+H)\times(G+H)$ with B_0 nonsingular

Γ is $(G+H)\times K$ with K the number of exogenous variables

Δ_i is $(G+H)\times(G+H)$ with $\Delta_0 = I_{G+H}$

(e_t) is a white noise process, mean zero, covariance Σ for t=1,..,T

(1) Y_T' and X_T' are the data matrices on endogenous and exogenous variables

(RF) $$Y_T' = \Pi_1 LY_T' + \Pi_0 X_T' + V_T', \quad V_T' = (v_t) = (B_0^{-1} u_t) = \Big(\sum_{k=0}^{q} \Xi_k B_0^{-1} e_{t-k}\Big)$$

$$\Pi_1 = -B_0^{-1} B_1, \quad \Pi_0 = -B_0^{-1}\Gamma \quad \Xi_k = B_0^{-1} \Delta_k B_0 \quad \Omega = B_0^{-1} \Sigma B_0'^{-1} = E(v_1 v_1')$$

(FF) $$Y_T' = M_T' + W_T'$$

(2) $$M_T' = (\mu_t) = \Big(\Pi_1^t \mu_0 + \sum_{s=0}^{t-1} \Pi_1^s \Pi_0 x_{t-s}\Big) = (\Pi_1 \mu_{t-1} + \Pi_0 x_t) \quad \text{given } \mu_0$$

(3) $$W_T' = (w_t) = \Big(\sum_{s=0}^{t-1} \Pi_1^s v_{t-s}\Big) = \Big(\sum_{s=0}^{t-1} \Pi_1^s \sum_{k=0}^{q} \Xi_k B_0^{-1} e_{t-s-k}\Big)$$

The final form expresses the endogenous variables as the sum of the means M_T' and the final form residuals W_T'. Each component w_t is a distributed lag of the sequence of residuals $B_0^{-1} e_t$, which is a white noise sequence with covariance Ω.

We show the relations (3) in more detail by rewriting them in the form

$$
(4)\quad \left[\begin{array}{c} w_1 \\ w_2 \\ \hline w_{q+1} \\ w_{q+2} \\ \hline w_T \end{array}\right] = \left[\begin{array}{cc|cc|c} I_{G+H} & 0 & 0 & 0 & \\ \Pi_1+\Xi_1 & I_{G+H} & 0 & 0 & \\ \hline \sum_{s=0}^{q}\Pi_1^s\Xi_{q-s} & \sum_{s=0}^{q-1}\Pi_1^s\Xi_{q-1-s} & I_{G+H} & 0 & \\ \Pi_1\sum_{s=0}^{q}\Pi_1^s\Xi_{q-s} & \sum_{s=0}^{q}\Pi_1^s\Xi_{q-s} & \Pi_1+\Xi_1 & I_{G+H} & \\ \hline \Pi_1^{T-q-1}\sum_{s=0}^{q}\Pi_1^s\Xi_{q-s} & & & & I_{G+H} \end{array}\right] \left[\begin{array}{c} B_0^{-1}e_1 \\ B_0^{-1}e_2 \\ \hline B_0^{-1}e_{q+1} \\ B_o^{-1}e_{q+2} \\ \hline B_0^{-1}e_T \end{array}\right]
$$

$$= D\,(I_T \otimes B_0^{-1})\ \mathrm{vec}E_T,$$

where $\mathrm{vec}E_T$ is the row of rows of E_T written as a column and $E_T' = (e_t)$. 2/

From (4) we verify that w_t is a moving average of $B_0^{-1}e_t$ with the first q+1 weights determined by the moving average weights in the reduced form residuals as well as by the reduced form matrix Π_1. Beyond q+1 the weights are geometrically changing with the matrix Π_1. An important property of the matrix of weights D is that it is recursive and the lag coefficients only depend on the lag and not on the time t.

Letting $D_{i,j}$ be the (i,j)th submatrix of D, from (4) we verify that the covariance matrix C(T) satisfies

$$(5a)\qquad C(T) = (E(w_t w_{t'}')) = D(I_T \otimes \Omega)D', \qquad t,t'=1,..,T,$$

$$(5b)\qquad E(w_t w_1') = D_{t,1}\Omega = \Big(\sum_{s=0}^{t-1}\Pi_1^s\Xi_{t-1-s}\Big)\Omega = \Xi_{t-1}\Omega + \Pi_1 E(w_{t-1}w_1'), \qquad t \le q+1$$

$$(5)\qquad\qquad = D_{t,1}\Omega = \Pi_1^{t-q-1}\,E(w_{q+1}w_1') \qquad\qquad t > q+1$$

$$(5c)\qquad E(w_{t+1}w_{t'+1}') = E(w_t w_{t'}') + D_{t+1,1}\,\Omega\, D_{t'+1,1}' \qquad t,t'=1,..,T-1.$$

2/ $(\mathrm{vec}Z)'$ is the row of rows of a matrix Z and $(\mathrm{vec}^*Z)'$ is the row of rows of a symmetric matrix omitting the elements below the diagonal. The matrices Q_{G+H} and P_{G+H}' are matrices of zeroes, ones and halves satisfying $\mathrm{vec}^*Z=Q_{G+H}\mathrm{vec}Z$ and $\mathrm{vec}Z= P_{G+H}'\mathrm{vec}^*Z$, where Z is symmetric of order G+H. The operator vec has the property that $\mathrm{vec}Z_1Z_2Z_3=(Z_1\otimes Z_3')\mathrm{vec}Z_2$.

The recursive relation (5c) implies that if we were to start the dynamic stochastic process one period later, i.e. in period 2 given $w_t=0$, $t \leq 1$, instead of in period 1 given $w_t=0$, $t \leq 0$, the conditional covariance structure is exactly the same as the original covariance structure, with one period delay. Property (5c) follows from (5a) and the recursive nature of D. Conversely if a covariance satisfies (5c) it can be shown that there exist a recursive matrix D such that (5a) holds.

The properties (5c) and (5b) imply further that all the covariance matrices in C(T) can be expressed in terms of the covariances $E(w_t w_1')$, $t=1,..,q+1$, for given Π_1. It follows that the information contained in C(T) is the same as the information contained in $E(w_t w_1')$, $t=1,..,q+2$, including the information concerning Π_1. Properties (5c) and (5b) are the basis for determining a minimal number of external covariance characteristics.

If the true model is an ARMAX(p,q)-Model with a p-th order autoregression in G^* endogenous variables

$$(6)\quad b_0 Y_T^{*\prime} + b_1\, L Y_T^{*\prime} + \ldots + b_p\, L^p Y_T^{*\prime} + \gamma X_T' = U_T^{*\prime}, \quad U_T^{*\prime} = (u_t^*) = \Big(\sum_{k=0}^{q} \Delta_k^* e_{t-k}^*\Big), \quad \Delta_0^* = I_{G^*},$$

we can always write this in the form of (1) by letting $G+H=pG^*$ and defining

$$(7)\quad Y_T' = \begin{bmatrix} Y_T^{*\prime} \\ LY_T^{*\prime} \\ L^{p-1}Y_T^{*\prime} \end{bmatrix},\ B_0 = \begin{bmatrix} b_0 & b_1 & b_{p-1} \\ 0 & I_{G^*} & 0 \\ 0 & 0 & I_{G^*} \end{bmatrix},\ B_1 = \begin{bmatrix} 0 & 0 & b_p \\ -I_{G^*} & 0 & 0 \\ 0 & -I_{G^*} & 0 \end{bmatrix},\ \Gamma = \begin{bmatrix} \gamma \\ 0 \\ 0 \end{bmatrix},\ \Delta_k = \begin{bmatrix} \Delta_k^* & 0 & 0 \\ 0 & I_{G^*} & 0 \\ 0 & 0 & I_{G^*} \end{bmatrix},\ e_t = \begin{bmatrix} e_t^* \\ 0 \\ 0 \end{bmatrix}$$

Fisher [4] considers a first-order difference equation but nothing is known about the error structure. This is formally equivalent to an ARMAX(1,T-1)-Model in which the final form residuals w_t are moving averages of the white noise process $(B_0^{-1} e_t)$ as in (4) and satisfying (5c). In particular w_T would depend on Π_1 and on T-1 arbitrary weighting matrices Ξ_k. In Hatanaka [7] both p and K are unknown and q=T-1.

2. The Identification Concept

The structural parameter α is the vector of parameters

(8) $$\alpha' = (\beta_0', \beta_1', \gamma', \delta_1', \ldots, \delta_q', \sigma')$$

where

$$\beta_0 = \text{vec}B_0, \quad \beta_1 = \text{vec}B_1, \quad \gamma = \text{vec}\Gamma, \quad \delta_k = \text{vec}\Delta_k, \quad \sigma = \text{vec}^*\Sigma \quad .$$

The reduced form parameter θ is the vector of parameters

(9) $$\theta' = (\pi_1', \pi_0', \xi_1', \ldots, \xi_q', \omega)$$

where

$$\pi_1 = \text{vec}\Pi_1, \quad \pi_0 = \text{vec}\Pi_0, \quad \xi_k = \text{vec}\Xi_k, \quad \omega = \text{vec}^*\Omega \ .$$

By definition and B_o nonsingular the relation between structural and reduced form parameters

(10) $$\alpha' = (\beta_0', -\pi_1'(B_0' \otimes I_{G+H}), -\pi_0'(B_0' \otimes I_K), \xi_1'(B_0' \otimes B_0^{-1}), \ldots, \xi_q'(B_0' \otimes B_0^{-1}), \omega' P_{G+H}(B_0' \otimes B_0')Q_{G+H}'$$

is one-to-one.

The structural parameter restrictions are a list of k_r equations $\phi(\alpha)=0$, where $\phi = (\phi_i)$, $i=1,\ldots,k_r$, is continuously differentiable.

We adopt the following definitions that are modifications of concepts in [9] to reflect our interest in reducing external information requirements.

Definition 1.

The structure $s=\{\alpha, \phi(\alpha)=0, X_T', U_T'\}$ with moments M_T', $C(T)$ and the structure $s^*=\{\alpha^*, \phi(\alpha^*)=0, X_T', U_T^{*\prime}\}$ with moments $M_T^{*\prime}$, $C^*(T)$ are (T_1,T_2)-observationally equivalent if $M_{T_1}' = M_{T_1}^{*\prime}$ and $C_{T_2} = C_{T_2}^*$, where $T_1 \leq T$, $T_2 \leq T$.

Definition 2

The parameter α of the true structure s is (T_1,T_2)-identifiable whenever s^* (T_1,T_2)-observationally equivalent to s, implies $\alpha^*=\alpha$.

Definition 3

The pair (T_1,T_2) is observationally efficient if α (T_1^*,T_2^*)-identifiable implies α is (T_1,T_2)-identifiable, where $T_1^* \geq T_1$, $T_2^* \geq T_2$ with strict inequality at least once.

In general one would not expect the observationally efficient pairs (T_1,T_2) to be unique. Trade-offs between T_1 and T_2 can exist.

Remark 1

In Definition 1 the external information is T_1 first moments $M_{T_1}'=(E(y_t))$, $t=1,..,T_1$, and the matrix $C(T_2)=(E(y_t-E(y_t))(y_{t'}-E(y_{t'}))')$, $t,t'=1,..,T_2$. In contrast in the classical papers by Hannan [5], [6], also Deistler [1], the identification concept is asymptotic identifiability, the external information being an infinite number of first and second moments. These enter in their analysis through assuming that the spectral density matrix functions or the z-transformed transfer functions $(I_{G+H}-\Pi_1 z)^{-1}\,\Gamma$ and $(I_{G+H}-\Pi_1 z)^{-1}\sum_{k=0}^{q}\Xi_k z^k$ are given. This in turn pre-supposes stability requirements and restrictions that either the historical values of the exogenous variables are known outside the sample period or it is known that they are stationary processes.

Remark 2

From the relation (10) between structural and reduced form parameter, the structural parameter α is identifiable if and only if the parameter$((\mathrm{vec}B_0)',\theta')'$ with B_0 nonsingular, is identifiable. In static models with intertemporally uncorrelated residuals the reduced form is in final form already and the study of identification proceeds by finding conditions under which the structural parameter α is uniquely determined given the reduced form parameter θ. The latter is assumed supplied by the external information. This exercise produces the classical conditions for identification in static models.

Likewise we can perform this same exercise for the ARMAX(p,q)-Model and state the conditions under which α is locally uniquely determined given θ. These conditions are that the matrix

$$(11)\quad \Psi = \phi_{\beta_0'}(I_{G+H}\otimes B_0')+\phi_{\beta_1'}(I_{G+H}\otimes B_1')+\phi_{\gamma'}(I_{G+H}\otimes\Gamma')+\sum_{k=1}^{q}\phi_{\delta_k'}[(I_{G+H}\otimes\Delta_k')-(\Delta_k\otimes I_{G+H})]$$

$$+\ 2\phi_{\sigma'}Q_{G+H}(I_{G+H}\otimes\Sigma)$$

have rank $(G+H)^2$, where $\Psi\ (I_{G+H}\otimes B_0'^{-1})$ is the Jacobian matrix of k_r restrictions

$$(12)\quad 0=\phi(\alpha')=\phi\,((\mathrm{vec}B_0)',-(\mathrm{vec}B_o\Pi_1)',-(\mathrm{vec}B_0\Pi_0)',..,(\mathrm{vec}B_o\Xi_k B_o^{-1})',..,(\mathrm{vec}^*B_0\Omega B_0')')$$

in the $(G+H)^2$ unknowns $\beta_0=\mathrm{vec}B_0$ for given reduced form parameters and where

$$\phi_{\beta_0'}=\frac{\partial\phi(\alpha)}{\partial(\mathrm{vec}B_0)'}\qquad \phi_{\beta_1'}=\frac{\partial\phi(\alpha)}{\partial(\mathrm{vec}B_1)'}\qquad \phi_{\gamma'}=\frac{\partial\phi(\alpha)}{\partial(\mathrm{vec}\Gamma)'}$$

$$\phi_{\delta_k'}=\frac{\partial\phi(\alpha)}{\partial(\mathrm{vec}\Delta_k)'}\qquad \phi_{\sigma'}=\frac{\partial\phi(\alpha)}{\partial(\mathrm{vec}^*\Sigma)'}\ .$$

With linear ϕ, independent of Δ_k and of Σ, this condition is also globally necessary and sufficient since (12) is linear in the unknowns $\mathrm{vec}B_0$.

Many results in the literature on the identification of dynamic systems are statements of conditions under which $\rho(\Psi)=(G+H)^2$ **is** the only condition for identification. Taking together however the conditions for identification of the parameter of dynamic systems must guarantee that $\rho(\Psi)=(G+H)^2$ as well as that the reduced form parameter θ itself is identifiable given the external information about the final form moments. This is what makes the study of identification of dynamic model parameters different from that of static model parameters.

3. Proposition 1.

A. The structural parameter α is locally (T_1,T_2)-identifiable if and for constant rank matrices only if the matrix

(13)

Information \ Parameter	β_0'	π_1'	π_0'	ξ_1'	ξ_q'	ω'
Means	0	$I_{G+H} \otimes LM_{T_1}$	$I_{G+H} \otimes X_{T_1}$	0	0	0
$E(w_1 w_1')$	0	0	0	0	0	$I_{(G+H)(G+H+1)/2}$
$E(w_2 w_1')$	0	$I_{G+H} \otimes E(w_1 w_1')$	0	$I_{G+H} \otimes E(w_1 w_1')$	0	0
- - -	- -	- - - - - - -	- - - - -	- - - - - -	- - - - - -	- - - - - - - -
$E(w_{q+1} w_1')$	0	$I_{G+H} \otimes E(w_1 w_q')$	0	0	$I_{G+H} \otimes E(w_1 w_1')$	0
- - -	- -	- - - - - - -	- - - - -	- - - - - -	- - - - - -	- - - - - - - -
$E(w_{T_2} w_1')$	0	$I_{G+H} \otimes E(w_1 w_{T_2-1}')$	0	0	0	0
Prior	Ψ	$-\phi_{\beta_1'}(B_0 \otimes I_{G+H})$	$-\phi_{\gamma'}(B_0 \otimes I_K)$	$\phi_{\delta_1'}(B_0 \otimes B_0'^{-1})$	$\phi_{\delta_q'}(B_0 \otimes B_0'^{-1})$	$\phi_{\sigma'} Q_{G+H}(B_0 \otimes B_0) P_{G+H}'$

has rank equal to the number of structural parameters. With linear prior information independent of Δ_k, $k=1,..,q$, and of Σ, this rank condition is necessary and sufficient for global identifiability.

B. If α is (T_1,T_2)-identifiable and the pair (T_1,T_2) is observationally efficient, $T_2 \leq q+2$.

Proof. By definition the structural parameter α is (T_1,T_2)-identifiable if the system (2),(5a) and (12) for given M_{T_1}', $C(T_2)$ and ϕ has a unique solution in $\zeta' = ((\text{vec}B_0)', \theta')$.

The system (2) $\mu_t=\Pi_1\mu_{t-1}+\Pi_0 x_t$ is the linear equations system

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ -- \\ \mu_{T_1} \end{bmatrix} = \begin{bmatrix} I_{G+H}\otimes\mu_0' & I_{G+H}\otimes x_1' \\ I_{G+H}\otimes\mu_1' & I_{G+H}\otimes x_2' \\ ------ & ------ \\ I_{G+H}\otimes\mu_{T_1-1}' & I_{G+H}\otimes x_{T_1}' \end{bmatrix} \begin{bmatrix} \mathrm{vec}\Pi_1 \\ \mathrm{vec}\Pi_0 \end{bmatrix}$$

for given first moments (μ_t), $t=0,..,T_1$. After re-ordering of the rows, the matrix of this linear system is reproduced under the first $(G+H)T_1$ rows of (13).

As discussed above the matrix $C(T_2)$ depends on the matrices $E(w_t w_1')$, $t=1,..,T_2$ only. From (5b) $E(w_t w_1')=\Xi_{t-1}\Omega + \Pi_1 E(w_{t-1}w_1')$ for $t \leq q+1$ and $E(w_t w_1')= \Pi_1^{t-q-1}\, E(w_{q+1}w_1')$ for $t>q+1$, where $E(w_1 w_1')= \Omega$. The rows in (13) corresponding to the external covariance matrices are **successively**

$$\frac{\partial \mathrm{vec}^*\,\Omega}{\partial\zeta'}$$

$$\frac{\partial \mathrm{vec}E(w_t w_1')}{\partial\zeta'} - (\Pi_1\otimes I_{G+H})\ \frac{\partial \mathrm{vec}E(w_{t-1}w_1')}{\partial\zeta'} - (\Xi_{t-1}\otimes I_{G+H})P_{G+H}'\,\frac{\partial \mathrm{vec}^*\Omega}{\partial\zeta'}$$

for $t=2,..,T_2$, where we put $\Xi_t=0$ for $t>q$.

The last row are the partial derivatives of the k_r prior restrictions $\phi(\alpha)=0$ with respect to ζ', through the relation (10), and where Ψ is defined at (11).

To show B. it suffices to verify that the contribution of external information on $E(w_{T_2} w_1')$ for $T_2>q+2$ is proportional to the contribution from $E(w_{q+2}w_1')$ and does not add to the rank of the matrix (13).

Discussion

1. As discussed above, we get back $\rho(\Psi)=(G+H)^2$ as part of the rank conditions for identifiability. This is to render inadmissible transformations of the structural equations (1). In comparison to static models, the matrix Ψ now includes restrictions involving the autoregressive matrix B_1 and the moving average matrices Δ_k as well as the usual parameter restrictions of the static model.

2. The reduced form parameter matrix Π_1 can be identified through external information on the means and also on the covariances as well as through prior restrictions on $\mathbf{B}_1$ that are not used to meet the condition before.

3. A failure of the data matrix X'_{T_1} to have rank K implies that the reduced form parameter Π_0 is not identifiable on basis of the external information. This can be remedied by extra restrictions on the matrix Γ of coefficients of the exogenous variables, which is indeed the remedy against perfect multicollinearity.

4. The moving average matrix Ξ_k is identifiable from the external information $E(w_{k+1}w_1')$ provided $E(w_1w_1')=\Omega$ is nonsingular. If Ω is singular, prior information is needed to identify Ξ_k. If the source of singularity of Ω is the singularity of Σ due to the presence of identities in the structural model (1), the corresponding submatrix in Δ_k is a submatrix of I_{G+H} and enough prior restrictions are available to identify Ξ_k in such cases. If the source of singularity of Ω does not lead to justifiable prior restrictions on Δ_k, then Ξ_k is not identified.

5. The parameter Ω is always identifiable since it is given in the external information matrix $E(w_1w_1')$.

The leading question dealt with in the literature on the identification of dynamic model parameters is what are the conditions under which the structural parameter is identifiable under the sole condition $\rho(\Psi)=(G+H)^2$, and without having to use extra (over-identifying) prior restrictions to identify the reduced form parameter θ, except perhaps in the presence of identities, the trivial restrictions that make up for the deficiency in the rank of Ω. Assigning the external information on the covariances $E(w_k w_1')$, $k=1,..,q+1$, as well as the trivial restrictions associated with identities, to Ω and to Ξ_k, we can assign the externally given $E(w_{q+2} w_1')$ to Π_1. We have from Proposition 1 :

Corollary 1

The structural parameter α is locally $(T_1,q+2)$-identifiable without the help of extra (over-identifying) (B_1,Γ)-restrictions only if condition

$$\text{(C1)} \qquad G+H+K = \rho \begin{pmatrix} LM_{T_1}' & E(w_{q+1} w_1') \\ X_{T_1}' & 0 \end{pmatrix}$$

holds.

Unfortunately the condition (C1) is not easily verifiable.

Let the p-th order autoregression with an r-th order moving average in the K^* exogenous variables $Z_T = (z_t)$ be written as

$$(14) \qquad \sum_{k=0}^{p} b_k L^k Y_T^{*\prime} + \sum_{k=0}^{r} \gamma_k L^k Z_T' = U_T^{*\prime}, \quad U_T^{*\prime} = (u_t^*) = \left(\sum_{k=0}^{q} \Delta_k^* e_{t-k}^* \right) .$$

In addition to the variables defined at (7), let

$$(15) \quad X_T' = \begin{vmatrix} Z_T' \\ LZ_T' \\ \\ L^r Z_T' \end{vmatrix} \qquad \Gamma = \begin{vmatrix} \gamma_0 & \gamma_1 & \gamma_r \\ 0 & 0 & 0 \\ \\ 0 & 0 & 0 \end{vmatrix} = ((\Gamma)_0 \quad (\Gamma)_1 \qquad (\Gamma)_r).$$

Lemma 1.

In terms of the parameters of the structural model (1) a necessary condition for the condition (C1) to hold is that $\rho(B_1,\Gamma,\Delta_q)=G+H$ and $\rho(X_{T_1}')=K$. In terms of the parameters of Model (14) a necessary condition for the condition (C1) to hold is that $\rho(b_p,\Delta_q^*, \gamma_r)=G^*$ and $\rho(Z_{T_1})=K^*$, where G^* and K^* is the number of unlagged endogenous variables and exogenous variables respectively.

Proof. From $\mu_t=\Pi_1\mu_{t-1}+\Pi_0 x_t$ and from $E(w_{q+1}w_1')=(\sum_{s=0}^{q}\Pi_1^s\Xi_{q-s})\Omega$, we have

$$(16)\qquad \rho\begin{bmatrix} LM_{T_1}' & E(w_{q+1}w_1') \\ X_{T_1}' & 0 \end{bmatrix} = \rho\begin{bmatrix} \Pi_1 & \Pi_0 & 0 & \sum_{s=2}^{q}\Pi_1^s\Xi_{q-s} & \Xi_q \\ 0 & 0 & I_K & 0 & 0 \end{bmatrix}\begin{bmatrix} L^2M_{T_1}' & \Xi_{q-1}\Omega \\ LX_{T_1}' & 0 \\ X_{T_1}' & 0 \\ 0 & \Omega \\ 0 & \Omega \end{bmatrix}$$

$$\leq \rho(\Pi_1,\ \Xi_q,\ \Pi_0)+\rho(X_{T_1}')=\rho(B_1,\Gamma,\Delta_q)+\rho(X_{T_1}').$$

For the Model (14) X_{T_1}' is the matrix of K^* variables Z_{T_1}' and its r lagged values so that the matrices X_{T_1}' and LX_{T_1}' have $K^*(r-1)$ overlapping components. After elimination of these we have that (16) equals

$$\rho\begin{bmatrix} \Pi_1 & (\Pi_0)_0 & (\Pi_0)_1 & (\Pi_0)_r & 0 & \sum_{s=2}^{q}\Pi_1^s\Xi_{q-s} & \Xi_q \\ 0 & 0 & 0 & 0 & I_{K^*} & 0 & 0 \\ 0 & I_{K^*} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{K^*} & 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} L^2M_{T_1}' & \Xi_{q-1}\Omega \\ LZ_{T_1}' & 0 \\ L^2Z_{T_1}' & 0 \\ L^{r+1}Z_{T_1}' & 0 \\ Z_{T_1}' & 0 \\ 0 & \Omega \\ 0 & \Omega \end{bmatrix}$$

$$\leq \rho(\Pi_1,\ (\Pi_0)_r,\ \Xi_q)+rK^*+\rho(Z_{T_1}')=\rho(B_1,\ (\Gamma)_r,\Delta_q)+rK^*+\rho(Z_{T_1}'),$$

where $((\Pi_0)_k)$ is the partition of Π_0 corresponding to the partition of Γ. By definitions (7) the result follows, noting that $G+H=pG^*$ and $K=(r+1)K^*$.

The conditions under which it is possible to identify the structural parameter without over-identifying prior restrictions is thus seen to be limited to models with parameter matrices of highest order meeting a full row rank condition.

Aspecial case discussed in the literature is Model (14) with q=T-1, i.e. when the order of the error moving average process is not known. Covariance external information does not help in identifying Π_1, but Proposition 1 surprisingly specializes to:

Corollary 2.

The structural parameter of the model (14) satisfying $\rho(X_T')=K=(r+1)K^*$, $q=T-1$ and G^* endogenous variables is locally (T,T)-identifiable if and for constant rank matrices only if the matrix

$$(17)\qquad \left[\; \frac{\partial\phi}{\partial(\mathrm{vec}A)'}(I_{G^*}\otimes A') \qquad \frac{\partial\phi}{\partial(\mathrm{vec}(b_1,\ldots,b_p))'}(b_0\otimes I_{pG^*}) - \frac{\partial\phi}{\partial(\mathrm{vec}(\gamma_0,\ldots,\gamma_r))'}(b_0\otimes R) \;\right]$$

has rank $(p+1)G^{*2}$, where $A=(b_0,\ldots,b_p,\gamma_0,\ldots,\gamma_r)$ and $R = (X_T'X_T)^{-1}X_T'LM_T$. With linear restrictions this rank condition is necessary and sufficient for global identifiability. This rank condition is sufficient and for constant rank matrices necessary for the system $\phi(\mathrm{vec}A)=0$, $E(U_T^{*\prime}X_T)=0$ to have a locally unique solution in vecA.

Proof. When $q=T-1$ only the external information on the means helps to identify Π_1. The rank condition on the matrix (17) follows after elimination of the restrictions on (B_0,B_1,Γ) that are implied by the definitions (7) and (15) and after pre-multiplying the rows corresponding to the external information on the means by $(I_{pG^*}\otimes X_T')$. The second part of the Corollary follows from considering the Jacobian matrix of the system of equations $\phi(\mathrm{vec}A)=0$, $E(U_T^{*\prime}X_T)=0$.

Corollary 2 is an illustration of how in Proposition 1 the conditions for identifiability and estimability are closely linked together. The estimator that emerges is very standard.

We now discuss how our Proposition 1 contains the results from the literature.

Literature References

1. As discussed above, Fisher's Model in [4] is formally equivalent to an ARMAX(1,T-1)-Model within the concept of (T,T)-identifiability. In our terminology, Fisher derives the result that the condition

$$\text{(C2)} \qquad G+H+K = \rho\begin{pmatrix} LM_T' \\ X_T' \end{pmatrix}$$

is necessary and sufficient for the classical rank condition $\rho(\Psi)=(G+H)^2$ to become the sole condition for identifiability. As stated in Corollary 1 above the necessity of (C2) follows only if there are no over-identifying (B_1,Γ)-restrictions. If there are more than $(G+H)^2$ prior parameter restrictions, it is possible to identify α even when (C2) fails.

2. Hannan [6], Deistler [1], Koch [8] and others have considered the ARMAX(p,q)-Model from the perspective of asymptotic identifiability which is equivalent to (T,T)-identifiability, when $T\to\infty$. Often the exogenous variables are assumed generated by a stationary stochastic process and identifiability is to be understood as identifiability within the class of stochastically generated exogenous variables. Within the definitions of the pth order autoregression (14), the external information assumed is $(b(z))^{-1}\gamma(z)$ and $(b(z))^{-1}\delta(z)$, where

$$b(z)=\sum_{s=0}^{p} b_s z^s \qquad \gamma(z) = \sum_{s=0}^{r} \gamma_s z^s \qquad \delta(z) = \sum_{s=0}^{q} \Delta_s^* z^s$$

are the z-transforms of the autoregressive and moving average processes. In analyzing the leading question discussed in our Corollary 1, the authors found it necessary to introduce the following conditions:

(C3) $\rho(b_p, \gamma_r, \Delta_q^*) = G^*$ (Rank Condition on Highest Order Matrices)

(C4) $b(z)$ and $\gamma(z)$ have no roots in common (Minimality Condition for Means)

(C5) $b(z)$ and $\delta(z)$ have no roots in common (Minimality Condition for Covariances)

Conditions (C4) and (C5) play a crucial role in the method of analysis followed in Hannan [6] and others. Failure of (C4) implies that the final form transfer function $(b(z))^{-1}\gamma(z) = (b^a(z))^{-1}\gamma^a(z)$, where $b^a(z)$ and $\gamma^a(z)$ are of degree p-1 and r-1 respectively. If this is the case the final form asymptotic means could be perceived as being generated by different ARMAX-Models of degrees (p-1,q-1) and (p,q). A minimality condition would assume that the true model corresponds to the lowest possible orders after all common roots are removed. A similar remark applies to justify (C5).

Hannan [5] found the justification of (C3) from requiring that transformations of model (1) of the type

$$(18) \qquad (F_0+F_1L)(B_0+B_1L)Y_T' + (F_0+F_1L)\sum_{k=0}^{r}(\Gamma)_k\, L^k Z_T' = (F_0+F_1L)\sum_{k=0}^{q}\Delta_k\, L^k E_T' ,$$

with matrices F_0, F_1 of order G+H and $F_1 \neq 0$, be ruled out by the prior information. This requirement is that the coefficients of L^2Y_T', $L^{r+1}Z_T'$ and $L^{q+1}E_T'$ i.e. $F_1(B_1, (\Gamma)_r, \Delta_k)$ be all zeroes, for some $F_1 \neq 0$. Under condition (C3) this is impossible.

Without further discussion of the role played by the conditions in Hannan and in other references, we will state here how the failure of the above conditions imply a failure of our condition (C1):

Lemma 2.

In terms of the parameters of Model (14) we have $\rho\begin{pmatrix} LM_T' & E(w_{q+1}w_1') \\ X_T' & 0 \end{pmatrix}$

$\leq$ G+H+K-G* + $\rho(b_p, \gamma_r, \Delta_q^*)$

$\leq$ G+H+K-G* + $\rho((E(w_{q+1}w_1'))_1 - ((\Pi_1^a \;\; 0)E(w_q w_1'))_1)$ if (C4) fails and I_{G^*}+az is a common factor in b(z) and $\gamma(z)$, with $(Z)_1$ the first G* rows of :

$\leq$ G+H+K-G* if (C4) and (C5) fail and I+az is a common factor in all three b(z), $\gamma(z)$ and $\delta(z)$, where Π_1^a is the reduced form parameter matrix when the Model (14) is a (p-1)-th order autoregression with parameter matrices $(b_0^a, b_1^a, \ldots, b_{p-1}^a)$ satisfying $b(z)=(I_{G*}+az)b^a(z)$.

Proof. The first inequality of Lemma 2 is a restatement of Lemma 1 for the Model (14). To show the second inequality, following the definitions at (7) and the algebra in Lemma 1, write $\begin{pmatrix} LM_T' & E(w_{q+1}w_1') \\ X_T' & 0 \end{pmatrix} = R_1(p,r,b,\gamma)R_2(p,r)$, where

$$R_1(p,r,b,\gamma) = \begin{pmatrix} -b_0^{-1}b_1 & -b_0^{-1}b_2 & -b_0^{-1}b_p & -b_o^{-1}\gamma_0 & -b_0^{-1}\gamma_1 & -b_0^{-1}\gamma_r & 0 & \\ I_{G*} & 0 & 0 & 0 & 0 & 0 & 0 & E(w_{q+1}w_1') \\ 0 & I_{G*} & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & 0 & I_{K*} & 0 \\ 0 & 0 & 0 & I_{K*} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{K*} & 0 & 0 & 0 \end{pmatrix}$$

$$(R_2(p,r))' = \begin{pmatrix} L^2M_T^* & L^3M_T^* & L^{p+1}M_T^* & LZ_T & L^2Z_T & L^{r+1}Z_T & Z_T & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_{pG*} \end{pmatrix}$$

and $M_T^*=E(Y_T^*)$ is the T×G* matrix of expected values of Y_T^*.

If (C4) fails and I+az is a common factor in b(z) and $\gamma(z)$, with $b(z)=(I+az)b^a(z)$ and $\gamma(z)=(I+az)\gamma^a(z)$ identically in z, it must also be true that $\begin{pmatrix} LM_T' & E(w_{q+1}w_1') \\ X_T' & 0 \end{pmatrix} = R_1(p-1,r-1,b^a,\gamma^a)R_2(p-1,r-1)$. Lagging the last relation by one period and solving the first G* equations for $ab_{p-1}^a L^{p+1}M_T^{*\prime} + a\gamma_{r-1}^a L^{r+1}Z_T'$, eliminate $b_o^{-1}b_p\, L^{p+1}M_T^{*\prime} + b_0^{-1}\gamma_r\, L^{r+1}Z_T'$ from the product above, using $b_p=ab_{p-1}^a$, $\gamma_r=a\gamma_{r-1}^a$. The result is

$$\begin{pmatrix} LM_T' & E(w_{q+1}w_1') \\ X_T' & 0 \end{pmatrix} = R_3\, R_2(p-1,r-1)$$

where

$$R_3 = \begin{bmatrix} -b_0^{-1}b_1^a & -b_0^{-1}b_2^a & -b_0^{-1}b_{p-1}^a & -b_0^{-1}\gamma_0^a & -b_0^{-1}\gamma_1^a & -b_0^{-1}\gamma_{r-1}^a & 0 & \\ I_{G*} & 0 & 0 & 0 & 0 & 0 & 0 & E(w_{q+1}w_1') \\ 0 & I_{G*} & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & I_{G*} & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & 0 & I_{K*} & 0 \\ 0 & 0 & 0 & I_{K*} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{K*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{K*} & 0 & 0 \end{bmatrix}$$

The second rank inequality of the Lemma now follows by eliminating the $(p-1)G^*$ rows that follow the first G^* rows of R_3, remembering that the first $(p-1)G^*$ rows of $E(w_q w_1')$ are identical to the last $(p-1)G^*$ rows of $E(w_{q+1}w_1')$.

The third inequality follows from (5b) since if $I+az$ is a common factor in all three, the moving average error process is a $(q-1)$-th order process with $\Xi_q^a=0$.

In general for finite (p,q) a common factor in all three $b(z)$, $\gamma(z)$ and $\delta(z)$ results in a failure of our condition (C1). If $q=T-1$, then asymptotically a common factor in $b(z)$ and $\gamma(z)$ alone makes condition (C1) i.e. Fisher's condition (C2) fail. This is the main difference between (C4) and (C5). But for finite (p,q) a failure of (C4) and (C5) with different common factors is not shown to result in a failure of (C1).

Practically speaking the rank condition on the highest order matrices (C3) is the more useful condition and failure of this necessarily makes (C1) fail also.

3. The last comparison is with Hatanaka [7]. His Model is (14) with given G^* and K^* but unknown p, r=p and q=T-1 i.e. nothing is known about the stationary error structure and nothing is known about the order of the stable autoregression. External information is all asymptotic means and covariances and X_T' has full row rank. Hatanaka shows that the parameter is identifiable under a system of exclusion and normalization restrictions that prescribes the coefficients of at least G^* components in each row of (b_k, γ_k), these prescribed coefficients standing in the same columns of (b_k, γ_k), $k=0,1,2,\ldots$.

Therefore this is a system of $(p+1)G^{*2}$ restrictions when the order of the autoregression is p. This result is formally contained in our Corollary 2 to Proposition 1 and the rank condition of (17) stated there should hold for every integer p. It is also obvious that many alternative $(p+1)G^{*2}$ exclusion and normalization restrictions have the property of satisfying the rank condition stated at (17) and Hatanaka's case is special.

We end this survey of the literature with a reminder that Proposition 1 holds independently of any minimality conditions or without conditions on the rank of the parameter matrices of highest order.

We now turn to the problem of latent variables.

II. Identifiability Conditions with Latent Variables

In this part of the paper we derive the necessary and sufficient conditions for the local identifiability of the structural parameter when G endogenous variables are observed and H are not. In dynamic models the missing components are missing also in the vector of lagged endogenous variables and this makes the problem more complicated than the latent variables problem in static models. An interesting intermediate case is when observations become available after one period, but this is not treated here. When the variables are generically latent as in factor analysis, the identification problem must be dealt with as follows.

Assume the model is the same as model (1). Partition the endogenous variables Y_T' and the equations in G and H components. The three forms are :

$$\text{(SF)} \quad \begin{bmatrix} (B_0)_1 & (B_0)_2 \\ (B_0)_3 & (B_0)_4 \end{bmatrix} \begin{bmatrix} (Y_T')_I \\ (Y_T')_{II} \end{bmatrix} + \begin{bmatrix} (B_1)_1 & (B_1)_2 \\ (B_1)_3 & (B_1)_4 \end{bmatrix} \begin{bmatrix} (LY_T')_I \\ (LY_T')_{II} \end{bmatrix} + \begin{bmatrix} \Gamma_I \\ \Gamma_{II} \end{bmatrix} X_T' = \begin{bmatrix} (U_T')_I \\ (U_T')_{II} \end{bmatrix},$$

$$(19) \quad \text{(RF)} \quad \begin{bmatrix} (Y_T')_I \\ (Y_T')_{II} \end{bmatrix} = \begin{bmatrix} (\Pi_1)_1 & (\Pi_1)_2 \\ (\Pi_1)_3 & (\Pi_1)_4 \end{bmatrix} \begin{bmatrix} (LY_T')_I \\ (LY_T')_{II} \end{bmatrix} + \begin{bmatrix} (\Pi_0)_I \\ (\Pi_0)_{II} \end{bmatrix} X_T' + \begin{bmatrix} (V_T')_I \\ (V_T')_{II} \end{bmatrix}, \quad (v_t) = \left(\sum_{k=0}^{q} \begin{bmatrix} (\Xi_k)_I \\ (\Xi_k)_{II} \end{bmatrix} B_0^{-1} e_{t-k} \right)$$

$$\text{(FF)} \quad \begin{bmatrix} (Y_T')_I \\ (Y_T')_{II} \end{bmatrix} = \begin{bmatrix} (M_T')_I \\ (M_T')_{II} \end{bmatrix} + \begin{bmatrix} (W_T')_I \\ (W_T')_{II} \end{bmatrix}, \quad M_T' = \begin{bmatrix} (\mu_t)_I \\ (\mu_t)_{II} \end{bmatrix}, \quad E(w_t w_{t'}') = \begin{bmatrix} (E(w_t w_{t'}'))_I \\ (E(w_t w_{t'}'))_{II} \end{bmatrix} = \begin{bmatrix} (\ )_1 & (\ )_2 \\ (\ )_3 & (\ )_4 \end{bmatrix}$$

with all symbols as defined at (1).

The initial values $(\mu_0)_{II}$ are included in the list of structural parameters and the prior restrictions ϕ may or may not depend on $(\mu_0)_{II}$. Following [10], in the presence of latent variables the definition of observationally equivalent parameters is now changed to:

Definition 4.

With latent variables the structure $s=\{\alpha,\ \phi(\alpha)=0,\ X_T',\ U_T'\}$ with moments M_T', $C(T)$, and the structure $s^*=\{\alpha^*,\ \phi(\alpha^*)=0,\ X_T',\ U_T^{*\prime}\}$ with moments $M_T^{*\prime}$, $C^*(T)$ are (T_1,T_2)-observationally equivalent if $(M_{T_1}')_I=(M_{T_1}^{*\prime})_I$ and $(E(w_t w_{t'}'))_1 = (E(w_t^* w_{t'}^{*\prime}))_1$, $t,t' = 1,..,T_2$, where $T_1 \leq T$, $T_2 \leq T$.

Only the first G components of the mean vector are given externally and the principal submatrix of the covariance matrices. The remaining definitions are as before.

Proposition 2.

A. The structural parameter of the latent variables model (19) is locally (T_1,T_2)--identifiable if and for constant rank matrices only if the matrix

(20)

$(\mu_0)_{II}'$	β_0'	$(\pi_1)_I'$	$(\pi_1)_{II}'$	$(\pi_0)_I'$	$(\pi_0)_{II}'$	ξ_i', $i=1,..,q$	ω'
Λ_0	0	$I_G \otimes \mu_0'$	0	$I_G \otimes x_1'$	0	0	0
Λ_1	0	$I_G \otimes \mu_1'$	$\Lambda_0 \otimes \mu_0'$	$I_G \otimes x_2'$	$\Lambda_0 \otimes x_1'$	0	0
Λ_2	0	$I_G \otimes \mu_2'$	$\sum_{s=0}^{1} \Lambda_s \otimes \mu_{1-s}'$	$I_G \otimes x_3'$	$\sum_{s=0}^{1} \Lambda_s \otimes x_{2-s}'$	0	0
- - -	-	- -					
Λ_{T_1-1}	0	$I_G \otimes \mu_{T_1-1}'$	$\sum_{s=0}^{T_1-2} \Lambda_s \otimes \mu_{T_1-s-2}'$	$I_G \otimes x_{T_1}'$	$\sum_{s=0}^{T_1-2} \Lambda_s \otimes x_{T_1-s-1}'$	0	0
0	0	$2Q_{T_2G} \sum_{k=0}^{q} \sum_{s=k+1}^{T_2-1} (C_s \otimes \Theta\Omega\Pi_1'^{\,s-k-1})$		0		$2Q_{T_2G}(C_i \otimes \Theta\Omega)$	$Q_{T_2G}(\Theta \otimes \Theta)P_{G+H}'$
$\phi_{(\mu_0)_{II}'}$	Ψ	$-\phi_{\beta_1'}\ (B_0 \otimes I_{G+H})$		$-\phi_{\gamma'}\ (B_0 \otimes I_{G+H})$		$\phi_{\delta_i'}(B_0 \otimes B_0'^{-1})$	$\phi_{\sigma'}Q_{G+H}(B_0 \otimes B_0)P_{G+H}'$

has rank equal to the number of structural parameters, where $\Theta= \sum_{i=0}^{q} C_i \Xi_i$, $\Lambda_s=(\Pi_1)_2((\Pi_1)_4)^{s}$
$C_i' = (\ 0 \quad 0 \quad ((\Pi_1^0)_I)' \quad ((\Pi_1^1)_I)' \quad . . . \quad ((\Pi_1^{T_2-i-1})_I)'\ )$ with $((\Pi_1^0)_I)'$ in the (i+1)th set of G columns.

B. If the pair (T_1,T_2) is observationally efficient, $T_2 \leq q+2+H$.

Proof. By definitions (2) the derivatives of $(\mu_t)_I$ with respect to the structural parameters are all zero except

$$\frac{\partial(\mu_t)_I}{\partial(\mu_0)'_{II}} = (\Pi_1^t)_2\,, \qquad \frac{\partial(\mu_t)_I}{\partial\pi_1'} = \sum_{s=0}^{t-1}((\Pi_1^s)_I \otimes \mu'_{t-s-1}, \qquad \frac{\partial(\mu_t)_I}{\partial\pi_0'} = \sum_{s=0}^{t-1}(\Pi_1^s)_I \otimes x'_{t-s}.$$

The first rows in (20) are $\frac{\partial(\mu_1)_I}{\partial\alpha'}$. The second rows are $\frac{\partial(\mu_2)_I}{\partial\alpha'}$ minus $(\Pi_1)_1$ times the first rows of (20). The third rows are $\frac{\partial(\mu_3)}{\partial\alpha'}$ minus $(\Pi_1)_1$ times the second rows of (20) and minus $(\Pi_1^2)_1$ times the first rows of (20). In general the t-th rows are equal to $\frac{\partial(\mu_t)_I}{\partial\alpha'}$ minus the preceding rows $(1),\ldots,(t-1)$, multiplied by $(\Pi_1^{t-1})_1,\ldots,(\Pi_1)_1$ respectively where we made use of the relation

$$(\Pi_1^t)_2 = \sum_{s=0}^{t-1}(\Pi_1^s)_1(\Pi_1)_2((\Pi_1)_4)^{t-1-s} = \sum_{s=0}^{t-1}(\Pi_1^s)_1\Lambda_{t-1-s}.$$

The contribution to the rank criterion coming from the external information on the covariances is stated in the second part of the matrix (20). From the relations (5b) and (5c) the matrix

$$\Theta\Omega\Theta' = \begin{pmatrix} (E(w_1w_1'))_1 & (E(w_1w_2'))_1 & \cdots & (E(w_1w_{T_2}'))_1 \\ (E(w_2w_1'))_1 & & & \\ & & (E(w_{t+1}w'_{t'+1})-E(w_tw'_{t'}))_1 & \\ (E(w_{T_2}w_1'))_1 & & t,t'=1,\ldots,T_2-1 & \end{pmatrix}$$

contains all the relevant information. The second to last rows in (20) are the derivatives $\frac{\partial \mathrm{vec}^*\,\Theta\Omega\Theta'}{\partial\alpha'}$. In particular verify that

$$\frac{\partial\mathrm{vec}^*\Theta\Omega\Theta'}{\partial\pi_1'} = 2Q_{T_2G}(I_{GT_2}\otimes\Theta\Omega)\frac{\partial\mathrm{vec}\Theta}{\partial\pi_1'} = 2Q_{T_2G}\sum_{k=0}^{q}(I_{GT_2}\otimes\Xi_k')\frac{\partial C_k}{\partial\pi_1'}$$

$$\frac{\partial C_k}{\partial\pi_1'} = \sum_{s=k+1}^{T_2-1}(C_s\otimes\Pi_1'^{\,s-k-1}).$$

The last rows in (20) is the contribution from the prior information exactly as in the model with no latent variables.

To show part B. of the Proposition, observe that the contribution from the covariances to the identifiability of Ξ_i is proportional to C_i, $i=1,..,q$, and by definition of Θ, the contribution to Ω is proportional to C_0. The contribution to Π_1 is a linear combination of the matrices $C_{q+1},\ldots, C_{T_2-1}$. But we have

$$C_t = \left|\begin{array}{cc} 0 & 0 \\ \hline I_G & 0 \\ (\Pi_1^1)_1 & (\Pi_1^1)_2 \\ \hline (\Pi_1^{T_2-t-1})_1 & (\Pi_1^{T_2-t-1})_2 \end{array}\right| = \left|\begin{array}{cccc} 0 & 0 & 0 & \\ \hline (\Pi_1^0)_1 & 0 & 0 & \\ (\Pi_1^1)_1 & (\Pi_1^0)_1 & 0 & \\ \hline (\Pi_1^{T_2-t-1})_1 & (\Pi_1^{T_2-t-2})_1 & (\Pi_1^0)_1 & \cdots \end{array}\right| \left|\begin{array}{cc} I_G & 0 \\ 0 & \Lambda_0 \\ 0 & \Lambda_1 \\ 0 & \Lambda_2 \\ \hline 0 & \Lambda_{T_2-t-2} \end{array}\right|$$

so that any linear combination of $(C_{q+1},.., C_{T_2-1})$ can not have rank more than the rank of the second matrix in the product above. Since this matrix reaches its maximal rank for $T_2-t-2=H-1$ and $t=q+1$, the Proposition follows.

Discussion

1. The unobservable initial values $(\mu_0)_{II}$ are identifiable provided the observability condition

(C6) $\quad H = \rho(\Lambda_0' \quad \Lambda_1' \; .. \; \Lambda_{T_1-1}') \leq \rho((I_H \quad (\Pi_1)_4' \; .. \; ((\Pi_1)_4')^{H-1})(\Pi_1)_2')$

holds. In the language of linear system theory the condition (C6) holds if the pair $((\Pi_1)_4', (\Pi_1)_2')$ is completely reachable or $((\Pi_1)_4, (\Pi_1)_2')$ is completely observable. It plays the same role here. When (C6) fails prior information will have to be supplied to identify $(\mu_0)_{II}$.

2. The external information on the means $(\mu_t)_I$ contributes to identify not only $(\Pi_1)_I$ but also $(\Pi_1)_{II}$, when the observability condition (C6) holds. Although we do not observe H components $(\mu_t)_{II}$ of (μ_t), the information is not lost since $(\mu_t)_I$ depends partly on $(\mu_{t-1})_{II}$ through the matrix $(\Pi_1)_2$. It also depends on $(\mu_{t-2})_{II}$ through $(\Pi_1)_2(\Pi_1)_4$ and on $(\mu_{t-i})_{II}$ through Λ_{i-1}, $i=1,..$. Indirectly $(\mu_t)_{II}$ becomes

available when the observability conditions hold. This is unlike the static model where not observing $(\Pi_0)_{II}$ implies that prior information is the only source to identify the parameter $(\Pi_0)_{II}$.

Similar remarks apply as far as the identifiability of Π_0 is concerned. Under the observability condition (C6) it is possible to identify all components of $(\mu_0)_{II}$, Π_1 and Π_0 without needing any over-identifying prior restrictions from the external information on the means alone. In the static models this is not possible and prior information related directly or indirectly to $(\Pi_0)_{II}$ is needed there.

3. The external information on the covariances adds at most $\rho(C_i)\rho(\Theta\Omega)$ to the rank requirement for Ξ_i, and at most $\rho(\Theta)(\rho(\Theta)+1)/2$ to the rank requirement for Ω. Under (C6) $\rho(C_i)=G+H$ and the contribution to each Ξ_i could be $(G+H)^2$ as required. Similarly the contribution to Ω could be equal to $(G+H)(G+H+1)/2$, as required for identifiability.

These contributions are however not independent. The rank contribution to $(\Pi_1, \Xi_1, .., \Xi_q, \Omega)$ is directly related to the rank of $(C_{q+1}, C_1, .., C_q, C_0)$ respectively. But note that $C_i - C_{i+1}\Pi_1 = \begin{bmatrix} 0 & 0 \\ I_G & 0 \\ 0 & 0 \end{bmatrix}$, implying that the rank of any pair (C_i, C_{i+1}) is limited to 2G+H. Similarly the rank of any subset of k matrices (C_i), i=1,..,k, is limited to kG+H. This means that if Π_1 is identifiable through the external information on the means, we will need H(G+H) prior restrictions on each matrix Δ_i, i=1,..,q, if there are no prior restrictions on Ω, assuming the observability condition (C6) holds and Ω is nonsingular. More restrictions will be needed if either (C6) fails or Ω is singular.

Again this differs from the static model where we always have to supply GH+H(H+1)/2 restrictions on Ω_{II} to identify Ω. Here in the dynamic model it is possible to identify Ω without direct or indirect prior restrictions on Ω_{II}. However the prior information requirements directly or indirectly on $(\Xi_i)_{II}$ are new requirements not present in the static model.

III. Dynamic Factor Analysis Model with Fixed Regressors

As an illustration we restate Proposition 2 when the structural model is

$$(21)\quad \begin{pmatrix} I_G & -\Lambda_y \\ 0 & I_H \end{pmatrix}\begin{pmatrix} y_t \\ \eta_t \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -\P \end{pmatrix}\begin{pmatrix} y_{t-1} \\ \eta_{t-1} \end{pmatrix} + \begin{pmatrix} \Gamma_I \\ \Gamma_{II} \end{pmatrix} x_t = \begin{pmatrix} (u_t)_I \\ (u_t)_{II} \end{pmatrix} = \begin{pmatrix} \varepsilon_t \\ \zeta_t \end{pmatrix} + \Delta_1 \begin{pmatrix} \varepsilon_{t-1} \\ \zeta_{t-1} \end{pmatrix},$$

with $\begin{pmatrix} \varepsilon_t \\ \zeta_t \end{pmatrix}$ a white noise process with mean zero and variance $\Sigma = \begin{pmatrix} \Sigma_{\varepsilon\varepsilon'} & 0 \\ 0 & \Sigma_{\zeta\zeta'} \end{pmatrix}$.
We will assume that $\Delta_1 = \begin{pmatrix} \Delta_{\varepsilon\varepsilon'_{-1}} & 0 \\ 0 & \Delta_{\zeta\zeta'_{-1}} \end{pmatrix}$ implying that the structural residuals $(u_t)_I$ and $(u_t)_{II}$ are not correlated.

This is a generalization of the static factor analysis model with fixed exogenous variables analyzed in [10].The model (21) relates a vector of G observed variables y_t, such as test scores, consumption items, to a vector η_t of H unobserved variables, such as ability characteristics, permanent income, the so-called factors.
The factors themselves follow a dynamic process over time and are influenced by a vector of K regressors through the coefficient matrix Γ_{II}.

Model (21) has the form of the models introduced in Kalman-filtering theory, where η_t are unobserved state variables and y_t are measurements on certain linear combinations of the components of the state variables and the control variables x_t. Model (21) is however different from the Kalman-filtering model in that all the matrices Λ_y, $\P$, Σ and Δ_1 are all unknown parameters to be estimated.

The reduced form corresponding to (21) is

$$(22)\quad \begin{pmatrix} y_t \\ \eta_t \end{pmatrix} = \Pi_1 \begin{pmatrix} y_{t-1} \\ \eta_{t-1} \end{pmatrix} + \Pi_0 x_t + v_t, \quad v_t = \sum_{k=0}^{1} \begin{pmatrix} (\Xi_k)_I \\ (\Xi_k)_{II} \end{pmatrix}\begin{pmatrix} I_G & \Lambda_y \\ 0 & I_H \end{pmatrix}\begin{pmatrix} \varepsilon_{t-k} \\ \zeta_{t-k} \end{pmatrix},$$

where the parameters are

$$\Pi_1 = \begin{pmatrix} 0 & \Lambda_y \P \\ 0 & \P \end{pmatrix}, \quad \Pi_0 = \begin{pmatrix} -\Gamma_I - \Lambda_y \Gamma_{II} \\ -\Gamma_{II} \end{pmatrix}, \quad \Xi_1 = \begin{pmatrix} \Delta_{\varepsilon\varepsilon'_{-1}} & \Delta_{\varepsilon\varepsilon'_{-1}}\Lambda_y + \Lambda_y \Delta_{\zeta\zeta'_{-1}} \\ 0 & \Delta_{\zeta\zeta'_{-1}} \end{pmatrix},$$

$$\Omega = E(v_1 v_1') = \begin{pmatrix} \Sigma_{\varepsilon\varepsilon'} + \Lambda_y \Sigma_{\zeta\zeta'} \Lambda_y & \Lambda_y \Sigma_{\zeta\zeta'} \\ \Sigma_{\zeta\zeta'} \Lambda_y' & \Sigma_{\zeta\zeta'} \end{pmatrix}.$$

Corollary 3.

The structural parameter of the dynamic factor analysis model (21) is locally (T_1,T_2)-identifiable if and for constant rank matrices only if the matrix (Ψ_1, Ψ_2) has rank equal to the number of its columns, where

$$
\Psi_1 = \begin{array}{c|cc|c|c}
-(\mathrm{vec}\Lambda_y)' & (\mathrm{vec}\ \Pi_0)' & & (\mathrm{vec}^*\Sigma_{\varepsilon\varepsilon'})' & (\mathrm{vec}^*\Sigma_{\zeta\zeta'})' \\
\hline
-I_G \otimes (\mu_0)'_{II}\P' & I_G \otimes x_1' & 0 & 0 & 0 \\
-I_G \otimes (\mu_1)'_{II}\P' & I_G \otimes x_2' & \Lambda_y \P \otimes x_1' & 0 & 0 \\
-I_G \otimes (\mu_{T_1-1})'_{II}\P' & I_G \otimes x'_{T_1} & \sum_{s=0}^{T_1-2} \Lambda_y \P^{s+1} \otimes x'_{T_1-s-1} & 0 & 0 \\
0 & 0 & 0 & Q_{T_2G}(\Theta^I \otimes \Theta^I)P'_G & Q_{T_2G}((\Theta^{II} + 2\Theta^I\Lambda_y) \otimes \Theta^{II})P'_H \\
\frac{2\partial\phi\ Q_G(I_G \otimes \Lambda_y \Sigma_{\zeta\zeta'})}{\partial(\mathrm{vec}^*\Sigma_{\varepsilon\varepsilon'})'} - \frac{\partial\phi}{\partial(\mathrm{vec}\Lambda_y)'} + \frac{\partial\phi\ (I_G \otimes \Gamma'_{II})}{\partial(\mathrm{vec}\Gamma_I)'} & -\frac{\partial\phi}{\partial(\mathrm{vec}\Gamma)'}(B_0 \otimes I_K) & & \frac{\partial\phi}{\partial(\mathrm{vec}^*\Sigma_{\varepsilon\varepsilon'})'} & \frac{\partial\phi}{\partial(\mathrm{vec}^*\Sigma_{\zeta\zeta'})'} - \frac{\partial\phi\ Q_G(\Lambda_y \otimes \Lambda_y)P'_H}{\partial(\mathrm{vec}^*\Sigma_{\varepsilon\varepsilon'})'}
\end{array}
$$

with $\Theta = (\ \Theta^I \quad \Theta^{II}) = \begin{vmatrix} I_G & 0 \\ \Delta_{\varepsilon\varepsilon'_{-1}} & \Lambda_y(\P + \Delta_{\zeta\zeta'_{-1}}) - \Delta_{\varepsilon\varepsilon'_{-1}}\Lambda_y \\ 0 & \Lambda_y \P^{T_2-2}(\P + \Delta_{\zeta\zeta'_{-1}}) \end{vmatrix}$ a $T_2G \times (G+H)$ matrix and

$\Psi_2 =$

$(\mu_0)'_{II}$	$-(\text{vec}\P)'$	$(\text{vec}\Delta_{\varepsilon\varepsilon'_{-1}})'$	$(\text{vec}\Delta_{\zeta\zeta'_{-1}})'$
$\Lambda_y \P$	$\Lambda_y \otimes (\mu_0)'_{II}$	0	0
$\Lambda_y \P^2$	$\Lambda_y \P \otimes (\mu_0)'_{II} + \Lambda_y \otimes (\mu_1)'_{II}$	0	0
$\Lambda_y \P^{T_1}$	$\sum_{s=0}^{T_1-1} \Lambda_y \P^s \otimes (\mu_{T_1-1-s})'_{II}$	0	0
0	$2\Omega_{T_2G} \sum_{k=0}^{1} \sum_{s=k+1}^{T_2-1} ((C_s^I \Lambda_y + C_s^{II}) \otimes (\Theta\Omega\Pi_1'^{\,s-k-1})_{II})$	$2\Omega_{T_2G}(C_1^I \otimes \Theta^I \Sigma_{\varepsilon\varepsilon'})$	$2\Omega_{T_2G}((C_1^I \Lambda_y + C_1^{II}) \otimes \Theta\Omega^{II})$
$\dfrac{\partial\phi}{\partial(\mu_0)'_{II}}$	$\dfrac{\partial\phi}{\partial(\text{vec}\P)'}$	$\dfrac{\partial\phi}{\partial(\text{vec}\Delta_{\varepsilon\varepsilon'_{-1}})'}$	$\dfrac{\partial\phi}{\partial(\text{vec}\Delta_{\zeta\zeta'_{-1}})'}$

and $C_s = (\, C_s^I \quad C_s^{II} \,) = \begin{vmatrix} 0 & 0 \\ I_G & 0 \\ 0 & \Lambda_y \P^{T_2-s-1} \end{vmatrix}$ with the matrix $(I_G \quad 0)$ in the (s+1)th set of G rows, is a $GT_2 \times (G+H)$ matrix.

The proof is straightforward and follows from Proposition 2 after elimination of the $(B_0, B_1, \Delta_1, \Sigma)$-restrictions that are explicitly shown in the model statement (21).

The matrices (Ψ_1, Ψ_2) contain the criteria for the identifiability of the parameters belonging to the static and to the dynamic parts of the model. In Ψ_1 we find back exactly all the different ways to identify the static model as may be compared by checking Proposition 4 of [10]. The rank conditions contained under Ψ_2 are the additional conditions.

It is possible for the structural parameter of the dynamic factor analysis model to be identified without any further prior restrictions beyond the zero--one-restrictions shown explicitly at (21). This is the case if in addition to the observability condition $H=\rho(\P'\Lambda_y', \P'^2\Lambda_y', \ldots, \P'^{T_1}\Lambda_y')$, we also have $H=\rho(\P(\mu_0)_{II}, \P(\Pi_0)_{II}x_1, \ldots, \P(\Pi_0)_{II}x_{T_1-1})$. This is in contrast to the static model with $\P=0$, where at least $GH+HK+H(H+1)/2$ prior restrictions on Λ_y, $\Gamma, \Sigma_{\varepsilon\varepsilon'}$, $\Sigma_{\zeta\zeta'}$ are always needed. This is an example of a dynamic model where the zero correlation between the structural observational and state variables equations is sufficient for identifiability.

IV. Concluding Remarks.

The main result of this paper is Proposition 1 and its Corollary 2 which contain necessary and sufficient conditions for the identifiability of the parameter of the ARMAX(p,q)-Model when the observational external characteristics are a finite number of first and second-order moments, instead of the transfer functions employed in the studies of asymptotic identifiability. We discuss how our results contain the results in asymptotic identifiability stated in the literature.

In the second part of the paper we developed necessary and sufficient conditions for the identifiability of the parameter of the ARMAX(p,q)-Model when some endogenous variables are latent variables. In comparison to the static model analyzed in [10], the conditions for identifiability of the dynamic model parameter are far less stringent. We illustrate this by specializing our results to the dynamic factor analysis model, which is a Kalman-filtering model with unknown coefficients, and show how observability conditions together with zero correlations between structural errors are sufficient for identifiability.

Dynamic features of the model help to identify, generally speaking.

REFERENCES

[1] Deistler M., The structural Identifiability of Linear Models with Autocorrelated Errors in the Case of Cross-Equation Restrictions, Journal of Econometrics 8 (1978) pp 23-31.

[2] Deistler M. and H-G Seifert, Identifiability and Consistent Estimability in Econometric Models, Econometrica, Vol. 46, No 6, July 1978.

[3] Deistler M. and J. Schrader, Linear Models with Autocorrelated Errors: Structural Identifiability in the Absence of Minimality Assumptions, Econometrica, Vol. 47, No 2, March 1979.

[4] Fisher F.M., The Identification Problem in Econometrics, McGraw Hill, 1966.

[7] Hatanaka M., On the Global Identification of the Dynamic Simultaneous Equations Model with Stationary Disturbances, International Economic Review, Vol.16, No 3, October 1975, pp 545-554.

[5] Hannan, E.J., The Identification of Vector Mixed Autoregressive Moving Average Systems, Biometrika, LVI, March 1969, pp 223-225.

[6] Hannan, E.J., The Identification Problem for Multiple Equation Systems with Moving Average Errors, Econometrica 39, 1971, pp 751-765.

[8] Kohn, R., Identification Results for Armax Structures, Econometrica, Vol.47, No 5, Sept. 1979, pp 1295-1304.

[9] Schönfeld, P., A Survey of Recent Concepts of Identification, Discussion Paper , University Bonn, 1975.

[10] Wegge, L., Identification with Latent Variables, Working Paper Sonderforschungsbereich 21, University of Bonn, January 1981.