The Identification Problem for Linear Rational Expectations Models

Majid M. Al-Sadoon
Piotr Zwiernik

September 2019

Barcelona GSE Working Paper Series

Working Paper nº 1114

The Identification Problem for Linear Rational Expectations Models*

Majid M. Al-Sadoon†
Durham University Business School

Piotr Zwiernik‡
Universitat Pompeu Fabra & BGSE

August 26, 2019

Abstract

We consider the problem of the identification of stationary solutions to linear rational expectations models from the second moments of observable data. Observational equivalence is characterized and necessary and sufficient conditions are provided for: (i) identification under affine restrictions, (ii) generic identification under affine restrictions of analytically parametrized models, and (iii) local identification under non-linear restrictions. The results strongly resemble the classical theory for VARMA models although significant points of departure are also documented.

JEL Classification: C10, C22, C32.

Keywords: Identification, linear rational expectations models, linear systems, vector autoregressive moving average models.

*Thanks are due to Manfred Deistler, Benedikt Pötscher, Hashem Pesaran, Fabio Canova, Barbara Rossi, Christian Brownlees, Geert Mesters, Davide Debortoli, Gábor Lugosi, Omiros Papaspiliopoulos, Albert Satorra, Bernd Funovits, Peter C. B. Phillips, Michele Piffer, and to seminar participants at Universitat Pompeu Fabra, University of Vienna, UC San Diego, University of Southampton, Universitat de les Illes Balears, NYU Abu Dhabi, London School of Economics, and Queen Mary University of London.

†MMA acknowledges support by Spanish Ministry of Economy and Competitiveness projects ECO2012-33247 and ECO2015-68136-P (MINECO/FEDER, UE) and Fundación BBVA Scientific Research Grant PR16-DAT-0043.

‡PZ acknowledges the support by the Spanish Government grants (RYC-2017-22544,PGC2018-101643-B-I00,SEV-2015-0563), and Ayudas Fundación BBVA a Equipos de Investigación Cientifica 2017.

1 Introduction

The linear rational expectations model (LREM) is distinguished among dynamic linear systems in that the present state depends not only on events leading up to the present but also on endogenously formulated expectations of the future. Such models are used today by researchers, practitioners, and policy-makers for causal and counter-factual analysis as well as forecasting. Yet the statistical properties of LREMs are poorly understood and identification in particular has remained an open problem throughout what is now known as the rational expectations revolution.

Establishing identifiability of this model class is important for several reasons. From the parametrization point of view, the parameters of LREMs codify the decision making of economic actors (e.g. households or firms) and it is important to know whether or not this behaviour can be learned from the available data. From the estimation point of view, lack of identifiability leads to ill-conditioned optimization procedures when employing extremum estimators; Bayesian methods are not immune to identification failure either as the posterior retains the shape of the prior along observationally equivalent directions in the parameter space. Finally, inference is substantially more difficult in the absence of identification in both the frequentist and Bayesian perspectives.

Partial results for identification of LREMs were derived by Muth (1981), Wallis (1980) and Pesaran (1981). However, these results apply to very restricted LREMs and cannot be employed in the study of modern LREMs. Subsequent econometric work in the area completely ignored identification until Canova & Sala (2009) called attention to serious identification problems plaguing many LREMs used in practice, a point echoed by Pesaran & Smith (2011), Romer (2016), and Blanchard (2018). This spurred a number of researchers to provide computational diagnostics for local and global identification (Iskrev, 2010; Komunjer & Ng, 2011; Qu & Tkachenko, 2017; Kociecki & Kolasa, 2018). Unfortunately, this work has not attempted an analytical examination of the mapping from parameters to observables of LREMs. Consequently, it has failed to uncover the underlying reasons for identification failure and has resorted instead to detecting the symptoms. A further consequence of this is that it has not been possible to make strong connections to classical identification results (e.g. the work of Hannan & Deistler (2012)).

The present work builds on recent results by Onatski (2006), Anderson et al. (2012), Anderson et al. (2016), and Al-Sadoon (2018) to explain why identification failure occurs in LREMs, provide a characterization of observational equivalence, and provide analytical diagnostic tests of identification that extend classical results for vector autoregressive moving average (VARMA) models. The key idea is that the mapping from parameters to observables of LREMs involves an initial Wiener-Hopf factorization but is otherwise identical to the mapping for VARMA. Once this is recognized, the theory proceeds almost exactly analogously to the classical theory.

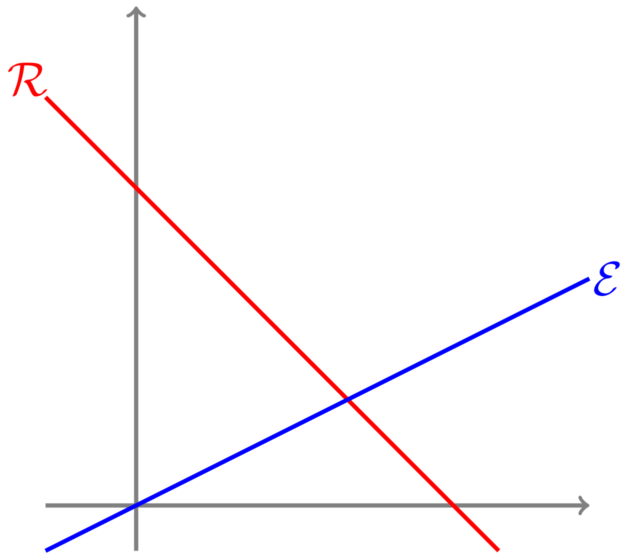
LREMs are subject to identification failure for some of the same reasons that simultaneous equations models and VARMA are. This is to be expected since the class of LREMs nests the aforementioned classes of models. However, there is a new source of identification failure that afflicts only LREMs and that has to do with the endogeneity of expectations. The fact that expectations in LREMs are functions of other endogenous variables necessitates more restrictions than might be called for in a classical model. Said differently, *forward dependence is not identified*. This result has been known to many authors in the literature, although it receives its most general treatment in this paper.

Our characterization of observational equivalence extends well-known results in the VARMA literature. We find that every class of observationally equivalent models sits inside a particular subspace of the parameter space, the dimension of which then gives the number of restrictions necessary for identification. This number is determined for both specific and generic points in the parameter space.

We then consider the identification problem under affine restrictions. These restrictions include zero restrictions, normalization to one restrictions, and (more generally) restrictions on linear combinations of parameters, possibly across equations. This leads to the geometric picture in Figure 1. Every point in the parameter space can now be thought of as lying at the intersection of two subspaces, the set of observationally equivalent parameters, $\mathcal{E}$, and the set of parameters satisfying the affine restrictions, $\mathcal{R}$. When these two subspaces intersect at a single point, that point is identified. Otherwise, they intersect along a subspace and local identification fails. This observation allows us to obtain necessary and sufficient conditions for identification under affine restrictions. It also immediately implies the equivalence of global and local identification for LREMs identified by affine restrictions.

Most LREMs in practice consist of parameters that are themselves rational functions of

Figure 1: The Geometry of Affine Restrictions.



more fundamental parameters, so-called "deep parameters." In this case, it will be difficult to provide necessary and sufficient conditions for identification under affine restrictions that would be easily testable. However, we show that necessary and sufficient conditions for identification under affine restrictions are still possible for generic systems. When the restrictions are non-linear, we provide necessary and sufficient conditions for local identification. The geometry of Figure 1 continues to be helpful here even as $\mathcal{R}$ is no longer an affine space.

It is perhaps worth pointing out some distinctive features of our analysis at the outset. Unlike previous approaches, ours does not requires any special assumptions on the number of exogenous shocks relative to observables (i.e. regularity) or redundant dynamics (i.e. minimality). Our approach does, however, make a strong identifying assumption on the first impulse response of the system. Nevertheless, we believe the new approach paves the way for much further progress on identification of LREMs as we discuss later on.

The reader wishing to have the full picture of the theory is advised to begin with Appendix A before starting Section 3. Section 2 sets the notation. Section 3 introduces the LREM and its solution. Section 4 characterizes observational equivalence in LREMs. Section 5 uses these results to provide conditions for identifiability of a parameter under affine constraints. Section 6 extends the set-up to non-linear parametrizations and constraints. Section 7 concludes.

2 Notation

Denote by $\mathbb{Z} \subset \mathbb{R} \subset \mathbb{C}$ the sets of integers, real numbers, and complex numbers respectively. We will need $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, and $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}$, the unit circle, the open unit disk, and the closed unit disk respectively. Complements of sets will be denoted by a superscript c. Denote by $\mathbb{R}[z] \subset \mathbb{R}[z, z^{-1}] \subset \mathbb{R}(z)$ the sets of real polynomials in z, real Laurent polynomials in z, and real irreducible rational functions in z respectively. Similarly, $\mathbb{R}[z^{-1}]$ is the set of real polynomials in z^{-1}. When forming arrays populated by elements of a given set, we will use the superscript $n \times m$ (e.g. $\mathbb{R}[z]^{n\times m}$ is the set of $n \times m$ polynomial matrices). When $m = 1$, we will simply use the superscript n (e.g. $\mathbb{R}^n$ is the set of n-dimensional real vectors). For a non-zero $B \in \mathbb{R}[z, z^{-1}]^{n\times m}$, we denote by $\max\deg(B)$ the highest power of z that appears in B, while $\min\deg(B)$ is the lowest power of z that appears in B. Finally, we will denote by I_n and $0_{n\times m}$ the $n \times n$ identity matrix and the $n \times m$ matrix of zeros respectively.

3 The Linear Rational Expectations Model

The linear rational expectations model (LREM) is characterized by the structural equations

$$(1) \qquad \sum_{i=-q}^{p} B_i \, \mathbb{E}_t(Y_{t-i}) = \sum_{i=0}^{k} A_i \varepsilon_{t-i}, \qquad t \in \mathbb{Z}.$$

Here ε is a sequence of m-dimensional, exogenous, and unobserved i.i.d. variables of mean zero and $\mathrm{var}(\varepsilon_0) = I_m$, while Y a sequence of n-dimensional endogenous observed variables. The variables in ε are understood as exogenous inputs to the system (e.g. shocks to productivity, monetary policy, etc.), while the variables in Y are understood as the output of the system (e.g. inflation, interest rates, etc.).

The coefficient matrices $A_0, \dots, A_k \in \mathbb{R}^{n\times m}$ codify the direct contemporaneous and lagged effects of the shocks on the system. The coefficient matrices $B_{-q}, \dots, B_p \in \mathbb{R}^{n\times n}$ codify how the expected, current, and lagged values of Y are directly related to each other. It will be convenient to encode the parameters as

$$B(z) = \sum_{i=-q}^{p} B_i z^i, \qquad\qquad A(z) = \sum_{i=0}^{k} A_i z^i.$$

Later we will introduce restrictions to these parameters and allow them to depend in a possibly non-linear way on another set of parameters. Following the usual convention we often omit the argument in the notation writing simply B and A.

Example 3.1. The most common specification of the LREM takes the form

$$B = \begin{bmatrix} \Phi & 0_{l\times(n-l)} \\ \Gamma & \mathsf{B} \end{bmatrix}, \qquad A = \begin{bmatrix} \Theta & 0_{l\times(m-l)} \\ 0_{(n-l)\times l} & \mathsf{A} \end{bmatrix}.$$

Here all submatrices except for $\mathsf{B} \in \mathbb{R}[z, z^{-1}]^{(n-l)\times(n-l)}$ are polynomial matrices. Thus, the first l variables of Y are considered exogenous in the last $n-l$ equations, which model the economic behaviour of primary interest. The first l variables of ε affect the dynamics of the exogenous variable, while the rest enter into the system directly.

In the special case where B is a polynomial matrix, we have the most general formulation of the classical VARMAX model with exogenous variables that have rational spectral density. Specializing further to the case where all matrices are constant, we obtain the classical simultaneous equations model.

A solution to (1) will be understood to be an n-dimensional stochastic process Y satisfying the causality condition that Y_t be measurable with respect to the σ-algebra generated by $\varepsilon_t, \varepsilon_{t-1}, \ldots$ for all $t \in \mathbb{Z}$, in addition to the structural equations (1) with $\mathbb{E}_t(\,\cdot\,) = \mathbb{E}(\,\cdot \mid \varepsilon_t, \varepsilon_{t-1}, \ldots)$ for all $t \in \mathbb{Z}$.

The parameter space of the LREM, denoted by Ω_{LREM}, is a set of pairs

$$(B, A) \in \mathbb{R}[z, z^{-1}]^{n\times n} \times \mathbb{R}[z]^{n\times m}$$

characterized by three restrictions, which we carefully introduce in this and the next section. The first of these restrictions is

$$\begin{aligned} & B = B_- B_+, \quad \text{where} \\ \text{(EU-LREM)} \quad & B_- \in \mathbb{R}[z^{-1}]^{n\times n}, \qquad \operatorname{rank}(B_-(z)) = n \text{ for all } z \in \mathbb{D}^c \text{ and } \lim_{z\to\infty} B_-(z) = I_n, \\ & B_+ \in \mathbb{R}[z]^{n\times n}, \qquad \operatorname{rank}(B_+(z)) = n \text{ for all } z \in \overline{\mathbb{D}}. \end{aligned}$$

Restriction (EU-LREM) is equivalent to the existence and uniqueness of a stationary solution to (1) (see Proposition 1 of Onatski (2006) and Theorem 3.2 of Al-Sadoon (2018)). This is because it is equivalent to the existence of a Wiener-Hopf factorization with zero partial indices for B (Clancey & Gohberg, 1981, Theorems I.1.1, I.1.2, and I.2.1).

To express the solution, we will need to recall the operators

$$\left[\sum_{i=-\infty}^{\infty} D_i z^i\right]_+ = \sum_{i=0}^{\infty} D_i z^i, \qquad\qquad \left[\sum_{i=-\infty}^{\infty} D_i z^i\right]_- = \sum_{i=-\infty}^{-1} D_i z^i,$$

where $\sum_{i=-\infty}^{\infty} D_i z^i$ converges in an annulus $\{z \in \mathbb{C} : \rho < |z| < 1\}$ for some $\rho \in (0,1)$ (Hansen & Sargent, 1981, Appendix A). Under restriction (EU-LREM), Onatski (2006) obtains the solution,

$$Y_t = B_+^{-1}(L) \left[B_-^{-1}A\right]_+ (L)\varepsilon_t, \qquad t \in \mathbb{Z}, \tag{2}$$

where L is the lag operator, $B_+^{-1}(L)$ is the composition of L and the Taylor series expansion of $B_+^{-1}(z)$ in a neighbourhood of $z = 0$, and $[B_-^{-1}A]_+(L)$ is the composition of L and the Taylor series expansion of $[B_-^{-1}A]_+(z)$ in a neighbourhood of $z = 0$.

Now defining

$$A^+ = B_-[B_-^{-1}A]_+,$$

we can rewrite the solution in the more familiar form

$$Y_t = B^{-1}(L)A^+(L)\varepsilon_t, \qquad t \in \mathbb{Z}, \tag{3}$$

where $B^{-1}(L)$ is the composition of L and the Laurent series expansion of $B^{-1}(z)$ in a neighbourhood of $\mathbb{T}$ and $A^+(L)$ is the composition of L and the Laurent series expansion of $A^+(z)$ in a neighbourhood of $\mathbb{T}$. This expression is similar to the solution of the classical VARMA model save for the substitution of A for A^+.

From (2) and (3) we see that there are two analytical expressions for the transfer function of the solution, $B_+^{-1}[B_-^{-1}A]_+$ and $B^{-1}A^+$. Each of these expressions plays a crucial role in the theory of identification of LREMs.

For future reference, we collect some of the algebraic consequences of (EU-LREM) in the following lemma.

Lemma 3.1. *Let* $(B, A) \in \mathbb{R}[z, z^{-1}]^{n\times n} \times \mathbb{R}[z]^{n\times m}$ *and let* B *satisfy* (EU-LREM)*. Then:*

(i) $\min\deg(B_-) = \min\deg(B)$ *and* $\max\deg(B_+) = \max\deg(B)$.

(ii) $[B_-^{-1}A]_+ \in \mathbb{R}[z]^{n\times m}$ *and* $\max\deg([B_-^{-1}A]_+) = \max\deg(A)$.

(iii) $A^+ \in \mathbb{R}[z, z^{-1}]^{n\times m}$, $\min\deg(A^+) \geq \min\deg(B)$, *and* $\max\deg(A^+) = \max\deg(A)$.

(iv) $[A^+]_+ = A$.

Proof. (i) Let the terms of lowest and highest degree of B_- and B_+ be $B_{-,\mu}z^{-\mu}$ and $B_{+,\nu}z^{\nu}$ respectively. Then the terms of lowest and highest degrees of $B = B_-B_+$ are $B_{-,\mu}B_{+,0}z^{-\mu}$ and $B_{-,0}B_{+,\nu}z^{\nu}$ respectively as $B_{+,0} = B_+(0)$ and $B_{-,0} = I_n$ are non-singular.

(ii) The highest degree term of the Laurent series expansion of B_-^{-1} in an annulus containing $\mathbb{T}$ is I_n. Thus, the highest degree term of $B_-^{-1}A$ is the highest degree term of A and the result follows.

(iii) Follows directly from (i) and (ii).

(iv) Compute

$$[A^+]_+ = [B_-[B_-^{-1}A]_+]_+ = [B_-(B_-^{-1}A)]_+ - [B_-[B_-^{-1}A]_-]_+ = [A]_+ - 0 = A.$$

The second equality follows from the fact that $B_-^{-1}A = [B_-^{-1}A]_+ + [B_-^{-1}A]_-$. The third equality follows from the fact that the Laurent series expansion of $B_-[B_-^{-1}A]_-$ in an annulus containing $\mathbb{T}$ consists of only negative powers of z. □

It follows from Lemma 3.1 (i) that the factors $B_\pm(z)$ are polynomial matrices in $z^{\pm 1}$ with degrees determined by the minimum and maximum degrees of B. Thus, in the VARMA setting where $\min\deg(B) = 0$, then $B_- = I_n$ and $B_+ = B$ so (EU-LREM) reduces to the condition for existence and uniqueness of a causal stationary solution to the VARMA model (condition (EU-VARMA) in Appendix A). Lemma 3.1 (ii) implies that the solution (2) can be viewed equivalently as a solution to the VARMA model with autoregressive part B_+ and moving average part $[B_-A]_+$, a fact that will play a crucial role in our analysis. Lemma 3.1 (iii) proves that A^+ is a Laurent polynomial with degrees bounded by the minimum degree of B and the maximum degree of A. It follows, again, that in the VARMA setting $A^+ = A$. Finally, Lemma 3.1 (iv) is the remarkable property of the solution that even though the mapping $(B, A) \mapsto (B, A^+)$ is highly non-linear, it is one-to-one and its left inverse is linear.

Example 3.2. Consider the LREM (1) with $n = m = p = q = k = 1$ under restriction (EU-LREM). Then Lemma 3.1 (i) implies that we may write $B = B_-B_+$ with $B_-(z) = 1 - b_-z^{-1}$, $B_+(z) = b_0(1 - b_+z)$, and $|b_\pm| < 1$. Note that in this parameterization $B_{-1} = -b_-b_0$, $B_0 = b_0(1 + b_-b_+)$, and $B_1 = -b_0b_+$. Writing $A(z) = a_0 + a_1z$, we have

$$B_-^{-1}(z)A(z) = \frac{a_1b_-^2 + a_0b_-}{z - b_-} + a_1b_- + a_0 + a_1z = (a_1b_-^2 + a_0b_-)\sum_{i=1}^{\infty} b_-^{i-1}z^{-i} + a_1b_- + a_0 + a_1z$$

in an annulus containing $\mathbb{T}$. This implies that

$$\left[B_-^{-1}A\right]_+(z) = a_1 b_- + a_0 + a_1 z.$$

In consequence,

$$A^+(z) = \frac{-a_1 b_-^2 - a_0 b_-}{z} + a_0 + a_1 z.$$

The solution finally is given by

$$Y_t = \left(\frac{a_1 b_- + a_0 + a_1 L}{b_0(1 - b_+ L)}\right)\varepsilon_t,$$

which is the same as the solution of the ARMA model with autoregressive part $B_+(z) = b_0(1 - b_+ z)$ and the moving average part $[B_-^{-1}A]_+(z) = a_1 b_- + a_0 + a_1 z$.

4 Observational Equivalence

Before completing the characterization of the parameter space, it is helpful to consider observational equivalence as a means to motivate the second and third properties of the parameter space.

The spectral density of the observed data,

$$f_{YY}(z) = \sum_{j=-\infty}^{\infty} \operatorname{cov}(Y_j, Y_0) z^j,$$

satisfies

$$f_{YY}(z) = B^{-1}(z)A^+(z)A^{+\prime}(z^{-1})B^{-1\prime}(z^{-1}).$$

We say that two parameters (B, A) and $(\tilde{B}, \tilde{A})$ are *observationally equivalent* and denote this by $(B, A) \sim (\tilde{B}, \tilde{A})$ if both produce the same spectral density; that is, if and only if

$$B^{-1}(z)A^+(z)A^{+\prime}(z^{-1})B^{-1\prime}(z^{-1}) = \tilde{B}^{-1}(z)\tilde{A}^+(z)\tilde{A}^{+\prime}(z^{-1})\tilde{B}^{-1\prime}(z^{-1}).$$

Evidently observationally equivalent parameters are related in a very complicated way. The traditional way forward in the VARMA literature has been to first impose restrictions that identify the transfer function before considering how to identify the parameters from the transfer function (Hannan & Deistler, 2012, Theorem 1.3.3). In particular, it is typically imposed that for every parameter (B, A) the transfer function $B^{-1}(z)A^+(z)$ is of full rank for all $z \in \mathbb{D}$. This condition is known variably in the literature as the *invertibility*, *fundamentalness*, or *minimum phase* condition. The transfer function then corresponds to a Wold representation of Y (Hannan & Deistler, 2012, p. 25). Imposing this additional restriction allows us to conclude that $(B, A) \sim (\tilde{B}, \tilde{A})$ if and only if there exists an orthogonal matrix $V \in \mathbb{R}^{m \times m}$ such that

$$\tilde{B}^{-1}\tilde{A}^{+} = B^{-1}A^{+}V.$$

See e.g. Theorems 4.6.8 and 4.6.11 of Lindquist & Picci (2015). Thus, we have identified the transfer function from the spectral density matrix up to an orthogonal transformation.

Restrictions that eliminate V are very well understood in the VARMA literature (Lütkepohl, 2005, Chapter 9). The simplest and most convenient choice for our purposes is to restrict the first coefficient matrix of the transfer function to be canonical quasi-lower triangular as in Anderson et al. (2012). Here, the first non-zero element of column j is positive and occurs in row i_j with $1 \leq i_1 < \cdots < i_m \leq n$. In the special case $n = m$, this is just the Cholesky identification scheme commonly attributed in the literature to Sims (1980). Thus, we arrive at the second property of all $(B, A) \in \Omega_{LREM}$,

(CF-LREM)

$$\begin{aligned}&\operatorname{rank}(B^{-1}(z)A^{+}(z)) = m \text{ for all } z \in \mathbb{D} \text{ and}\\ &B^{-1}(z)A^{+}(z)|_{z=0} \text{ is canonical quasi-lower triangular.}\end{aligned}$$

Note that under (EU-LREM), (CF-LREM) is equivalent to $[B_{-}^{-1}A]_{+}(z)$ having rank m for all $z \in \mathbb{D}$ and $B_{+}(0)[B_{-}^{-1}A]_{+}(0)$ being canonical quasi-lower triangular.

The invertbility part of (CF-LREM) is not restrictive for Gaussian ε, which is a typical specification in the LREM literature (Herbst & Schorfheide, 2016). This is because the distribution of a Gaussian stationary process is completely determined by its spectral density and so it is not possible to distinguish invertible from non-invertible Gaussian models (Rosenblatt, 2000, p. 11). The second part of (CF-LREM) imposing canonical quasi-lower triangularity of the first impulse response uniformly over the parameter space is, on the other hand, very restrictive and not likely to be satisfied for the typical multivariate LREM. However, it permits us a great deal of mathematical traction and we hope that it will be possible to relax it in future work. The reader who still finds (CF-LREM) objectionable may consider that we have solved the identification problem for LREMs up to an orthogonal matrix transformation, as in Anderson et al. (2016). We might add that this problem, which is mathematically equivalent to the global identification of the simplest static model (9) under general linear restrictions remains an open problem in the literature (Rubio-Ramírez et al. (2010) and Bacchiocchi & Kitagawa (2019) provide results in the special case of exact identification).

Even though we have not yet completed our characterization of Ω_{LREM}, we can already simplify observational equivalence based on conditions (EU-LREM) and (CF-LREM).

Theorem 4.1. *Let* (B, A) *and* $(\tilde{B}, \tilde{A})$ *satisfy* (EU-LREM) *and* (CF-LREM). *Then* $(B, A) \sim (\tilde{B}, \tilde{A})$ *if and only if*

$$\tilde{B}_+^{-1}[\tilde{B}_-^{-1}\tilde{A}]_+ = B_+^{-1}[B_-^{-1}A]_+$$

or equivalently

$$\tilde{B}^{-1}\tilde{A}^+ = B^{-1}A^+.$$

Theorem 4.1 generalizes well known classical results. When B and A are polynomial matrices, it reduces exactly to the result for the VARMA model (Theorem A.6) and when B and A are constants, it reduces to the result for the classical linear simultaneous equations model (Theorem A.2). Unfortunately Theorem 4.1 does not have the simple algebraic flavour of the classical models, which leads naturally to the notion of coprimeness and other useful and elegant algebraic ideas. However, we will see that with special care, these complications are surmountable.

Theorem 4.1 makes clear what the identification problem is for LREMs. Two parameters are observationally equivalent if and only if they generate the same transfer function. But the mapping from parameters to transfer functions factorizes as

$$(4) \quad (B, A) \overset{\varphi_1}{\longmapsto} (B_-, B_+, A) \overset{\varphi_2}{\longmapsto} (B_-, B_+, [B_-^{-1}A]_+) \overset{\varphi_3}{\longmapsto} (B_+, [B_-^{-1}A]_+) \overset{\varphi_4}{\longmapsto} B_+^{-1}[B_-^{-1}A]_+.$$

Since the Wiener-Hopf factorization is a well defined function and $B = B_-B_+$, φ_1 is one-to-one and so no identification problems are possible at this stage. Similarly, by Lemma 3.1 (iv), φ_2 is one-to-one so no identification problems are possible at this stage either. If we consider now φ_4, this mapping is not one-to-one due to the identification problem for VARMA models (see Section A for a review). This is to be expected because the set of VARMA is a subset of the set of LREMs. The only new aspect to the identification problem for LREMs is φ_3, where B_- is dropped. Since B_- determines the forward dependence of the solution to (1), we arrive at the following fundamental observation.

The Fundamental Aspect of the Identification Problem for LREMs:

Forward dependence is not identified.

The claim here goes beyond the well known fact that the effect of endogenous variables (e.g. expectations) in a classical structural equations model is not identified. What distinguishes LREMs is that expectations are endogenous variables which are determined by other endogenous variables. Thus, they necessitate more restrictions than necessary to identify the classical structural equation models. This point has already been made by several authors including the father of rational expectations himself, Muth (1981), in a paper written in 1960. Of course these earlier realizations of this point were in the context of much more restrictive LREMs relative to what we consider in this paper. By the end of this section we will characterize exactly how many additional restrictions are necessary for the identification of LREMs.

Before we can do that, however, we must introduce the final characterization of our parameter space. Notice that the set of elements in $\mathbb{R}[z, z^{-1}]^{n\times n} \times \mathbb{R}[z]^{n\times m}$ satisfying (EU-LREM) and (CF-LREM) is infinite dimensional. The sets of observationally equivalent parameters, as described in Theorem 4.1, are also infinite dimensional. In practice, however, LREMs are specified with a finite number of leads and lags. Thus, it is typically assumed that there exist non-negative integers κ and λ such that for every parameter (B, A),

$$\text{(L-LREM)} \qquad \min\deg(B) \geq -\lambda, \qquad \max\deg(B) \leq \kappa, \qquad \max\deg(A) \leq \kappa.$$

Thus, Ω_{LREM} is the set of pairs $(B, A) \in \mathbb{R}[z, z^{-1}]^{n\times n} \times \mathbb{R}[z]^{n\times m}$ satisfying (EU-LREM), (CF-LREM), and (L-LREM). The condition (L-LREM) allows us think of Ω_{LREM} as a subset of $\mathbb{R}^{n^2(\kappa+\lambda+1)+nm(\kappa+1)}$ with the Euclidean topology. The following result provides a parametrization of Ω_{LREM} analogous to the parametrization of VARMA models.

Proposition 4.2. *Ω_{LREM} is homeomorphic to a subset of $\mathbb{R}^{n^2(1+\lambda+\kappa)+nm(1+\kappa)-\frac{1}{2}m(m-1)}$, the interior of which consists of two connected components.*

Proof. Let

$$\Theta_{LREM} = \Bigg\{ (F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa, C_0, A_1, \ldots, A_\kappa) :$$

$$F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa \in \mathbb{R}^{n \times n}, C_0, A_1, \ldots, A_\kappa \in \mathbb{R}^{n \times m},$$

$$\text{rank}(F(z)) = n \text{ for all } z \in \mathbb{D}^c, \text{ where } F(z) = I_n + \sum_{i=1}^{\lambda} F_i z^{-i},$$

$$\text{rank}(G(z)) = n \text{ for all } z \in \overline{\mathbb{D}}, \text{ where } G(z) = \sum_{i=0}^{\kappa} G_i z^i,$$

$$C_0 \text{ is canonical quasi-lower triangular},$$

$$A(z) = G_0 C_0 - A_1 F^1 - \cdots - A_\kappa F^\kappa + \sum_{i=1}^{\kappa} A_i z^i,$$

$$\text{where } F^{-1}(z) = \sum_{i=0}^{\infty} F^i z^{-i}, \text{ for all } z \in \mathbb{D}^c,$$

$$\text{and rank}\left([F^{-1}A]_+(z)\right) = m \text{ for all } z \in \mathbb{D} \Bigg\}.$$

Then Θ_{LREM} can be viewed as a subset of $\mathbb{R}^{n^2(1+\lambda+\kappa)+nm(1+\kappa)-\frac{1}{2}m(m-1)}$. We claim that the mapping $\phi_{LREM} : \Theta_{LREM} \to \Omega_{LREM}$, defined by

$$(F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa, C_0, A_1, \ldots, A_\kappa) \mapsto$$
$$\left(\left(I_n + \sum_{i=1}^{\lambda} F_i z^{-i} \right) \left(\sum_{i=0}^{\kappa} G_i z^i \right), G_0 C_0 - A_1 F^1 - \cdots - A_\kappa F^\kappa + \sum_{i=1}^{\kappa} A_i z^i \right),$$

is a homeomorphism. Since $\left(I_n + \sum_{i=1}^{\lambda} F_i z^{-i} \right) \left(\sum_{i=0}^{\infty} F^i z^{-i} \right) = I_n$ for all $z \in \mathbb{D}^c$, we have that $F^0 = I_n$ and $F^i = -\sum_{j=1}^{\min\{\lambda,i\}} F_j F^{i-j}$ for $i \geq 1$ and so the elements of $F^1, \ldots, F^\kappa$ are polynomials in the elements of $F_1, \ldots, F_\lambda$ and are therefore continuous. It follows that ϕ_{LREM} is continuous. Next consider $\phi_{LREM}^{-1}(B, A)$. By the uniqueness of the Wiener-Hopf factorization, it must be that $F(z) = B_-(z)$ and $G(z) = B_+(z)$. The continuity of the Wiener-Hopf factorization (Clancey & Gohberg, 1981, Proposition X.1.1) then ensures that B is mapped continuously to $(F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa)$. On the other hand, it must be the case that $C_0 = G_0^{-1} \left(A_0 + A_1 F^1 + \cdots + A_\kappa F^\kappa \right)$ so that ϕ_{LREM}^{-1} is a function. Finally, C_0 is a rational function of the coefficient matrices of B_-, B_+, and A and therefore continuous over $(B, A) \in \Omega_{LREM}$. Thus, ϕ_{LREM}^{-1} is continuous.

Next, we claim that

$$\Theta_{LREM}^{\circ} = \Big\{ (F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa, C_0, A_1, \ldots, A_\kappa) \in \Theta_{LREM} : \\ \operatorname{rank}\big([F^{-1}A]_+(z)\big) = m \text{ for all } z \in \mathbb{T} \Big\}$$

is the interior of Θ_{LREM}. By the continuity of zeros of a polynomial with respect to its coefficients (Horn & Johnson, 1985, Appendix D), Θ_{LREM}° is open. Now pick any point $(F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa, C_0, A_1, \ldots, A_\kappa) \in \Theta_{LREM} \backslash \Theta_{LREM}^{\circ}$, then $\operatorname{rank}\big([F^{-1}A]_+(z_0)\big) < m$ for some $z_0 \in \mathbb{T}$. Now define

$$\begin{bmatrix} I_n & 0 & \cdots & 0 & 0 \\ F^1 & I_n & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & I_n & 0 \\ F^\kappa & \cdots & \cdots & F^1 & I_n \end{bmatrix} \begin{bmatrix} A_\kappa(\rho) \\ A_{\kappa-1}(\rho) \\ \vdots \\ A_2(\rho) \\ A_1(\rho) \end{bmatrix} = \\ \begin{bmatrix} \rho^\kappa & 0 & \cdots & 0 & 0 \\ 0 & \rho^{\kappa-1} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \rho^2 & 0 \\ 0 & \cdots & \cdots & 0 & \rho \end{bmatrix} \begin{bmatrix} I_n & 0 & \cdots & 0 & 0 \\ F^1 & I_n & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & I_n & 0 \\ F^\kappa & \cdots & \cdots & F^1 & I_n \end{bmatrix} \begin{bmatrix} A_\kappa \\ A_{\kappa-1} \\ \vdots \\ A_2 \\ A_1 \end{bmatrix}.$$

Then for any $\rho > 1$, $(F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa, C_0, A_1(\rho), \ldots, A_\kappa(\rho)) \notin \Theta_{LREM}$ because the moving average part of its VARMA representation is $[F^{-1}A(\rho)](z) = [F^{-1}A]_+(\rho z)$, which has a zero in $\mathbb{D}$. It follows that $\Theta_{LREM} \backslash \Theta_{LREM}^{\circ}$ are boundary points of Θ_{LREM}.

Finally, let $(F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa, C_0, A_1, \ldots, A_\kappa) \in \Theta_{LREM}^{\circ}$. Then, we may follow a similar reasoning to show that

$$(F_\lambda, \ldots, F_1, G_0, G_1(1-t), \ldots, G_\kappa(1-t), C_0, A_1(1-t), \ldots, A_\kappa(1-t)) \in \Theta_{LREM}^{\circ}, \qquad t \in [0,1],$$

where $G_j(\rho) = \rho^j G_j$ for $j = 1, \ldots, \kappa$. Thus, $(F_\lambda, \ldots, F_1, G_0, \ldots, G_\kappa, C_0, A_1, \ldots, A_\kappa) \in \Theta_{LREM}^{\circ}$ is in the same connected component as $(F_\lambda, \ldots, F_1, G_0, 0, \ldots, 0, C_0, 0, \ldots, 0)$. And following the same idea again, we may connect this point to $(0, \ldots, 0, G_0, 0, \ldots, 0, C_0, 0, \ldots, 0)$ by the path $(F_\lambda(1-t), \ldots, F_1(1-t), G_0, 0, \ldots, 0, C_0, 0, \ldots, 0)$ for $t \in [0,1]$, where $F_j(\rho) = \rho^j F_j$ for $j = 1, \ldots, \lambda$. The final claim then follows from Proposition A.1. □

In the course of proving Proposition 4.2, we find that Ω_{LREM} is neither open nor closed. The boundary points that are also elements of Ω_{LREM} are exactly those parameters where the transfer function has a zero on $\mathbb{T}$. The point $(0_{n\times n}, 0_{n\times m})$ is a boundary point of Ω_{LREM} that is not an element of Ω_{LREM}. The fact that Ω_{LREM} consists of two connected components is a property inherited from the classical simultaneous equations model (see Proposition A.1); for example $\left(I_n, \left[\begin{smallmatrix} I_m \\ 0 \end{smallmatrix}\right]\right)$ and $\left(\left[\begin{smallmatrix} -1 & 0 \\ 0 & I_{n-1} \end{smallmatrix}\right], \left[\begin{smallmatrix} I_m \\ 0 \end{smallmatrix}\right]\right)$ cannot be connected by a path in Ω_{LREM}.

Now that we have the full characterization of the parameter space we may consider simplifying the conditions for observationally equivalence even further than what we have seen in Theorem 4.1. Let $(B, A) \sim (\tilde{B}, \tilde{A})$ and let

$$C = B^{-1}A^{+}.$$

Then Theorem 4.1 implies that

$$\tilde{A}^{+} = \tilde{B}C.$$

If $\tilde{B}(z) = \sum_{j=-\lambda}^{\kappa} \tilde{B}_i z^i$, $\tilde{A}^{+}(z) = \sum_{i=-\lambda}^{\kappa} \tilde{A}_i^{+} z^i$, and $C(z) = \sum_{j=0}^{\infty} C_i z^i$ in an annulus containing $\mathbb{T}$, then equating term by term above we arrive to the following equivalent expression

$$\begin{bmatrix} \tilde{A}_{-\lambda}^{+} & \cdots & \tilde{A}_{\kappa}^{+} & 0 & \cdots \end{bmatrix} = \begin{bmatrix} \tilde{B}_{-\lambda} & \cdots & \tilde{B}_{\kappa} & 0 & \cdots \end{bmatrix} \begin{bmatrix} C_0 & C_1 & C_2 & C_3 & \cdots \\ 0 & C_0 & C_1 & C_2 & \ddots \\ 0 & 0 & C_0 & C_1 & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Although this is an infinite dimensional system, (L-LREM) will allow us to restrict attention to a finite dimensional subsystem, which will then allow us to provide a simpler characterization of observational equivalence. The key idea is a familiar one from linear systems theory (see e.g. Lemma A.4 of Dufour & Renault (1998)).

Lemma 4.3. *Suppose $f, g \in \mathbb{R}[z]$, $\max\deg(f) \leq p$ and $g(0) \neq 0$. Then $h = f/g$ is identically zero if and only if the first $p+1$ terms in the Taylor series expansion of $h(z)$ in a neighbourhood of $z = 0$ are zero.*

The essential point of Lemma 4.3 is that the coefficients of any Taylor series expansion of a rational function are linearly recursive and therefore determined by the initial coefficients.

Before deriving the simplification of observational equivalence, we will also need the following lemma which develops some properties of a submatrix of the infinite matrix we have identified above. The reader may wish to review the concepts of coprimeness and the McMillan degree, δ, of a rational matrix discussed in Section A.2 (c.f. (15)).

Lemma 4.4. *Let $(B, A) \in \Omega_{LREM}$, let $C = B^{-1}A^{+}$ have a Taylor series expansion $C(z) = \sum_{i=0}^{\infty} C_i z^i$ in a neighbourhood of $z = 0$, and let*

$$H = \begin{bmatrix} C_{\kappa+\lambda+1} & C_{\kappa+\lambda+2} & \cdots & C_{(n+1)\kappa+\lambda} \\ \ddots & \ddots & \ddots & \vdots \\ C_2 & \ddots & \ddots & C_{n\kappa+1} \\ C_1 & C_2 & \ddots & C_{n\kappa} \end{bmatrix}.$$

Then:

(i) $\operatorname{rank}(H) = \delta(C(z^{-1}) - C(0)) \leq n\kappa$.

(ii) The set of parameters satisfying $\operatorname{rank}(H) = n\kappa$ *is generic (i.e. contains an open and dense subset of* Ω_{LREM}*).*

Proof. (i) The conditions for $(B, A) \in \Omega_{LREM}$ together with Lemma 3.1 (i) and (ii) imply that $(B_+, [B_-^{-1}A]_+) \in \Omega_{VARMA}$, the parameter space for VARMA models developed in Section A. The result is obtained in the course of proving Lemma A.8 (i).

(ii) Lemma A.8 (ii) identifies a generic subset of Ω_{VARMA} where the result holds. We will prove that its preimage under the mapping $(B, A) \mapsto (B_+, [B_-^{-1}A]_+)$ is generic in Ω_{LREM}. This will follow if we can show that $\varphi_3 \circ \varphi_2 \circ \varphi_1$ in (4) is continuous and open (observe that if $\mathscr{X}$ and $\mathscr{Y}$ are topological spaces, $f : \mathscr{X} \to \mathscr{Y}$ is continuous and open, and $\mathscr{D}$ is open and dense in $\mathscr{Y}$, then $f^{-1}(\mathscr{D})$ is open and dense in $\mathscr{X}$). First,

$$(B, A) \xmapsto{\varphi_1} (B_-, B_+, A)$$

viewed as a mapping $(\mathbb{R}^{n\times n})^{1+\lambda+\kappa} \times (\mathbb{R}^{n\times m})^{1+\kappa} \to (\mathbb{R}^{n\times n})^{\lambda} \times (\mathbb{R}^{n\times n})^{1+\kappa} \times (\mathbb{R}^{n\times m})^{1+\kappa}$ is continuous by Proposition X.1.1 of Clancey & Gohberg (1981). Since multiplication is continuous, φ_1^{-1} is also continuous and φ_1 is therefore a homeomorphism. Next,

$$(B_-, B_+, A) \xmapsto{\varphi_2} (B_-, B_+, [B_-^{-1}A]_+),$$

viewed as mapping $(\mathbb{R}^{n\times n})^{\lambda} \times (\mathbb{R}^{n\times n})^{1+\kappa} \times (\mathbb{R}^{n\times m})^{1+\kappa}$ to itself, is a composition of inversion, multiplication, and $[\,\cdot\,]_+$ all of which are continuous in the domain we consider. By

Lemma 3.1 (iii), φ_2^{-1} is a composition of multiplication and $[\ \cdot\ ]_+$, which are continuous. Thus, φ_2 is also a homeomorphism. Finally,

$$(B_-, B_+, [B_-^{-1}A]_+) \overset{\varphi_3}{\longmapsto} (B_+, [B_-^{-1}A]_+)$$

is a projection and therefore continuous and open. The composition of the three mappings is then continuous and open. $\square$

We are now in a position to simplify Theorem 4.1 and characterize the set

$$(B,A)/\sim = \left\{(\tilde{B},\tilde{A}) \in \Omega_{LREM} : (\tilde{B},\tilde{A}) \sim (B,A)\right\}, \qquad (B,A) \in \Omega_{LREM}.$$

Theorem 4.5. *Let $(B,A),(\tilde{B},\tilde{A}) \in \Omega_{LREM}$, let $C = B^{-1}A^+$ have a Taylor series expansion $C(z) = \sum_{i=0}^{\infty} C_i z^i$ in a neighbourhood of $z = 0$, and let*

$$T = \begin{bmatrix} C_0 & C_1 & \cdots & C_{\kappa+\lambda} \\ 0 & C_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & C_1 \\ 0 & \cdots & 0 & C_0 \end{bmatrix}, \qquad H = \begin{bmatrix} C_{\kappa+\lambda+1} & C_{\kappa+\lambda+2} & \cdots & C_{(n+1)\kappa+\lambda} \\ \ddots & \ddots & \ddots & \vdots \\ C_2 & \ddots & \ddots & C_{n\kappa+1} \\ C_1 & C_2 & \ddots & C_{n\kappa} \end{bmatrix},$$

$$P = \begin{bmatrix} -T & -H \\ I_{m(\kappa+\lambda+1)} & 0_{m(\kappa+\lambda+1)\times nm\kappa} \end{bmatrix}.$$

Then:

(i) $(\tilde{B},\tilde{A}) \sim (B,A)$ *if and only if*

$$\operatorname{vec}\left(\begin{bmatrix} \tilde{B}_{-\lambda} & \cdots & \tilde{B}_\kappa & \tilde{A}^+_{-\lambda} & \cdots & \tilde{A}^+_\kappa \end{bmatrix}\right) \in \ker\left(P' \otimes I_n\right).$$

(ii) $(B,A)/\sim$ *is a relatively open subset of the subspace*

$$\operatorname{mat}\left(\ker\left(P' \otimes I_n\right)\right),$$

where

$$\operatorname{mat} : \operatorname{vec}\left(\begin{bmatrix} B_{-\lambda} & \cdots & B_\kappa & A^+_{-\lambda} & \cdots & A^+_\kappa \end{bmatrix}\right) \mapsto \left(\sum_{i=-\lambda}^{\kappa} B_i z^i, \sum_{i=0}^{\kappa} A_i^+ z^i\right).$$

(iii) $\dim((B,A)/\sim) = n^2(\kappa+\lambda+1) - n\,\delta\big(C(z^{-1}) - C(0)\big) \geq n^2(1+\lambda)$ *and for generic points in the parameter space* $\dim((B,A)/\sim) = n^2(1+\lambda)$.

Proof. (i) If $(\tilde{B}, \tilde{A}) \sim (B, A)$, then Theorem 4.1 implies that

$$\begin{aligned}0 &= -z^{\lambda}\tilde{B}(z)C(z) + z^{\lambda}\tilde{A}^{+}(z)\\ &= -z^{\lambda}\tilde{B}(z)B_{+}^{-1}(z)[B_{-}^{-1}A]_{+}(z) + z^{\lambda}\tilde{A}^{+}(z)\\ &= \frac{-z^{\lambda}\tilde{B}(z)\mathrm{adj}(B_{+}(z))[B_{-}^{-1}A]_{+}(z) + \det(B_{+}(z))z^{\lambda}\tilde{A}^{+}(z)}{\det(B_{+}(z))}.\end{aligned}$$

By Lemma 3.1 and (EU-LREM), each element of the right hand side can be expressed as a ratio of a polynomial (of degree at most $\max\left\{\lambda + \max\deg(\tilde{B}) + \max\deg(\mathrm{adj}(B_{+})) + \max\deg([B_{-}^{-1}A]_{+}), \lambda + \max\deg(\det(B_{+})) + \max\deg(\tilde{A}^{+})\right\} \leq \max\{\lambda + \kappa + (n-1)\kappa + \kappa, \lambda + n\kappa + \kappa\} = (n+1)\kappa + \lambda$) and $\det(B_{+})$, which satisfies $\det(B_{+}(0)) \neq 0$. By Lemma 4.3, this is equivalent to the first $1 + (n+1)\kappa + \lambda$ Taylor series coefficients equating to zero. Thus, observational equivalence is equivalent to

$$-\begin{bmatrix} \tilde{B}_{-\lambda} & \dots & \tilde{B}_{\kappa} \end{bmatrix}\begin{bmatrix} T & H \end{bmatrix} + \begin{bmatrix} \tilde{A}^{+}_{-\lambda} & \dots & \tilde{A}^{+}_{\kappa} & 0_{n\times m} & \dots & 0_{n\times m} \end{bmatrix} = 0_{n\times(1+(n+1)\kappa+\lambda)m}$$

or equivalently

$$\begin{bmatrix} \tilde{B}_{-\lambda} & \cdots & \tilde{B}_{\kappa} & \tilde{A}^{+}_{-\lambda} & \cdots & \tilde{A}^{+}_{\kappa} \end{bmatrix} P = 0_{n\times(1+(n+1)\kappa+\lambda)m}.$$

Vectorizing we obtain

$$(P' \otimes I_n)\,\mathrm{vec}\left(\begin{bmatrix} \tilde{B}_{-\lambda} & \cdots & \tilde{B}_{\kappa} & \tilde{A}^{+}_{-\lambda} & \cdots & \tilde{A}^{+}_{\kappa} \end{bmatrix}\right) = 0_{nm(1+(n+1)\kappa+\lambda)\times 1}.$$

(ii) If $(\breve{B}, \breve{A}) \in \mathrm{mat}(\ker(P' \otimes I_n))$ then it satisfies (L-LREM). We claim that if, additionally, it satisfies (EU-LREM), then (CF-LREM) is also satisfied. To see this, note that the first $1+(n+1)\kappa+\lambda$ Taylor series coefficients of $\breve{A} - [\breve{B}C]_{+}$ equate to zero and, following the same argument as used in (i), it can be shown that $\breve{A} = [\breve{B}C]_{+}$. Therefore

$$\begin{aligned}\breve{A}^{+} &= \breve{B}_{-}[\breve{B}_{-}^{-1}\breve{A}]_{+}\\ &= \breve{B}_{-}[\breve{B}_{-}^{-1}[\breve{B}C]_{+}]_{+}\\ &= \breve{B}_{-}[\breve{B}_{-}^{-1}\breve{B}C]_{+} - \breve{B}_{-}[\breve{B}_{-}^{-1}[\breve{B}C]_{-}]_{+}\\ &= \breve{B}_{-}[\breve{B}_{+}C]_{+}\\ &= \breve{B}_{-}\breve{B}_{+}C\\ &= \breve{B}C,\end{aligned}$$

where we have used the property that $[\,\cdot\,]_+ + [\,\cdot\,]_-$ is the identity mapping on $\mathbb{R}(z)^{n\times m}$ as well as the fact that $\breve{B}_+$ and C are analytic in $\overline{\mathbb{D}}$. It follows that (CF-LREM) is satisfied as claimed. Thus, $(B,A)/\sim$ is the intersection of $\mathrm{mat}(\ker(P'\otimes I_n))$ with

$$\left\{(\breve{B},\breve{A})\in\mathbb{R}[z,z^{-1}]^{n\times n}\times\mathbb{R}[z]^{n\times m} : (\text{EU-LREM}) \text{ and } (\text{L-LREM}) \text{ are satisfied}\right\}.$$

The latter set is open in $\mathbb{R}^{n^2(\kappa+\lambda+1)+nm(\kappa+1)}$ due to the continuity of the Wiener-Hopf factorization with respect to entries of the matrix function (Clancey & Gohberg, 1981, Proposition X.1.1). Therefore, $(B,A)/\sim$ is relatively open in $\mathrm{mat}(\ker(P'\otimes I_n))$.

(iii) $\dim(\ker(P'))=\dim(\ker(H'))$ and so the result follows from Lemma 4.4 (i) and the standard properties of Kronecker products (Horn & Johnson, 1991, Theorem 4.2.15). For generic parameters the result follows from Lemma 4.4 (ii). □

Theorem 4.5 (i) provides a significantly simpler criterion for observational equivalence than Theorem 4.1; it substitutes a rational matrix equation for a simple linear algebraic criterion. Theorem 4.5 (ii) characterizes a set of observationally equivalent parameters as a relatively open subset of an affine subspace of Ω_{LREM} determined by the first $1+(n+1)\kappa+\lambda$ impulse responses. Finally, Theorem 4.5 (iii) gives the dimension of a set of observationally equivalent parameters, which can be understood as the number of restrictions that must be imposed on the parameter space in order to identify a parameter with a spectral density matrix. This number can be understood as the difference between the effective number of free parameters $n^2(\kappa+\lambda+1)$ (recall that $\tilde{A}^+$ is determined from C whenever $\tilde{B}$ is known) and the complexity of the transfer function $n\delta(C(z^{-1})-C(0))$.

When $\lambda=0$, Theorem 4.5 specializes to classical results for VARMA models as reviewed in Appendix A. Notice that the dimension of observationally equivalent parameters is larger by $n^2\lambda$ than in the VARMA setting (Theorem A.9 (iii)). This is exactly the number of free parameters in $B_-(z)$ and another manifestation of the fundamental aspect of the identification problem for LREMs.

Example 4.1. Consider the setting of Example 3.2. If $(\tilde{B},\tilde{A})\sim(B,A)$ then, by Theorem 4.1,

$$(\tilde{B}_{-1}z^{-1}+\tilde{B}_0+\tilde{B}_1z)\left(\frac{a_1b_-+a_0+a_1z}{b_0(1-b_+z)}\right)=\tilde{A}^+_{-1}z^{-1}+\tilde{A}^+_0+\tilde{A}^+_1z.$$

Multiplying through by $b_0(1-b_+z)z$ we obtain an equality of two polynomials of degree at most three, which implies four linear equations in the six variables $\tilde{B}_{-1}$, $\tilde{B}_0$, $\tilde{B}_1$, $\tilde{A}^+_{-1}$, $\tilde{A}^+_0$, and $\tilde{A}_1^+$. These are precisely the equations verified in Theorem 4.5 (i). We have,

$$P' \otimes I_n = \begin{bmatrix} -C_0 & 0 & 0 & 1 & 0 & 0 \\ -C_1 & -C_0 & 0 & 0 & 1 & 0 \\ -C_2 & -C_1 & -C_0 & 0 & 0 & 1 \\ -C_3 & -C_2 & -C_1 & 0 & 0 & 0 \end{bmatrix}.$$

Consider now the special case of the parameter $(B, A) = (1, 1)$ so that $C_0 = 1$ and $C_1 = C_2 = C_3 = 0$. Then $\left(0, 1, \frac{1}{2}, 0, 1, \frac{1}{2}\right)' \in \ker(P' \otimes I_n)$. This corresponds to $\left(1 + \frac{1}{2}z, 1 + \frac{1}{2}z\right) \in \Omega_{LREM}$ and it follows from Theorem 4.5 (i) that $\left(1 + \frac{1}{2}z, 1 + \frac{1}{2}z\right) \sim (1, 1)$. This illustrates the identification failure familiar for ARMA models due to non-coprimeness. A more interesting example in our context is $\left(\frac{1}{3}, 1, \frac{1}{2}, \frac{1}{3}, 1, \frac{1}{2}\right)' \in \ker(P' \otimes I_n)$, which corresponds to $\left(\frac{1}{3}z^{-1} + 1 + \frac{1}{2}z, 1 + \frac{1}{2}z\right) \in \Omega_{LREM}$ and is therefore also observationally equivalent to $(1, 1)$.

5 Identification by Affine Restrictions

In practice, LREMs are restricted in a variety of ways such as exclusion (setting a parameter to zero), normalization (setting a parameter to 1), and, more generally, affine restrictions that set linear combinations of the parameters (possibly across equations) to fixed values. Here we consider the ability of such restrictions to identify a single parameter.

Let Ω^R_{LREM} be a subset of Ω_{LREM} endowed with the relative topology. We say that $(B, A) \in \Omega^R_{LREM}$ is *identified in* Ω^R_{LREM} if every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in Ω^R_{LREM} is equal to (B, A). We say that a parameter (B, A) is *locally identified in* Ω^R_{LREM} if it has a neighbourhood N in Ω^R_{LREM} such that every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in N is equal to (B, A). Clearly, a parameter is locally identified in Ω^R_{LREM} if it is identified in Ω^R_{LREM} but the converse is not true in general.

Theorem 5.1. *Let Ω^R_{LREM} be the set of $(B, A) \in \Omega_{LREM}$ satisfying*

$$R \operatorname{vec}\left(\begin{bmatrix} B_{-\lambda} & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right) = u, \tag{5}$$

where $R \in \mathbb{R}^{r \times n^2(\kappa+\lambda+1)+nm(\kappa+1)}$ and $u \in \mathbb{R}^r$. Partition R as $\begin{bmatrix} R_B & R_A \end{bmatrix}$, where $R_A \in \mathbb{R}^{r \times n^2(\kappa+\lambda+1)}$ and $R_B \in \mathbb{R}^{r \times nm(\kappa+1)}$, and set $\overline{R} = \begin{bmatrix} R_B & 0_{r \times nm\lambda} & R_A \end{bmatrix}$. If $(B, A) \in \Omega^R_{LREM}$ and P is defined as in Theorem 4.5, then (B, A) is identified in Ω^R_{LREM} if and only if

$$M = \begin{bmatrix} P' \otimes I_n \\ \overline{R} \end{bmatrix}$$

is of full column rank $n(n+m)(\kappa+\lambda+1)$*,*

Proof. Let $\zeta = \text{vec}\left(\begin{bmatrix} B_{-\lambda} & \cdots & B_\kappa & A^+_{-\lambda} & \cdots & A^+_\kappa \end{bmatrix}\right)$. If M is of full column rank, then ζ is the only point in $\ker(P' \otimes I_n)$ that satisfies (5). Theorem 4.5 (i) then implies that (B, A) is identified in Ω^R_{LREM}. If M is not of full rank, then there exists $0 \neq \xi \in \ker(P' \otimes I_n) \cap \ker(\overline{R})$. If $c > 0$ is sufficiently small, Theorem 4.5 (ii) implies that $(B, A) \sim \text{mat}(\zeta + c\xi) \neq (B, A)$ and since $\text{mat}(\zeta + c\xi)$ satisfies (5), (B, A) is not identified in Ω^R_{LREM}. □

Theorem 5.1 is a direct generalization of classical results for simultaneous equations models (the case $\lambda = \kappa = 0$) and for VARMA (the case $\lambda = 0$) derived in Theorems A.4 and A.10 respectively. Theorem 5.1 also exhibits similar geometry to the classical results (Figure 1). Any given parameter $(B, A) \in \Omega_{LREM}$ lies in the intersection of two affine subspaces. The first affine subspace, denoted by $\mathcal{E}$, is $\text{mat}\left(\ker(P' \otimes I_n)\right)$ and contains the set of parameters observationally equivalent to (B, A) by Theorem 4.5 (i). The second affine subspace is the space Ω^R_{LREM}, denoted by $\mathcal{R}$, which contains the set of parameters satisfying the restrictions (5). When (B, A) is the only point of intersection then it is identified in Ω^R_{LREM}. Otherwise, the two affine subspaces intersect along an affine a subspace, which contains a line segment in Ω^R_{LREM} by Theorem 4.5 (ii) and so every neighbourhood of (B, A) contains infinitely many observationally equivalent parameters that also satisfy the given restrictions. Thus, *for the LREM subject to affine restrictions, a parameter is identified if and only if it is locally identified.*

Theorem 5.1 leaves unstated whether or not u in (5) is equal to the zero vector. In fact, it is meaningless to allow u to be the zero vector. If $u = 0$, then Theorem 4.5 (i) and (ii) imply that $(cB, cA) \in \Omega^R_{LREM} \cap (B, A)/\sim$ for all c in some neighbourhood of 1, thus (B, A) cannot be identified. Stated differently, if $u = 0$ then M cannot be of full rank because this would force $(B, A) = (0_{n\times n}, 0_{n\times m}) \notin \Omega_{LREM}$.

Example 5.1. Consider the setting of Example 3.2. By Theorem 4.5 and the discussion following it, to obtain identifiability we need to impose at least $n^2(1+\lambda) = 2$ restrictions.

Thus, we may consider fixing B_{-1} and A_0. This implies that

$$M = \begin{bmatrix} -C_0 & 0 & 0 & 1 & 0 & 0 \\ -C_1 & -C_0 & 0 & 0 & 1 & 0 \\ -C_2 & -C_1 & -C_0 & 0 & 0 & 1 \\ -C_3 & -C_2 & -C_1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

We have $\det(M) = C_0 C_1$ and so any point in the parameter space satisfying our two restrictions and such that its associated transfer function has a non-zero second impulse response is identified. Note that the two restrictions are not sufficient to identify the parameter $(B, A) = (1, 1)$ that we considered in Example 4.1. Since there are more free parameters than necessary in order to characterize this point, we will need two additional restrictions.

Suppose now that we are interested in identifying just the i-th equation of (1). Let Ω^R_{LREM} be as before and let $(B, A) \in \Omega^R_{LREM}$. We say that the i-th equation of (1) is *identified at* (B, A) *in* Ω^R_{LREM} if every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in Ω^R_{LREM} has the same i-th equation as (B, A). We say that the i-th equation of (1) is *locally identified at* (B, A) *in* Ω^R_{LREM} if (B, A) has a neighbourhood N in Ω^R_{LREM} such that every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in N has the same i-th equation as (B, A). Clearly, if all equations are (locally) identified at (B, A) in Ω^R_{LREM}, then (B, A) is (locally) identified in Ω^R_{LREM}.

Theorem 5.2. *Let Ω^R_{LREM} be the set of $(B, A) \in \Omega_{LREM}$ satisfying*

$$R_i \operatorname{vec}\left(e_i' \begin{bmatrix} B_{-\lambda} & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right) = u_i, \tag{6}$$

where $R_i \in \mathbb{R}^{r \times n(\kappa+\lambda+1)+m(\kappa+1)}$, $u_i \in \mathbb{R}^r$, and $e_i \in \mathbb{R}^n$ is the i-th standard unit vector. Partition R_i as $\begin{bmatrix} R_{iB} & R_{iA} \end{bmatrix}$, where $R_{iA} \in \mathbb{R}^{r \times n(\kappa+\lambda+1)}$ and $R_{iB} \in \mathbb{R}^{r \times m(\kappa+1)}$, and set $\overline{R}_i = \begin{bmatrix} R_{iB} & 0_{r \times m\lambda} & R_{iA} \end{bmatrix}$. If $(B, A) \in \Omega^R_{LREM}$ and P is defined as in Theorem 4.5, then the i-th equation of (1) is identified at (B, A) in Ω^R_{LREM} if and only if

$$M_i = \begin{bmatrix} P' \\ \overline{R}_i \end{bmatrix}$$

has full column rank $(n + m)(\kappa + \lambda + 1)$.

Proof. Let $\zeta = \text{vec}\left(\begin{bmatrix} B_{-\lambda} & \cdots & B_\kappa & A^+_{-\lambda} & \cdots & A^+_\kappa \end{bmatrix}\right)$. If M_i is of full column rank, then $\text{vec}\left(e_i'\begin{bmatrix} B_{-\lambda} & \cdots & B_\kappa & A^+_{-\lambda} & \cdots & A^+_\kappa \end{bmatrix}\right) = (I_{(n+m)(\kappa+\lambda+1)} \otimes e_i')\zeta$ is the only point in $\ker(P')$ that satisfies (6), since $P'(I_{(n+m)(\kappa+\lambda+1)} \otimes e_i') = (I_{m(1+(n+1)\kappa+\lambda)} \otimes e_i')(P' \otimes I_n)$. Theorem 4.5 (i) then implies that any parameter in Ω^R_{LREM} that is observationally equivalent to (B, A) must have the same i-th equation as (B, A). Thus the i-th equation is identified at (B, A) in Ω^R_{LREM}. If M_i is not of full rank, then there exists $0 \neq \xi_i \in \ker(P') \cap \ker(\overline{R}_i)$. If $c > 0$ is sufficiently small, Theorem 4.5 (ii) implies that $(B, A) \sim \text{mat}(\zeta + c\xi_i \otimes e_i)$ and since $\text{mat}(\zeta + c\xi_i \otimes e_i)$ satisfies (6) but has a different i-th equation than (B, A), the i-th equation is not identified at (B, A) in Ω^R_{LREM}. □

Theorem 5.2 provides necessary and sufficient conditions for the identification of an equation of an LREM. It has exactly the same flavour, interpretation, and geometry as Theorem 5.1. It also generalizes classical results for VARMA and simultaneous equations models and retains the property of equivalence of identification and local identification. Finally, for the same reason as before, $u_i = 0$ cannot be allowed.

We remark that Theorems 5.1 and 5.2 can be formulated in terms of different matrices than M and M_i respectively. In Corollary A.11 we show that, when attention is restricted to VARMA models, our result is equivalent to that of Deistler & Schrader (1979), who formulate identification in terms of the rank of a matrix populated not by impulse responses but by the coefficient matrices of (B, A). The direct generalization of Deistler & Schrader (1979) to the LREM setting would then involve a matrix populated by (B, A^+). Since this formulation would involve the negative terms of A^+, which have no economic interpretation, it is clear that our formulation in terms of impulse responses is the preferable one.

6 Generic and Local Identification

In practice it is usually the case that we are interested in the identification of all parameters, not just a single one. Moreover, most LREMs in practice are not only restricted by affine constraints but their free parameters are usually functions of more fundamental structural parameters. Similarly, the restrictions considered will not always be affine. In this section, we provide results for these scenarios. We begin with a motivating example.

Example 6.1. Hansen & Sargent (1981) study an LREM for the level of employment of a factor of production. Their model is of the form considered in Example 3.2, parametrized as

$$B(z) = \theta_1 z^{-1} - ((\theta_3/\theta_2) + 1 + \theta_1) + z, \qquad A(z) = \theta_2^{-1}.$$

Here, θ_1 is a time discount factor, θ_2 is a cost of adjustment, and θ_3 is a measure of returns to scale. This model has $\kappa = 1$ and $\lambda = 1$ and is subject to two affine restrictions, $B_1 = 1$ and $A_1 = 0$. The question then is whether every $\theta = (\theta_1, \theta_2, \theta_3)$ is identified.

We can attempt to answer the question posed in Example 6.1 as follows. Let the parameter space of interest be a subset $\Theta \subset \mathbb{R}^d$ that maps one-to-one to a subset of Ω_{LREM}. Thus, fixing a linear subspace as in the previous section, there is a one-to-one mapping $\phi : \Theta \to \Omega_{LREM}^R$, where $\Omega_{LREM}^R \subset \Omega_{LREM}$ is, as before, endowed with the relative topology. We refer to an LREM parameterized as above as a ϕ-LREM. We say that a ϕ-LREM is generically identified in Ω_{LREM}^R if there is a relatively open and dense subset $\Psi \subset \Theta$ such that every parameter in $\phi(\Psi)$ is identified in Ω_{LREM}^R. Clearly if every point of $\phi(\Theta)$ is identified in Ω_{LREM}^R, then the LREM is generically identified in Ω_{LREM}^R.

In order to develop results for this new notion of identification, we will need the following well known lemma for which we offer an elementary proof.

Lemma 6.1. *Let $\mathscr{X} \subset \mathbb{R}^d$ be non-empty, open, and connected. Let $f : \mathscr{X} \to \mathbb{R}$ be a non-zero real analytic function. Then $\{x \in \mathscr{X} : f(x) \neq 0\}$ is open and dense in $\mathscr{X}$.*

Proof. The continuity of f implies that the set is indeed open. It remains to prove denseness. Let $x, y \in \mathscr{X}$ be such that $f(x) = 0$ and $f(y) \neq 0$. The connectedness and openness of $\mathscr{X}$ imply that there exists a polygonal path in $\mathscr{X}$ with vertices $x = z_0, z_1, \ldots, z_{n-1}, z_n = y$. Suppose the first integer i such that $f(z_i) \neq 0$ is greater than 1. Then $f(tz_i + (1-t)z_{i-1})$ is a non-zero real analytic function in t defined over an open interval containing $[0, 1]$ with a zero at $t = 0$. Since the zeros of a non-zero real analytic function over an open interval are isolated (Krantz & Parks, 2002, Corollary 1.2.6), it follows that $f(tz_i + (1-t)z_{i-1}) \neq 0$ for any small enough $t \neq 0$. If we now perturb z_{i-1} along the direction of z_i by a non-zero but small enough amount to a point z'_{i-1}, then the polygonal path with vertices $z_0, \ldots, z'_{i-1}, \ldots, z_n$ remains in $\mathscr{X}$ and $f(z'_{i-1}) \neq 0$. Thus, we may assume that $f(z_1) \neq 0$. But then the previous argument may be repeated to show that arbitrarily near $x = z_0$ there are points at which f is non-zero. □

Theorem 6.2. *Under the assumptions and notation of Theorem 5.1, let $\Theta \subset \mathbb{R}^d$ be non-empty, open, and connected, and let $\phi : \Theta \to \Omega^R_{LREM}$ be analytic and one-to-one.*

(i) If M has full column rank $n(n+m)(\kappa+\lambda+1)$ for some point in Θ then the ϕ-LREM is generically identified.

(ii) If there is a non-empty open subset of Θ on which M is rank deficient, then no point in $\phi(\Theta)$ is identified in Ω^R_{LREM}.

Proof. (i) We claim that the elements of M are real analytic functions of $\theta \in \Theta$. Since θ enters into M through the matrix P, it suffices to show that each coefficient matrix of C is an analytic function of θ. Every element of every coefficient matrix of B and A is analytic in θ by assumption. Using the fact that a composition of real analytic maps is real analytic (Krantz & Parks, 2002, Proposition 2.2.8), it is enough to show that each of the maps in the factorization (4) is real analytic. Clearly φ_2, φ_3, and φ_4 define real analytic mappings because they are compositions of real analytic mappings (multiplications, inversions, and projections). The fact that φ_1 is real analytic follows from a simple extension of Proposition X.1.2 of Clancey & Gohberg (1981). Next, since there is a point in Θ at which M has full column rank, M has a non-zero minor of order $n(n+m)(\kappa+\lambda+1)$. This minor is a real analytic function of $\theta \in \Theta$ and since it is not identically zero, it is generically non-zero by Lemma 6.1. Thus, M is generically of full column rank. Finally, Theorem 5.1 and the injectivity of ϕ imply that the ϕ-LREM is generically identified in Ω^R_{LREM}.

(ii) If M is rank deficient on an open subset of Θ, all its minors of order $n(n+m)(\kappa+\lambda+1)$ are equal to zero on this subset. By the preceding analysis, each of these minors is a real analytic mapping on Θ that vanishes on an open set. Thus, the set of points in Θ where these minors do not vanish is not dense in Θ and therefore Lemma 6.1 implies that all of these minors vanish identically over Θ. It follows that M is rank deficient at every point in Θ. By Theorem 5.1, no point in $\phi(\Theta)$ is identified in Ω^R_{LREM}. □

The advantage of Theorem 6.2 is that it provides very simple conditions for generic identification. One need only find a single point in Θ whose associated M matrix is of full rank to conclude generic identification.

Example 6.2. Consider the setting of Example 6.1. We have

$$\phi : (\theta_1, \theta_2, \theta_3) \mapsto (\theta_1 z^{-1} - ((\theta_3/\theta_2) + 1 + \theta_1) + z, \theta_2^{-1})$$

and $\Theta = \phi^{-1}(\Omega_{LREM})$. Let $\rho_1(\theta)$ and $\rho_2(\theta)$ be the zeros of $\theta_1 z^{-1} - ((\theta_3/\theta_2)+1+\theta_1)+z$ ordered as $|\rho_1(\theta)| < |\rho_2(\theta)|$ (condition (EU-LREM) is equivalent to $|\rho_1(\theta)| < 1 < |\rho_2(\theta)|$ for all $\theta \in \Theta$). Thus, the solution has the transfer function $C(z) = \frac{1}{\theta_2(z-\rho_2(\theta))}$. Condition (CF-LREM) then requires that $C(0) = -\frac{1}{\theta_2\rho_2(\theta)} > 0$. Given the continuity of ρ_1 and ρ_2, it follows that

$$\Theta = \left\{\theta \in \mathbb{R}^3 : |\rho_1(\theta)| < 1 < |\rho_2(\theta)|, \quad \frac{1}{\theta_2\rho_2(\theta)} < 0\right\}$$

is an open subset of $\mathbb{R}^3$. Now consider the matrix associated with the given restrictions

$$M = \begin{bmatrix} -C_0 & 0 & 0 & 1 & 0 & 0 \\ -C_1 & -C_0 & 0 & 0 & 1 & 0 \\ -C_2 & -C_1 & -C_0 & 0 & 0 & 1 \\ -C_3 & -C_2 & -C_1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Since $\det(M) = C_2^2 - C_1C_3$ and $C_j = -\frac{1}{\theta_2\rho_2^{j+1}}$, $\det(M)$ is identically zero so no point of $\phi(\Theta)$ is identified in Ω^R_{LREM}.

Suppose we restrict θ_1 to equal 1. We now have

$$\phi : (\theta_2, \theta_3) \mapsto (z^{-1} - ((\theta_3/\theta_2) + 2) + z, \theta_2^{-1})$$

and $\Theta = \phi^{-1}(\Omega_{LREM}) = \{\theta_2 > 0, 4\theta_2 + \theta_3 < 0\} \cup \{\theta_2 < 0, \theta_3 < 0\}$. We also have

$$M = \begin{bmatrix} -C_0 & 0 & 0 & 1 & 0 & 0 \\ -C_1 & -C_0 & 0 & 0 & 1 & 0 \\ -C_2 & -C_1 & -C_0 & 0 & 0 & 1 \\ -C_3 & -C_2 & -C_1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

It is easily seen that M is of full rank if and only if $C_2 = -\frac{1}{\theta_2\rho_2^3} \neq 0$, which is the case throughout the new Θ. However, we need not rely on this observation to conclude generic identifiability. The new Θ is the union of two open connected sets and so we may check the rank of M at any randomly chosen points in either component (e.g. $(\theta_2, \theta_3) = (-2, -1)$ and $(1, -5)$) to conclude by Theorem 6.2 that the ϕ-LREM is generically identified.

The example above makes clear that the conditions for the applicability of Theorem 6.2 can be difficult to verify. In particular, we suspect that under the conditions of Theorem 5.1 and with ϕ taken as the identity mapping, the ϕ-LREM is generically identified whenever any of its elements is identified. However, we have not succeeded in obtaining a parametrization of Ω^R_{LREM} that satisfies the conditions of Theorem 6.2 due to difficulties created by the canonical quasi-lower triangular assumption on $C(0)$ in (CF-LREM). Thus, this must be left for future research.

Generic identification for the i-th equation can be defined analogously. We say that the i-th equation of a ϕ-LREM is generically identified in Ω^R_{LREM} if there is a relatively open and dense subset $\Psi \subset \Theta$ such that for every parameter in $\phi(\Psi)$ the i-the equation is identified in Ω^R_{LREM}.

Theorem 6.3. *Under the assumptions and notation of Theorem 5.2, let $\Theta \subset \mathbb{R}^d$ be non-empty, open, and connected, and let $\phi : \Theta \to \Omega^R_{LREM}$ be analytic and one-to-one.*

(i) If M_i has full column rank $(n+m)(\kappa+\lambda+1)$ for some point in Θ then the i-th equation of the ϕ-LREM is generically identified.

(ii) If there is a non-empty open subset of Θ on which M_i is rank deficient, then at no point in $\phi(\Theta)$ is the i-th equation identified in Ω^R_{LREM}.

Proof. The proof is identical to that of Theorem 6.2 and is omitted. □

If the parameter space is restricted by nonlinear constraints it is generally difficult to obtain conditions for identification. However, it is possible to obtain necessary and sufficient conditions for local identification.

Theorem 6.4. *Let Ω^R_{LREM} be the set of $(B, A) \in \Omega_{LREM}$ satisfying*

$$R\left(\operatorname{vec}\left(\left[\begin{array}{cccccc} B_{-\lambda} & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{array}\right]\right)\right) = 0, \tag{7}$$

where $R : \mathbb{R}^{n^2(\kappa+\lambda+1)+nm(\kappa+1)} \mapsto \mathbb{R}^r$ is continuously differentiable. Let

$$\pi : \operatorname{vec}\left(\left[\begin{array}{cccccc} B_{-\lambda} & \cdots & B_\kappa & A^+_{-\lambda} & \cdots & A^+_\kappa \end{array}\right]\right) \mapsto \operatorname{vec}\left(\left[\begin{array}{cccccc} B_{-\lambda} & \cdots & B_\kappa & A^+_0 & \cdots & A^+_\kappa \end{array}\right]\right),$$

and let $\overline{R} = R \circ \pi$. If $(B, A) \in \Omega^R_{LREM}$ and P is defined as in Theorem 4.5, then (B, A) is locally identified in Ω^R_{LREM} if

$$M = \left[\begin{array}{c} P' \otimes I_n \\ \nabla \overline{R} \end{array}\right]$$

is of full column rank $n(n+m)(\kappa+\lambda+1)$*, where* $\nabla\overline{R}$ *is the Jacobian of* $\overline{R}$ *evaluated at* (B, A)*. Conversely, if* $(B, A) \in \Omega^R_{LREM}$*,* P *is defined as in Theorem 4.5, and*

$$\begin{bmatrix} P' \otimes I_n \\ \nabla\overline{R}(\tilde{B}, \tilde{A}) \end{bmatrix}$$

is of fixed rank lower than $n(n+m)(\kappa+\lambda+1)$ *for all* $(\tilde{B}, \tilde{A})$ *in a neighbourhood of* (B, A)*, then* (B, A) *is not locally identified in* Ω^R_{LREM}*.*

Proof. If (B, A) is not locally identified, there exists a sequence $(\tilde{B}_j, \tilde{A}_j) \in \Omega_{LREM}$ converging to (B, A) such that for all $j \geq 1$, $(\tilde{B}_j, \tilde{A}_j) \in (B, A)/\sim$, $(\tilde{B}_j, \tilde{A}_j) \in \Omega^R_{LREM}$, and $(\tilde{B}_j, \tilde{A}_j) \neq (B, A)$. Now define

$$\zeta_j = c_j \operatorname{vec}\left(\begin{bmatrix} \tilde{B}_{j,-\lambda} - B_{-\lambda} & \cdots & \tilde{B}_{j,\kappa} - B_\kappa & \tilde{A}^+_{j,-\lambda} - A^+_{-\lambda} & \cdots & \tilde{A}^+_{j,\kappa} - A^+_\kappa \end{bmatrix}\right),$$

where $c_j \in \mathbb{R}$ simply ensures that $\|\zeta_j\| = 1$ for all $j \geq 1$. Since $(\tilde{B}_j, \tilde{A}_j) \in (B, A)/\sim$, Theorem 4.5 (i) implies that $\zeta_j \in \ker(P' \otimes I_n)$ for all $j \geq 1$. On the other hand, since $(\tilde{B}_j, \tilde{A}_j) \in \Omega^R_{LREM}$ for all $j \geq 1$, $\nabla\overline{R}\zeta_j$ converges to zero. It follows that $M\zeta_j$ converges to zero. But this is impossible because $\|M\zeta_j\|$ is bounded below by the smallest singular value of M, which is non-zero because M is of full rank (Horn & Johnson, 1985, Theorems 7.3.5 and 7.3.10). Thus, (B, A) is locally identified.

Conversely, for $(\tilde{B}, \tilde{A}) \in \mathbb{R}[z, z^{-1}]^{n\times n} \times \mathbb{R}[z]^{n\times m}$ satisfying (L-LREM) define

$$Z(\tilde{B}, \tilde{A}) = \begin{bmatrix} (P' \otimes I_n)\operatorname{vec}\left(\begin{bmatrix} \tilde{B}_{-\lambda} & \cdots & \tilde{B}_\kappa & \tilde{A}^+_{-\lambda} & \cdots & \tilde{A}^+_\kappa \end{bmatrix}\right) \\ R(\tilde{B}, \tilde{A}) \end{bmatrix}$$

and notice that $Z(B, A) = 0$, $\nabla Z(B, A) = M$, and the rank of $\nabla Z(\tilde{B}, \tilde{A})$ is constant and equal to the rank of M in a neighbourhood of (B, A). It follows from the Rank Theorem (Rudin, 1976, Theorem 9.32) that every neighbourhood of (B, A) contains points different from (B, A) where Z vanishes. By Theorem 4.5 (ii), the zero set of Z coincides with $\Omega^R_{LREM} \cap (B, A)/\sim$ in a neighbourhood of (B, A) and as a results (B, A) is not locally identified in Ω^R_{LREM}. □

The reason why the converse in Theorem 6.4 requires more a stringent condition is well understood in the identification literature (Hsiao, 1983, Section 5.1). The condition is known as regularity. Without it, it is not possible to conclude local non-identifiability when M is rank deficient. For example, let $n = m = 1$, $\kappa = \lambda = 0$, and $R(B_0, A_0) = (A_0 - 1)^2$. Clearly every point of Ω^R_{LREM}, defined as in Theorem 6.4, is identified. On the other hand, M is equal to $\left[\begin{smallmatrix} -C_0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ at any given $(B,A) \in \Omega^R_{LREM}$. Note, however, that in this case, ∇Z is not of fixed rank in any neighbourhood of any $(B,A) \in \Omega^R_{LREM}$.

Local identification results for the i-th equation are similar.

Theorem 6.5. *Let Ω^R_{LREM} be the set of $(B,A) \in \Omega_{LREM}$ satisfying*

$$R_i\left(\operatorname{vec}\left(e_i'\left[\begin{array}{cccccc} B_{-\lambda} & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{array}\right]\right)\right) = 0, \tag{8}$$

where $R_i : \mathbb{R}^{n(\kappa+\lambda+1)+m(\kappa+1)} \mapsto \mathbb{R}^r$ is continuously differentiable and $e_i \in \mathbb{R}^n$ is the i-th standard unit vector. Let

$$\pi_i : \operatorname{vec}\left(e_i'\left[\begin{array}{cccccc} B_{-\lambda} & \cdots & B_\kappa & A^+_{-\lambda} & \cdots & A^+_\kappa \end{array}\right]\right) \mapsto \operatorname{vec}\left(e_i'\left[\begin{array}{cccccc} B_{-\lambda} & \cdots & B_\kappa & A^+_0 & \cdots & A^+_\kappa \end{array}\right]\right),$$

and let $\overline{R}_i = R_i \circ \pi_i$. If $(B,A) \in \Omega^R_{LREM}$ and P is defined as in Theorem 4.5, then the i-th equation of (1) is locally identified at (B,A) in Ω^R_{LREM} if

$$M_i = \left[\begin{array}{c} P' \\ \nabla\overline{R}_i \end{array}\right]$$

is of full column rank $(n+m)(\kappa+\lambda+1)$, where $\nabla\overline{R}_i$ is the Jacobian of $\overline{R}_i$ evaluated at (B,A). Conversely, if $(B,A) \in \Omega^R_{LREM}$, P is defined as in Theorem 4.5, and

$$\left[\begin{array}{c} P' \\ \nabla\overline{R}_i(\tilde{B},\tilde{A}) \end{array}\right]$$

is of fixed rank lower than $(n+m)(\kappa+\lambda+1)$ for all $(\tilde{B},\tilde{A})$ in a neighbourhood of (B,A), then the i-the equation of (1) is is not locally identified at (B,A) in Ω^R_{LREM}.

Proof. If the i-th equation of (B,A) is not locally identified, there exists a sequence $(\tilde{B}_j,\tilde{A}_j) \in \Omega_{LREM}$ converging to (B,A) such that for all $j \geq 1$, $(\tilde{B}_j,\tilde{A}_j) \in (B,A)/\sim$, $(\tilde{B}_j,\tilde{A}_j) \in \Omega^R_{LREM}$, and the i-th equations of $(\tilde{B}_j,\tilde{A}_j)$ and (B,A) are different. Now define

$$\zeta_j = c_j \operatorname{vec}\left(\left[\begin{array}{cccccc} \tilde{B}_{j,-\lambda} - B_{-\lambda} & \cdots & \tilde{B}_{j,\kappa} - B_\kappa & \tilde{A}^+_{j,-\lambda} - A^+_{-\lambda} & \cdots & \tilde{A}^+_{j,\kappa} - A^+_\kappa \end{array}\right]\right),$$

where $c_j \in \mathbb{R}$ simply ensures that $\|(I_{(n+m)(\kappa+\lambda+1)} \otimes e_i')\zeta_j\| = 1$ for all $j \geq 1$. Since $(\tilde{B}_j,\tilde{A}_j) \in (B,A)/\sim$, Theorem 4.5 (i) implies that $\zeta_j \in \ker(P' \otimes I_n)$ for all $j \geq 1$. This implies that $(I_{(n+m)(\kappa+\lambda+1)} \otimes e_i')\zeta_j \in \ker(P')$, since $P'(I_{(n+m)(\kappa+\lambda+1)} \otimes e_i') = (I_{m(1+(n+1)\kappa+\lambda)} \otimes e_i')(P' \otimes I_n)$.

On the other hand, since $(\tilde{B}_j, \tilde{A}_j) \in \Omega^R_{LREM}$ for all $j \geq 1$, $\nabla \overline{R}(I_{(n+m)(\kappa+\lambda+1)} \otimes e_i')\zeta_j$ converges to zero. It follows that $M_i(I_{(n+m)(\kappa+\lambda+1)} \otimes e_i')\zeta_j$ converges to zero. But this is impossible because $\|M_i(I_{(n+m)(\kappa+\lambda+1)} \otimes e_i')\zeta_j\|$ is bounded below by the smallest singular value of M_i, which is non-zero because M_i is of full rank (Horn & Johnson, 1985, Theorems 7.3.5 and 7.3.10). Thus, the i-the equation is locally identified at (B, A) in Ω^R_{LREM}.

Conversely, for $(\tilde{B}, \tilde{A}) \in \mathbb{R}[z, z^{-1}]^{n\times n} \times \mathbb{R}[z]^{n\times m}$ satisfying (L-LREM) define

$$Z_i(\tilde{B}, \tilde{A}) = \begin{bmatrix} P'\text{vec}\left(e_i'\begin{bmatrix} \tilde{B}_{-\lambda} & \cdots & \tilde{B}_\kappa & \tilde{A}^+_{-\lambda} & \cdots & \tilde{A}^+_\kappa \end{bmatrix}\right) \\ R_i(\tilde{B}, \tilde{A}) \end{bmatrix}$$

and notice that $Z_i(B, A) = 0$, $\nabla Z_i(B, A) = M_i$, and the rank of $\nabla Z_i(\tilde{B}, \tilde{A})$ is constant and equal to the rank of M_i in a neighbourhood of (B, A). It follows from the Rank Theorem (Rudin, 1976, Theorem 9.32) that every neighbourhood of the parameters of the i-the equation of (B, A) contains points where Z_i vanishes. Thus, in every neighbourhood of (B, A) we can find a $(\tilde{B}, \tilde{A})$ that is identical to (B, A) except in the i-th equation and contained in $\Omega^R_{LREM} \cap (B, A)/\sim$. Thus, the i-th equation of (B, A) is not locally identified in Ω^R_{LREM}. $\square$

7 Conclusion

This paper's title is an hommage to the seminal paper of the VARMA identification literature (Hannan, 1971). Like Hannan's paper, the present work characterizes observational equivalence and provides conditions for identification in a variety of settings. More importantly, and again much like Hannan's paper, the present work has not succeeded in answering all of the questions surrounding the identification of the model under study. We now turn to some of the pending issues.

Recognizing the difficulty of the identification problem for VARMA models, the literature proposed a variety of canonical parametrizations (Hannan & Deistler, 2012, p. 67). These parametrizations allow the researcher to specify a model without having to worry about identification. It would be quite useful for empirical work to find similar parametrizations for LREMs.

Our framework has excluded measurement errors, which are commonly used in the LREM literature. This is not an insurmountable challenge as the literature on latent variables and measurement error is very well developed (Bollen, 1989; Fuller, 1987). Treating this material in the present work would have made it prohibitively long and complicated, not to mention distracting from the primary theoretical difficulties of the identification problem for LREMs. Thus, this is also left for further work.

It is tempting to think of the condition number of M or M_i from Sections 5 and 6 as measures of distance to non-identifiability. However, this remains to be proven. It would be interesting to look more closely at this problem especially in light of the weak identification of LREMs frequently encountered in empirical work.

Finally, the quasi-lower triangular assumption on the first impulse response in (CF-LREM) has allowed us to solve a long standing open problem, the identification problem for LREMs, up to another long standing open problem, the general identification problem for simultaneous equations models. This progress was made possible by recognizing that the appropriate mathematical framework for LREMs is Wiener-Hopf factorization theory. Interestingly, Hannan (1971) had a similar trajectory. As Hannan describes how he came about resolving the identification problem for VARMA,

> "It was really quite simple once you recognize what the underlying mathematical technique is. ... That's how it came about ... it is important to have in command the mathematics so you can solve the problem. Of course, the 64 dollar question is which mathematics to learn, because you can't learn all of it." (Pagan & Hannan, 1985, p. 273)

The 64 dollar question now is which new mathematics will allow us to solve the general identification problem for simultaneous equations models.

A A Review of Classical Identification Theory

This section reviews some of the basic theory of identification in linear systems, beginning with the classical simultaneous equations model and proceeding to VARMA models. For more detailed treatments, the reader may wish to consult Hsiao (1983) or Hannan & Deistler (2012).

A.1 The Classical Simultaneous Equation Model

The classical simultaneous equations model is the set of structural equations

$$BY_t = A\varepsilon_t, \qquad t \in \mathbb{Z}. \tag{9}$$

Here ε is an m-dimensional exogenous and unobserved i.i.d. sequences of mean zero and $\mathrm{var}(\varepsilon_0) = I_m$, while Y an n-dimensional endogenous observed sequence.

The parameter space of the classical simultaneous equations model, denoted by Ω_{SEM}, is a set of pairs

$$(B, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$$

characterized by two restrictions, the first of which is:

$$\mathrm{rank}(B) = n. \tag{EU-SEM}$$

Restriction (EU-SEM) is equivalent to the existence and uniqueness of the solution,

$$Y_t = B^{-1}A\varepsilon_t, \qquad t \in \mathbb{Z}.$$

The variance matrix of the observed data then satisfies

$$\mathrm{var}(Y_0) = B^{-1}AA'B^{-1\prime}.$$

Before introducing the second restriction, we must understand why it is needed. To that end, we say that two parameters (B, A) and $(\tilde{B}, \tilde{A})$ are *observationally equivalent* and denote this by $(B, A) \sim (\tilde{B}, \tilde{A})$ if both produce the same $\mathrm{var}(Y_0)$; that is, if and only if

$$\tilde{B}^{-1}\tilde{A}\tilde{A}'\tilde{B}^{-1\prime} = B^{-1}AA'B^{-1\prime}.$$

In order to make progress on this problem it is necessary to impose further restrictions. One such restriction requires $B^{-1}A$ to be of full column rank so that there are no redundant shocks in the system. This then implies the existence of an orthogonal matrix $V \in \mathbb{R}^{m\times m}$ such that

$$\tilde{B}^{-1}\tilde{A} = B^{-1}AV.$$

Now it is certainly possible to formulate the identification problem at this level of generality. However, this makes the problem substantially more difficult. We will opt, as most of the literature has done, for imposing additional restrictions on the parameter space in order to eliminate the dependence on V. In particular, the QR algorithm allows us to reduce $B^{-1}AV$ to a *canonical quasi-lower triangular form*, where the first non-zero element of the j-th column is positive and occurs on row i_j, with $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ (Anderson et al., 2012, Theorem 1). Note that when $n = m$, the canonical quasi-lower triangular form is the Cholesky factor of $\text{var}(Y_0)$. We will need the following additional restriction on the parameter space.

(CF-SEM) $\quad$ $B^{-1}A$ is of rank m and canonical quasi-lower triangular.

Thus Ω_{SEM} is the set of pairs $(B, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$ satisfying (EU-SEM) and (CF-SEM). We will endow it with the relative topology inherited from $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$. Some aspects of its topology are given in the following result

Proposition A.1. *Ω_{SEM} is homeomorphic to an open subset of $\mathbb{R}^{n(n+m)-\frac{1}{2}m(m-1)}$ consisting of two connected components.*

Proof. Let

$$\begin{aligned}\Theta_{SEM} = \big\{(B, C) : B \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{n\times m},& \\ \text{rank}(B) = n,& \\ \text{rank}(C) = m,& \\ \text{and } C \text{ is canonical quasi-lower triangular}\big\}.&\end{aligned}$$

Then Θ_{SEM} can be viewed as a subset of $\mathbb{R}^{n(n+m)-\frac{1}{2}m(m-1)}$ and the mapping $\phi_{SEM} : \Theta_{SEM} \to \Omega_{SEM}$ defined by $(B, C) \mapsto (B, BC)$ is a homeomorphism. Since the smallest singular value of a matrix is continuous with respect to the elements of the matrix (Horn & Johnson, 1991, Theorem 3.3.16), Θ_{SEM} is an open subset of $\mathbb{R}^{n(n+m)-\frac{1}{2}m(m-1)}$. The set of non-singular $n \times n$ matrices consists of two components, one containing I_n and another containing $\left[\begin{smallmatrix} -1 & 0 \\ 0 & I_{n-1} \end{smallmatrix}\right]$ (Hall, 2003, Proposition 1.12). In turn, the set of canonical quasi-lower triangular $n \times m$ matrices of full rank has a single component as every such element is connected by a straight line to $\left[\begin{smallmatrix} I_m \\ 0_{(n-m)\times m} \end{smallmatrix}\right]$. Thus, Θ_{SEM} consists of two connected components. $\square$

Proposition A.1 implies that Ω_{SEM} can be parametrized by $n(n + m) - \frac{1}{2}m(m - 1)$ free parameters.

We have already characterized observational equivalence as follows.

Theorem A.2. *Let* $(B, A), (\tilde{B}, \tilde{A}) \in \Omega_{SEM}$*. Then* $(\tilde{B}, \tilde{A}) \sim (B, A)$ *if and only if*

$$\tilde{B}^{-1}\tilde{A} = B^{-1}A. \tag{10}$$

We need a simpler characterization of spaces of observationally equivalent parameters denoted by

$$(B, A)/\sim = \left\{(\tilde{B}, \tilde{A}) \in \Omega_{SEM} : (\tilde{B}, \tilde{A}) \sim (B, A)\right\}, \qquad (B, A) \in \Omega_{SEM}.$$

That is the purpose of the next result.

Theorem A.3. *Let* $(B, A), (\tilde{B}, \tilde{A}) \in \Omega_{SEM}$ *and let*

$$C = B^{-1}A, \qquad\qquad P = \begin{bmatrix} -C \\ I_m \end{bmatrix}.$$

Then:

(i) $(\tilde{B}, \tilde{A}) \sim (B, A)$ *if and only if*

$$\text{vec}\left(\begin{bmatrix} \tilde{B} & \tilde{A} \end{bmatrix}\right) \in \ker\left(P' \otimes I_n\right).$$

(ii) $(B, A)/\sim$ *is a relatively open and dense subset of the subspace*

$$\text{mat}\left(\ker\left(P' \otimes I_n\right)\right),$$

where

$$\text{mat} : \text{vec}\left(\begin{bmatrix} B & A \end{bmatrix}\right) \mapsto (B, A).$$

(iii) $\dim\left((B, A)/\sim\right) = n^2$.

Proof. (i) By Theorem A.2, $(\tilde{B}, \tilde{A}) \sim (B, A)$ if and only if

$$-\tilde{B}C + \tilde{A} = 0_{n\times m}.$$

Vectorizing both sides we obtain

$$\underbrace{\begin{bmatrix} -C' \otimes I_n & I_{nm} \end{bmatrix}}_{P' \otimes I_n} \text{vec}\left(\begin{bmatrix} \tilde{B} & \tilde{A} \end{bmatrix}\right) = 0_{nm\times 1}.$$

(ii) If $(\breve{B}, \breve{A}) \in \text{mat}(\ker(P' \otimes I_n))$ then the preceding implies (CF-SEM) is satisfied if (EU-SEM) is satisfied. Therefore, $(B, A)/\sim$ is the intersection of $\text{mat}(\ker(P' \otimes I_n))$ with

$$\left\{ (\breve{B}, \breve{A}) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} : \text{(EU-SEM) is satisfied} \right\}.$$

Since the latter set is open in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$, $(B, A)/\sim$ is relatively open in $\text{mat}(\ker(P' \otimes I_n))$. If $(\breve{B}, \breve{A}) \in \text{mat}(\ker(P' \otimes I_n))$ and $\det(\breve{B}) = 0$, then arbitrarily near $\breve{B}$ we can find a non-singular $\bar{B}$ (Horn & Johnson, 1985, p. 312) and we can then define $\bar{A} = \bar{B}C$, which is then also arbitrarily near $\breve{A}$. It follows that $(B, A)/\sim$ is dense in $\text{mat}(\ker(P' \otimes I_n))$.

(iii) $\dim(\ker(P')) = n$ and so $\dim(\ker(P' \otimes I_n)) = n^2$ by the standard properties of Kronecker products (Horn & Johnson, 1991, Theorem 4.2.15). □

Theorem A.3 characterizes the spaces of observationally equivalent parameters as relatively open and dense subsets of subspaces of the parameter space. As the origin is not an element of Ω_{SEM}, these subsets are proper.

Theorem A.3 makes it clear that the parameter space needs to be further restricted in order to be able to identify a single parameter with a given $\text{var}(Y_0)$. Let Ω^R_{SEM} be a subset of Ω_{SEM} endowed with the relative topology. We say that $(B, A) \in \Omega^R_{SEM}$ is *identified in* Ω^R_{SEM} if every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in Ω^R_{SEM} is equal to (B, A). We say that a parameter (B, A) is *locally identified in* Ω^R_{SEM} if it has a neighbourhood N in Ω^R_{SEM} such that every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in N is equal to (B, A). Clearly, a parameter is locally identified in Ω^R_{SEM} if it is identified in Ω^R_{SEM} but the converse is not true in general.

Theorem A.4. *Let* Ω^R_{SEM} *be the set of* $(B, A) \in \Omega_{SEM}$ *satisfying*

$$R \,\text{vec}\left(\begin{bmatrix} B & A \end{bmatrix} \right) = u, \tag{11}$$

where $R \in \mathbb{R}^{r\times n(m+n)}$ *and* $u \in \mathbb{R}^r$. *If* $(B, A) \in \Omega^R_{SEM}$ *and* P *is defined as in Theorem A.3, then* (B, A) *is identified in* Ω^R_{SEM} *if and only if*

$$M = \begin{bmatrix} P' \otimes I_n \\ R \end{bmatrix}$$

is of full column rank $n(n + m)$.

Proof. Let $\zeta = \text{vec}\left(\begin{bmatrix} B & A \end{bmatrix} \right)$. If M is of full column rank, then ζ is the only point in $\ker(P' \otimes I_n)$ that satisfies (11). Theorem A.3 (i) then implies that (B, A) is identified in

Ω^R_{SEM}. If M is not of full rank, then there exists $0 \neq \xi \in \ker(P' \otimes I_n) \cap \ker(R)$. If $c > 0$ is sufficiently small, Theorem A.3 (ii) implies that $(B, A) \sim \mathrm{mat}(\zeta + c\xi) \neq (B, A)$ and since $\mathrm{mat}(\zeta + c\xi)$ satisfies (11), (B, A) is not identified in Ω^R_{SEM}. □

The geometry of Theorem A.4 is illustrated in Figure 1. Any given parameter (B, A) lies in the intersection of two affine subspaces. The first affine subspace, denoted by $\mathcal{E}$, contains the set of parameters observationally equivalent to (B, A) by Theorem A.3 (i). The second affine subspace is the space Ω^R_{SEM}, denoted by $\mathcal{R}$, which contains the set of parameters satisfying the restrictions (11). When (B, A) is the only point of intersection then it is identified in Ω^R_{SEM}. Otherwise, the two affine subspaces intersect along an affine a subspace, which contains a line segment in Ω^R_{SEM} by Theorem A.3 (ii) and so every neighbourhood of (B, A) contains infinitely many observationally equivalent parameters that also satisfy the given restrictions. Thus, *for the classical simultaneous equations model subject to affine restrictions, a parameter is identified if and only if it is locally identified.*

Suppose now that we are interested in identifying just the i-th equation of (9). Let Ω^R_{SEM} be as before and let $(B, A) \in \Omega^R_{SEM}$. We say that the i-th equation of (9) is *identified at* (B, A) *in* Ω^R_{SEM} if every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in Ω^R_{SEM} has the same i-th equation as (B, A). We say that the i-th equation of (9) is *locally identified at* (B, A) *in* Ω^R_{SEM} if (B, A) has a neighbourhood N in Ω^R_{SEM} such that every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in N has the same i-th equation as (B, A). Clearly, if all equations are (locally) identified at (B, A) in Ω^R_{SEM}, then (B, A) is (locally) identified in Ω^R_{SEM}.

Theorem A.5. *Let Ω^R_{SEM} be the set of $(B, A) \in \Omega_{SEM}$ satisfying*

$$R_i \operatorname{vec}\left(e_i' \begin{bmatrix} B & A \end{bmatrix}\right) = u_i, \tag{12}$$

where $R_i \in \mathbb{R}^{r \times (m+n)}$, $u_i \in \mathbb{R}^r$, and $e_i \in \mathbb{R}^n$ is the i-th standard unit vector. If $(B, A) \in \Omega^R_{SEM}$ and P is defined as in Theorem A.3, then the i-th equation of (9) is identified at (B, A) in Ω^R_{SEM} if and only if

$$M_i = \begin{bmatrix} P' \\ R_i \end{bmatrix}$$

is of full column rank $n + m$.

Proof. Let $\zeta = \text{vec}\left(\begin{bmatrix} B & A \end{bmatrix}\right)$. If M_i is of full column rank, then $\text{vec}\left(e_i' \begin{bmatrix} B & A \end{bmatrix}\right) = (I_{n+m} \otimes e_i')\zeta$ is the only point in $\ker(P')$ that satisfies (12), since $P'(I_{n+m} \otimes e_i') = (I_m \otimes e_i')(P' \otimes I_n)$. Theorem A.3 (i) then implies that any parameter in Ω^R_{SEM} that is observationally equivalent to (B, A) must have the same i-th equation as (B, A). Thus the i-th equation is identified at (B, A) in Ω^R_{SEM}. If M_i is not of full rank, then there exists $0 \neq \xi_i \in \ker(P') \cap \ker(R_i)$. If $c > 0$ is sufficiently small, Theorem A.3 (ii) implies that $(B, A) \sim \text{mat}(\zeta + c\xi_i \otimes e_i)$ and since $\text{mat}(\zeta + c\xi_i \otimes e_i)$ satisfies (12) but has a different i-th equation than (B, A), the i-th equation is not identified at (B, A) in Ω^R_{SEM}. □

The geometry of Theorem A.5 is exactly analogous to that of Theorem A.4, as is the equivalence of identification and local identification.

A.2 The Classical VARMA Model

The classical VARMA model generalizes (9) by allowing for dependence across time (i.e. dynamics).

$$\sum_{i=0}^{p} B_i Y_{t-i} = \sum_{i=0}^{k} A_i \varepsilon_{t-i}, \qquad t \in \mathbb{Z}, \tag{13}$$

Here ε is an m-dimensional exogenous and unobserved i.i.d. sequences of mean zero and $\text{var}(\varepsilon_0) = I_m$, while Y an n-dimensional endogenous observed sequence.

It will be convenient to collect the parameters $B_0, \ldots, B_p \in \mathbb{R}^{n\times n}$ and $A_0, \ldots, A_k \in \mathbb{R}^{n\times m}$ in the form of polynomial matrices $B(z) = \sum_{i=0}^{p} B_i z^i$ and $A(z) = \sum_{i=0}^{k} A_i z^i$. The parameter space of the VARMA model, denoted by Ω_{VARMA}, is then a set of pairs

$$(B, A) \in \mathbb{R}[z]^{n\times n} \times \mathbb{R}[z]^{n\times m}$$

characterized by three restrictions, the first of which is:

$$\text{rank}(B(z)) = n \text{ for all } z \in \overline{\mathbb{D}}. \tag{EU-VARMA}$$

Restriction (EU-VARMA) is equivalent to the existence and uniqueness of a stationary causal solution,

$$Y_t = B^{-1}(L)A(L)\varepsilon_t, \qquad t \in \mathbb{Z},$$

where L is the lag operator, $B^{-1}(L)$ is the composition of L and the Taylor series expansion of $B^{-1}(z)$ in a neighbourhood of $z = 0$, and $A(L)$ is the composition of L and A (Hannan & Deistler, 2012, pp. 10-11). This implies that the spectral density of the observed data,

$$f_{YY}(z) = \sum_{j=-\infty}^{\infty} \operatorname{cov}(Y_j, Y_0) z^j,$$

satisfies

$$f_{YY}(z) = B^{-1}(z)A(z)A'(z^{-1})B^{-1\prime}(z^{-1}).$$

We say that two parameters (B, A) and $(\tilde{B}, \tilde{A})$ are *observationally equivalent* and denote this by $(\tilde{B}, \tilde{A}) \sim (B, A)$ if both produce the same spectral density; that is, if and only if

$$B^{-1}(z)A(z)A'(z^{-1})B^{-1\prime}(z^{-1}) = \tilde{B}^{-1}(z)\tilde{A}(z)\tilde{A}'(z^{-1})\tilde{B}^{-1\prime}(z^{-1}).$$

Just as in the classical simultaneous equations model, in order to make progress here it is necessary to impose further restrictions. One such restriction requires the transfer function $B^{-1}(z)A(z)$ to be of full rank for all $z \in \mathbb{D}$ so that every shock can be reconstructed from the present and past values of Y. This condition is known variably in the literature as the *invertibility*, *fundamentalness*, or *minimum phase* condition. Proceeding then, it is well known in the literature that observational equivalence holds if and only if there exists an orthogonal matrix $V \in \mathbb{R}^{m \times m}$ such that

$$\tilde{B}^{-1}\tilde{A} = B^{-1}AV.$$

See e.g. Theorems 4.6.8 and 4.6.11 of Lindquist & Picci (2015). We may then eliminate V just as we did in the classical simultaneous equations model. Thus, we arrive at the second restriction on all $(B, A) \in \Omega_{VARMA}$,

(CF-VARMA)

$$\operatorname{rank}\left(B^{-1}(z)A(z)\right) = m \text{ for all } z \in \mathbb{D} \text{ and } B^{-1}(0)A(0) \text{ is canonical quasi-lower triangular.}$$

If (EU-VARMA) is maintained, then (CF-VARMA) is equivalent to the more familiar restriction that $A(z)$ have rank m for all $z \in \mathbb{D}$ and $B^{-1}(0)A(0)$ be canonical quasi-lower triangular (Anderson et al., 2016). However, our formulation is more convenient for LREM applications.

Note also that when $n = m$, assumption (CF-VARMA) sets $B^{-1}(0)A(0)$ equal to the Cholesky factor of the variance of the innovations of Y. Thus, it corresponds to the identification strategy commonly attributed to Sims (1980).

Even though we have not yet completed our characterization of Ω_{VARMA}, we can already simplify observational equivalence based on conditions (EU-VARMA) and (CF-VARMA).

Theorem A.6. *Let $(B, A), (\tilde{B}, \tilde{A}) \in \Omega_{VARMA}$. Then $(B, A) \sim (\tilde{B}, \tilde{A})$ if and only if*

$$\tilde{B}^{-1}\tilde{A} = B^{-1}A. \tag{14}$$

The set of points in $\mathbb{R}[z]^{n\times n} \times \mathbb{R}[z]^{n\times m}$ satisfying (EU-VARMA) and (CF-VARMA) is infinite dimensional. The sets of observationally equivalent parameters, as described in Theorem A.6, are also infinite dimensional. Therefore, in practice one usually specifies a non-negative integer κ such that for every $(B, A) \in \Omega_{VARMA}$,

$$\max\deg(B) \leq \kappa, \qquad \max\deg(A) \leq \kappa. \tag{L-VARMA}$$

Thus, Ω_{VARMA} is the set of pairs $(B, A) \in \mathbb{R}[z]^{n\times n} \times \mathbb{R}[z]^{n\times m}$ satisfying (EU-VARMA), (CF-VARMA), and (L-VARMA). The condition (L-VARMA) allows us to think of Ω_{VARMA} as a subset of $\mathbb{R}^{n(n+m)(\kappa+1)}$ with the topology induced by Euclidean topology.

Proposition A.7. *Ω_{VARMA} is homeomorphic to a subset of $\mathbb{R}^{n(n+m)(1+\kappa)-\frac{1}{2}m(m-1)}$, the interior of which consists of two connected components.*

Proof. Let

$$\Theta_{VARMA} = \Bigg\{(B_0, \ldots, B_\kappa, C_0, A_1, \ldots, A_\kappa) : B_0, \ldots, B_\kappa \in \mathbb{R}^{n\times n}, C_0, A_1, \ldots, A_\kappa \in \mathbb{R}^{n\times m},$$
$$\text{rank}\left(\sum_{i=0}^{\kappa} B_i z^i\right) = n \text{ for all } z \in \overline{\mathbb{D}},$$
$$C_0 \text{ is canonical quasi-lower triangular, and}$$
$$\text{rank}\left(B_0C_0 + \sum_{i=1}^{\kappa} A_i z^i\right) = m \text{ for all } z \in \mathbb{D}\Bigg\}.$$

Then Θ_{VARMA} can be viewed as a subset of $\mathbb{R}^{n(n+m)(1+\kappa)-\frac{1}{2}m(m-1)}$ and the mapping $\phi_{VARMA} : \Theta_{VARMA} \to \Omega_{VARMA}$ defined by

$$(B_0, \ldots, B_\kappa, C_0, A_1, \ldots, A_\kappa) \mapsto \left(\sum_{i=0}^{\kappa} B_i z^i, B_0C_0 + \sum_{i=1}^{\kappa} A_i z^i\right)$$

is a homeomorphism. Now let

$$\Theta^{\circ}_{VARMA} = \left\{ (B_0, \ldots, B_\kappa, C_0, A_1, \ldots, A_\kappa) \in \Theta_{VARMA} : \text{rank}\left(B_0 C_0 + \sum_{i=1}^{\kappa} A_i z^i \right) = m \text{ for all } z \in \mathbb{T} \right\}.$$

If we pick any point $(B_0, \ldots, B_\kappa, C_0, A_1, \ldots, A_\kappa) \in \Theta_{VARMA} \backslash \Theta^{\circ}_{VARMA}$, then for any $\rho > 1$, the point $(B_0, \ldots, B_\kappa, C_0, \rho A_1, \ldots, \rho^\kappa A_\kappa) \notin \Theta_{VARMA}$ because rank $\left(B_0 C_0 + \sum_{i=1}^{\kappa} A_i (\rho z)^i\right)$ will fall below m for some point in $\mathbb{D}$. Thus $\Theta_{VARMA} \backslash \Theta^{\circ}_{VARMA}$ are boundary points. In contrast, the continuity of zeros of a polynomial with respect to its coefficients (Horn & Johnson, 1985, Appendix D) ensures that Θ°_{VARMA} is open. Thus Θ°_{VARMA} is the interior of Θ_{VARMA}. By similar reasoning, $(B_0, (1-t)B_1, \ldots, (1-t)^\kappa B_\kappa, C_0, (1-t)A_1, \ldots, (1-t)^\kappa A_\kappa)$ is in Θ°_{VARMA} for any $t \in [0, 1]$. Thus, $(B_0, B_1, \ldots, B_\kappa, C_0, A_1, \ldots, A_\kappa) \in \Theta^{\circ}_{VARMA}$ is in the same connected component as $(B_0, 0, \ldots, 0, C_0, 0, \ldots, 0)$, which in turn falls into one of two components by Proposition A.1. □

The interior set referred to in Proposition A.7 is the set that parametrizes the *strictly invertible* processes in Ω_{VARMA}. Thus, the parameter space may be thought to have a boundary consisting of those elements of Ω_{VARMA}, which are invertible but not strictly invertible. Proposition A.7 is the analogue to Theorem 2.5.3 (ii) of Hannan & Deistler (2012). Note, however, that Hannan and Deistler's parametrization is canonical (i.e. they parametrize the equivalence classes of parameters) and ours is not.

The condition (L-VARMA) also simplifies observationally equivalence. For suppose $(B, A) \sim (\tilde{B}, \tilde{A})$ and let

$$C = B^{-1} A.$$

Then (14) can be rewritten as

$$\tilde{A} = \tilde{B} C.$$

If $\tilde{B}(z) = \sum_{j=0}^{\kappa} \tilde{B}_i z^i$, $\tilde{A}(z) = \sum_{i=0}^{\kappa} \tilde{A}_i z^i$, and $C(z) = \sum_{j=0}^{\infty} C_i z^i$ for $z \in \overline{\mathbb{D}}$, then equating Taylor series coefficients we arrive to the following equivalent expression

$$\begin{bmatrix} \tilde{A}_0 & \cdots & \tilde{A}_\kappa & 0 & \cdots \end{bmatrix} = \begin{bmatrix} \tilde{B}_0 & \cdots & \tilde{B}_\kappa & 0 & \cdots \end{bmatrix} \begin{bmatrix} C_0 & C_1 & C_2 & C_3 & \cdots \\ 0 & C_0 & C_1 & C_2 & \ddots \\ 0 & 0 & C_0 & C_1 & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Although there are infinitely many equations in this expression, (L-VARMA) will allow us to restrict attention to only the first $1+(n+1)\kappa$ equations, which will then allow us to determine the dimension of the sets of observationally equivalent models. This is achieved in the next result. Before we state it formally, we review the concepts of coprimeness and McMillan degree. If $B \in \mathbb{R}[z]^{n\times n}$, $A \in \mathbb{R}[z]^{n\times m}$, and $\det(B)$ is not identically zero, we say that the pair (B, A) is *coprime* if $\operatorname{rank}\left(\left[\begin{array}{cc} B(z) & A(z) \end{array}\right]\right) = n$ for all $z \in \mathbb{C}$. Every $C \in \mathbb{R}(z)^{n\times m}$ can be represented as $C = B^{-1}A$ for some coprime pair (B, A) and

$$\delta(C) = \max \deg(\det(B)) \tag{15}$$

is an invariant of such representations of C known as the *McMillan degree* of C (Hannan & Deistler, 2012, pp. 41 and 51).

Lemma A.8. *Let $(B, A) \in \Omega_{VARMA}$, let $C = B^{-1}A$ have a Taylor series expansion $C(z) = \sum_{i=0}^{\infty} C_i z^i$ in a neighbourhood of $z = 0$, and let*

$$H = \begin{bmatrix} C_{\kappa+1} & C_{\kappa+2} & \cdots & C_{(n+1)\kappa} \\ \ddots & \ddots & \ddots & \vdots \\ C_2 & \ddots & \ddots & C_{n\kappa+1} \\ C_1 & C_2 & \ddots & C_{n\kappa} \end{bmatrix}.$$

Then:

(i) $\operatorname{rank}(H) = \delta(C(z^{-1}) - C(0)) \leq n\kappa$*.*

(ii) The set of parameters satisfying $\operatorname{rank}(H) = n\kappa$ *is generic (i.e. contains an open and dense subset of Ω_{VARMA}).*

Proof. (i) The rank of the infinite Hankel matrix

$$\begin{bmatrix} C_1 & C_2 & C_3 & \iddots \\ C_2 & C_3 & C_4 & \iddots \\ C_3 & C_4 & C_5 & \iddots \\ \iddots & \iddots & \iddots & \iddots \end{bmatrix}$$

is equal to $\delta(C(z^{-1}) - C(0))$ (Hannan & Deistler, 2012, Theorem 2.4.1 (iii)). Now write

$$\begin{aligned} C(z^{-1}) - C(0) &= B^{-1}(z^{-1})\left(A(z^{-1}) - B(z^{-1})C(0)\right) \\ &= \left(z^{\kappa}B(z^{-1})\right)^{-1}\left(z^{\kappa}A(z^{-1}) - z^{\kappa}B(z^{-1})C(0)\right). \end{aligned}$$

Restriction (L-VARMA) implies that $z^\kappa B(z^{-1}) \in \mathbb{R}[z]^{n\times n}$ and $z^\kappa A(z^{-1}) \in \mathbb{R}[z]^{n\times m}$. Since $B_0 = B(0)$ is non-singular by restriction (EU-VARMA), $\max\deg\left(\det\left(z^\kappa B(z^{-1})\right)\right) = n\kappa$ (Hannan & Deistler, 2012, p. 42). It follows that $\delta(C(z^{-1}) - C(0)) \leq n\kappa$ (Hannan & Deistler, 2012, Lemma 2.2.1 (e)). By Theorem 2.4.1 (iii) of Hannan & Deistler (2012) again and reordering the blocks, $\delta(C(z^{-1}) - C(0))$ is the rank of the matrix

$$Q = \begin{bmatrix} C_{n\kappa} & C_{n\kappa+1} & \cdots & C_{2n\kappa-1} \\ \ddots & \ddots & \ddots & \vdots \\ C_2 & \ddots & \ddots & C_{n\kappa+1} \\ C_1 & C_2 & \ddots & C_{n\kappa} \end{bmatrix}.$$

Since $A = BC$ is a polynomial matrix of degree at most κ, the $\kappa+1, \kappa+2, \ldots, 2n\kappa-1$ coefficient matrices of BC are all zero. That is,

$$\sum_{i=0}^{\kappa} B_i C_{j-i} = 0, \qquad \kappa+1 \leq j \leq 2n\kappa - 1.$$

Since B_0 is non-singular,

$$C_j = -B_0^{-1} \sum_{i=1}^{\kappa} B_i C_{j-i}, \qquad \kappa+1 \leq j \leq 2n\kappa - 1.$$

This implies that all of the top blocks of Q are linear dependent on the bottom κ blocks of Q. This implies that

$$\delta(C(z^{-1}) - C(0)) = \mathrm{rank}(Q) = \mathrm{rank}(H) = \mathrm{rank}(\check{H}),$$

where

$$\check{H} = \begin{bmatrix} C_{\kappa} & C_{\kappa+1} & \cdots & C_{(n+1)\kappa-1} \\ \ddots & \ddots & \ddots & \vdots \\ C_2 & \ddots & \ddots & C_{n\kappa+1} \\ C_1 & C_2 & \ddots & C_{n\kappa} \end{bmatrix}$$

will be needed later.

(ii) The proof is in two steps.

STEP 1: *The set of coprime $(B, A) \in \Omega_{VARMA}$ with non-singular B_κ is generic in Ω_{VARMA}.*

Our technique for proving this follows the technique used to prove Theorem 3 of Anderson et al. (2016) although our proof is substantially more explicit. We are not able to rely on that result directly because of our additional assumption that $C(0)$ is quasi-lower triangular. Recall that

$$\Omega_{VARMA} = \Big\{(B, A) \in \mathbb{R}[z]^{n\times n} \times \mathbb{R}[z]^{n\times m} : \max\deg(B), \max\deg(A) \leq \kappa,$$
$$\operatorname{rank}(B) = n \text{ for all } z \in \overline{\mathbb{D}},\ \operatorname{rank}(A) = m \text{ for all } z \in \mathbb{D},$$
$$\text{and } B^{-1}(0)A(0) \text{ is canonical quasi-lower triangular}\Big\}.$$

Recall also that the topology on this set is the topology inherited from $\mathbb{R}^{n(n+m)(\kappa+1)}$. We will prove that the following subset is open and dense in Ω_{VARMA},

$$\check{\Omega}_{VARMA} = \Big\{(B, A) \in \Omega_{VARMA} : \det(B) \text{ has distinct zeros}, \det(B_\kappa) \neq 0,$$
$$\operatorname{rank}(A) = m \text{ for all } z \in \mathbb{T},\ \text{and } (B, A) \text{ is coprime}\Big\}.$$

To see that $\check{\Omega}_{VARMA}$ is open, let $(B, A) \in \check{\Omega}_{VARMA}$. We will construct a neighbourhood of (B, A) in $\check{\Omega}_{VARMA}$ as $N = N_1 \cap N_2 \cap N_3 \cap N_4$, where N_i is a neighbourhood of (B, A) satisfying the i-th additional condition in $\check{\Omega}_{VARMA}$.

By the continuity of the zeros of a polynomial with respect to its coefficients (Horn & Johnson, 1985, Appendix D), there is a neighbourhood N_1 of (B, A) such that for every $(\check{B}, \check{A}) \in N_1$, $\det(\check{B})$ has distinct zeros.

Since $\det(B_\kappa) \neq 0$, the continuity of the determinant implies that there is a neighbourhood N_2 of (B, A) such that for every $(\check{B}, \check{A}) \in N_2$, $\det(\check{B}_\kappa) \neq 0$.

Since $\operatorname{rank}(A) = m$ for all $z \in \overline{\mathbb{D}}$, it has a minor of order m that has no zeros in $\overline{\mathbb{D}}$. Using the continuity of the zeros of a polynomial with respect to its coefficients again, there is a neighbourhood N_3 of (B, A) such that for every $(\check{B}, \check{A}) \in N_3$, $\operatorname{rank}(\check{A}) = m$ for all $z \in \overline{\mathbb{D}}$.

Finally, since $\det(B)$ has distinct zeros and B_κ is non-singular, $\det(B)$ has $n\kappa$ distinct zeros $z_1, \ldots, z_{n\kappa} \in \mathbb{C}$. The fact that (B, A) is coprime then implies that

$$\Pi_i A(z_i) \neq 0, \qquad i = 1, \ldots, n\kappa,$$

where Π_i is the orthogonal projection matrix onto the left null space of $B(z_i)$. Since

$$\operatorname{rank}(B(z_i)) = n - 1, \qquad i = 1, \ldots, n\kappa,$$

small perturbations to $B(z_i)$ that leave the rank fixed at $n-1$ lead to small perturbations to Π_i (Gohberg et al., 2006, Theorem 13.5.1). Thus, there exists a neighbourhood N_4 of (B, A)

such that for every $(\check{B}, \check{A}) \in N_4$,

$$\check{\Pi}_i \check{A}(\check{z}_i) \neq 0, \qquad i = 1, \ldots, n\kappa,$$

where $\check{z}_1, \ldots, \check{z}_{n\kappa}$ are the zeros of $\det(\check{B})$ and $\check{\Pi}_i$ is the orthogonal projection matrix onto the left null space of $\check{B}(\check{z}_i)$. In other words, every $(\check{B}, \check{A}) \in N_4$ is coprime.

To see that $\check{\Omega}_{VARMA}$ is dense, let $(B, A) \in \Omega_{VARMA} \backslash \check{\Omega}_{VARMA}$. We propose an infinitesimal perturbation of B, followed by an infinitesimal perturbation A that leads to a point in $\check{\Omega}_{VARMA}$. Thus, any neighbourhood of (B, A) in Ω_{VARMA} will contain an element of $\check{\Omega}_{VARMA}$.

If $\det(B)$ has a zero of multiplicity greater than 1, we claim that there exists an infinitesimal perturbation of B into the set of polynomial matrices satisfying (EU-VARMA) and (L-VARMA) and whose determinants have distinct zeros. In the course of proving their Theorem 3, Anderson et al. (2016) prove that an infinitesimal perturbation exists indeed into the set of polynomial matrices satisfying (L-VARMA) whose determinants have distinct zeros. The claim then follows from the fact that the set of polynomial matrices satisfying (EU-VARMA) and (L-VARMA) is an open subset of its ambient space $\mathbb{R}^{n^2(\kappa+1)}$ by the continuity of zeros of polynomials with respect to their coefficients.

Next, if B_κ is singular, we can infinitesimally perturb its singular values at zero to obtain a non-singular matrix. By the continuity of zeros of polynomials with respect to their coefficients, this infinitesimal perturbation does not interfere with the $\det(B)$ having distinct zeros or condition (EU-VARMA).

Next, if $\text{rank}(A(z_0)) < m$ for some $z_0 \in \mathbb{T}$ then for $\rho < 1$ arbitrarily near 1, the polynomial matrix $A(\rho z)$ is such that $\text{rank}(A(\rho z)) = m$ for all $z \in \overline{\mathbb{D}}$. This perturbation has no effect on the rank of the transfer function in $\mathbb{D}$ and leaves the constant coefficient matrix invariant. Thus, the perturbation leaves condition (CF-VARMA) invariant.

Finally, suppose that after the sequence of infinitesimal perturbations above we arrive at a $(B, A) \in \Omega_{VARMA}$ that is not coprime. Enumerate the zeros of $\det(B)$ as $z_1, \ldots, z_{n\kappa}$ and choose non-zero vectors $v_1, \ldots, v_{n\kappa} \in \mathbb{C}^n$ spanning the left null spaces of $B(z_1), \ldots, B(z_{n\kappa})$ respectively. Since (B, A) is not coprime, there is an index i such that $v_i' A(z_i) = 0$. Choose $\Delta \in \mathbb{R}^{n \times m}$ satisfying $v_i' \Delta \neq 0$ for all i such that $v_i' A(z_i) = 0$. Then an infinitesimal perturbation of A_κ in the direction of Δ is sufficient to produce a coprime element pair. This infinitesimal perturbation has no effect on the condition that $\text{rank}(A(z))$ for all $z \in \overline{\mathbb{D}}$ by the continuity of zeros of polynomials with respect to their coefficients and it leaves $A(0)$ invariant. Thus (CF-VARMA) remains invariant.

STEP 2: *For all parameters in* $\check{\Omega}_{VARMA}$, $\operatorname{rank}(H) = n\kappa$.

We have already established that the matrix $\check{H}$ encountered in (i) is of rank $\delta(C(z^{-1}) - C(0)) \leq n\kappa$. If $\operatorname{rank}(\check{H}) < n\kappa$, then there exist vectors $x_i \in \mathbb{R}^n$, $i = 0, \ldots, \kappa - 1$, not all zero, such that

$$(x'_0, \ldots, x'_{\kappa-1})\check{H} = 0_{1\times nm\kappa}.$$

Setting

$$x(z) = \sum_{i=0}^{\kappa-1} x_i z^i,$$

this implies that the terms of degree $\kappa, \kappa+1, \ldots, (n+1)\kappa - 1$ of $x'C$ vanish. To see that indeed all higher degree terms vanish as well, notice that each element of the numerator of

$$y' = x'C = \frac{x'\operatorname{adj}(B)A}{\det(B)}$$

is expressible as a polynomial of degree bounded above by $\max\deg(x) + \max\deg(\operatorname{adj}(B)) + \max\deg(A) \leq (\kappa - 1) + (n-1)\kappa + \kappa = (n+1)\kappa - 1$. Since $\det(B(0)) \neq 0$ by (EU-VARMA), it follows from Lemma 4.3 that $y \in \mathbb{R}^m[z]$ and $\max\deg(y) \leq \kappa - 1$. Now setting

$$U = \begin{bmatrix} x'B^{-1} \\ S \end{bmatrix}$$

with $S \in \mathbb{R}^{(n-1)\times n}$ chosen so that $\det(U)$ is not identically zero (e.g. choose $z_0 \in \mathbb{D}$ such that $x(z_0) \neq 0$ and choose S as an orthogonal complement to $x'(z_0)B^{-1}(z_0)$). Then

$$\dot{B} = UB = \begin{bmatrix} x' \\ SB \end{bmatrix} \in \mathbb{R}[z]^{n\times n}, \qquad \dot{A} = UA = \begin{bmatrix} y' \\ SA \end{bmatrix} \in \mathbb{R}[z]^{n\times m},$$

and $\dot{B}^{-1}\dot{A} = B^{-1}A = C$. But this violates the minimality of $\delta(C) = n\kappa$ among all matrix fraction descriptions of C because $\max\deg(\dot{B}) < n\kappa$ (Hannan & Deistler, 2012, Lemma 2.2.1 (e)). Thus, $\operatorname{rank}(\check{H}) = n\kappa$ and therefore $\operatorname{rank}(H) = n\kappa$. □

We are now in a position to simplify Theorem A.6 and characterize the set

$$(B, A)/\sim = \left\{(\tilde{B}, \tilde{A}) \in \Omega_{VARMA} : (\tilde{B}, \tilde{A}) \sim (B, A)\right\}, \qquad (B, A) \in \Omega_{VARMA}.$$

Theorem A.9. *Let* $(B, A), (\tilde{B}, \tilde{A}) \in \Omega_{VARMA}$*, let* $C = B^{-1}A$ *have a Taylor series expansion* $C(z) = \sum_{i=0}^{\infty} C_i z^i$ *in a neighbourhood of* $z = 0$*, and let*

$$T = \begin{bmatrix} C_0 & C_1 & \cdots & C_\kappa \\ 0 & C_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & C_1 \\ 0 & \cdots & 0 & C_0 \end{bmatrix}, \qquad H = \begin{bmatrix} C_{\kappa+1} & C_{\kappa+2} & \cdots & C_{(n+1)\kappa} \\ \ddots & \ddots & \ddots & \vdots \\ C_2 & \ddots & \ddots & C_{n\kappa+1} \\ C_1 & C_2 & \ddots & C_{n\kappa} \end{bmatrix},$$

$$P = \begin{bmatrix} -T & -H \\ I_{m(\kappa+1)} & 0_{m(\kappa+1)\times nm\kappa} \end{bmatrix}.$$

Then:

(i) $(\tilde{B}, \tilde{A}) \sim (B, A)$ *if and only if*

$$\mathrm{vec}\left(\begin{bmatrix} \tilde{B}_0 & \cdots & \tilde{B}_\kappa & \tilde{A}_0 & \cdots & \tilde{A}_\kappa \end{bmatrix}\right) \in \ker\left(P' \otimes I_n\right).$$

(ii) $(B, A)/\sim$ *is a relatively open subset of the subspace*

$$\mathrm{mat}\left(\ker\left(P' \otimes I_n\right)\right),$$

where

$$\mathrm{mat} : \mathrm{vec}\left(\begin{bmatrix} B_0 & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right) \mapsto \left(\sum_{i=0}^{\kappa} B_i z^i, \sum_{i=0}^{\kappa} A_i z^i\right).$$

(iii) $\dim\left((B, A)/\sim\right) = n^2(\kappa + 1) - n\,\delta\left(C(z^{-1}) - C(0)\right) \geq n^2$ *and for generic points in the parameter space* $\dim\left((B, A)/\sim\right) = n^2$*.*

Proof. (i) If $(\tilde{B}, \tilde{A}) \sim (B, A)$, then Theorem A.6 implies that

$$0 = -\tilde{B}C + \tilde{A} = \frac{-\tilde{B}\mathrm{adj}(B)A + \det(B)\tilde{A}}{\det(B)}.$$

Each element of the right hand side can be expressed as a ratio of a polynomial (of degree at most $\max\left\{\max\deg(\tilde{B}) + \max\deg(\mathrm{adj}(B)) + \max\deg(A), \max\deg(\det(B)) + \max\deg(\tilde{A})\right\} \leq \max\{\kappa + (n-1)\kappa + \kappa, n\kappa + \kappa\} = (n+1)\kappa$) and $\det(B)$, which satisfies $\det(B(0)) \neq 0$ by (EU-VARMA). By Lemma 4.3, this is equivalent to the first $1 + (n+1)\kappa$ Taylor series coefficients equating to zero. Thus, observational equivalence is equivalent to

$$-\begin{bmatrix} \tilde{B}_0 & \ldots & \tilde{B}_\kappa \end{bmatrix}\begin{bmatrix} T & H \end{bmatrix} + \begin{bmatrix} \tilde{A}_0 & \ldots & \tilde{A}_\kappa & 0_{n\times m} & \ldots & 0_{n\times m} \end{bmatrix} = 0_{n\times(1+(n+1)\kappa)m}$$

or equivalently,

$$\begin{bmatrix} \tilde{B}_0 & \cdots & \tilde{B}_\kappa & \tilde{A}_0 & \cdots & \tilde{A}_\kappa \end{bmatrix} P = 0_{n\times(1+(n+1)\kappa)m}.$$

Vectorizing we obtain

$$(P' \otimes I_n)\,\mathrm{vec}\left(\begin{bmatrix} \tilde{B}_0 & \cdots & \tilde{B}_\kappa & \tilde{A}_0 & \cdots & \tilde{A}_\kappa \end{bmatrix}\right) = 0_{nm(1+(n+1)\kappa)\times 1}.$$

(ii) If $(\breve{B}, \breve{A}) \in \mathrm{mat}(\ker(P' \otimes I_n))$ then it satisfies (L-VARMA). If, additionally, it satisfies (EU-VARMA), the preceding implies that $\breve{A} = \breve{B}C$ and therefore (CF-VARMA) is satisfied. Thus, $(B, A)/\sim$ is the intersection of $\mathrm{mat}(\ker(P' \otimes I_n))$ with

$$\left\{(\breve{B}, \breve{A}) \in \mathbb{R}[z]^{n\times n} \times \mathbb{R}[z]^{n\times m} : \text{(EU-VARMA) and (L-VARMA) are satisfied}\right\}.$$

The latter set is open in $\mathbb{R}^{n(n+m)(\kappa+1)}$ due to the continuity of zeros of polynomials with respect to their coefficients (Horn & Johnson, 1985, Appendix D). Therefore, $(B, A)/\sim$ is relatively open in $\mathrm{mat}(\ker(P' \otimes I_n))$.

(iii) $\dim(\ker(P')) = \dim(\ker(H'))$ and so the result follows from Lemma A.8 (i) and the standard properties of Kronecker products (Horn & Johnson, 1991, Theorem 4.2.15). For generic parameters the result follows from Lemma A.8 (ii). □

Theorem A.9 is a direct generalization of Theorem A.3 to the VARMA setting. Theorem A.9 (i) characterizes the sets of observationally equivalent parameters as subsets of particular subspaces of the parameter space determined by the first $1 + (n + 1)\kappa$ impulse responses. This result is equivalent to Theorem 1 of Deistler & Schrader (1979) although Deistler and Schrader use the more traditional formulation of observational equivalence, $\tilde{B} = UB$ and $\tilde{A} = UA$ for some $U \in \mathbb{R}(z)^{n\times n}$. Theorem A.9 (ii) then shows that the sets of observationally equivalent parameters constitute relatively open although not necessarily dense subsets of the aforementioned subspaces (see Example A.1). Theorem A.9 (iii) finally characterizes the dimension of observationally equivalent parameters. Theorem A.9 (ii) and (iii) are analogues to Theorem 2.5.3 (v) of Hannan & Deistler (2012) (see also their Remark 2 on page 67). As noted earlier, however, Hannan and Deistler parametrize the VARMA model differently.

Example A.1. Let $n = m = \kappa = 1$ and suppose $(B, A) = (1, 1)$, then

$$P' \otimes I_n = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Clearly $(1, -2, 1, -2)' \in \ker(P' \otimes I_n)$. Thus, $(\tilde{B}, \tilde{A}) = (1 - 2z, 1 - 2z) \in \mathrm{mat}(\ker(P' \otimes I_n))$. However, this pair violates condition (EU-VARMA) and by the continuity of zeros of polynomials with respect to their coefficients (Horn & Johnson, 1985, Appendix D), there is a neighbourhood of $(\tilde{B}, \tilde{A})$ in $\mathbb{R}^4$ containing no parameters. Thus, $(B, A)/\sim$ is is not dense in $\ker(P' \otimes I_n)$.

We are now in a position to consider affine restrictions. Let Ω^R_{VARMA} be a subset of Ω_{VARMA} endowed with the relative topology. We say that $(B, A) \in \Omega^R_{VARMA}$ is *identified in* Ω^R_{VARMA} if every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in Ω^R_{VARMA} is equal to (B, A). We say that a parameter (B, A) is *locally identified in* Ω^R_{VARMA} if it has a neighbourhood N in Ω^R_{VARMA} such that every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in N is equal to (B, A). Again, a parameter is locally identified in Ω^R_{VARMA} if it is identified in Ω^R_{VARMA} but the converse is not true in general.

Theorem A.10. *Let Ω^R_{VARMA} be the set of $(B, A) \in \Omega_{VARMA}$ satisfying*

$$R \,\mathrm{vec}\left(\begin{bmatrix} B_0 & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right) = u, \tag{16}$$

where $R \in \mathbb{R}^{r \times n(n+m)(\kappa+1)}$ and $u \in \mathbb{R}^r$. If $(B, A) \in \Omega^R_{VARMA}$ and P is defined as in Theorem A.9, then (B, A) is identified in Ω^R_{VARMA} if and only if

$$M = \begin{bmatrix} P' \otimes I_n \\ R \end{bmatrix}$$

is of full column rank $n(n + m)(\kappa + 1)$.

Proof. Let $\zeta = \mathrm{vec}\left(\begin{bmatrix} B_0 & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right)$. If M is of full column rank, then ζ is the only point in $\ker(P' \otimes I_n)$ that satisfies (16). Theorem A.9 (i) then implies that (B, A) is identified in Ω^R_{VARMA}. If M is not of full rank, then there exists $0 \neq \xi \in \ker(P' \otimes I_n) \cap \ker(R)$. If $c > 0$ is sufficiently small, Theorem A.9 (ii) implies that $(B, A) \sim \mathrm{mat}(\zeta + c\xi) \neq (B, A)$ and since $\mathrm{mat}(\zeta + c\xi)$ satisfies (16), (B, A) is not identified in Ω^R_{VARMA}. □

Theorem A.10 is a direct generalization of Theorem A.4. The geometry of Theorem A.10 is also exactly analogous to that of Theorem A.4. Thus, *for the classical VARMA model subject to affine restrictions, a parameter is identified if and only if it is locally identified.*

Corollary A.11 (Deistler & Schrader (1979))**.** *Let*

$$E : \operatorname{vec}\left(\begin{bmatrix} Y_0 & \cdots & Y_{1+(n+1)\kappa} & X_0 & \cdots & X_{1+(n+1)\kappa} \end{bmatrix}\right) \mapsto \operatorname{vec}\left(\begin{bmatrix} Y_0 & \cdots & Y_\kappa & X_0 & \cdots & X_\kappa \end{bmatrix}\right),$$

where $Y_0, \ldots, Y_{1+(n+1)\kappa} \in \mathbb{R}^{n\times n}$ and $X_0, \ldots, X_{1+(n+1)\kappa} \in \mathbb{R}^{n\times m}$, let $E_\perp$ be an orthogonal complement to E, let

$$R_{DS} = \begin{bmatrix} RE \\ E_\perp \end{bmatrix},$$

and let

$$D = \left.\begin{bmatrix} B_0 & \cdots & B_\kappa & 0 & \cdots & 0 & A_0 & \cdots & A_\kappa & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & & \vdots & 0 & \ddots & & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 & \vdots & \ddots & \ddots & & \ddots & 0 \\ \vdots & & \ddots & \ddots & & B_\kappa & \vdots & & \ddots & \ddots & & A_\kappa \\ \vdots & & & \ddots & \ddots & \vdots & \vdots & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & B_0 & 0 & \cdots & \cdots & \cdots & 0 & A_0 \end{bmatrix}\right\} 1 + (n+1)\kappa \textit{ blocks.}$$

Then (B, A) is identified in Ω^R_{VARMA} if and only if $R_{DS}(D' \otimes I_n)$ is of full column rank $n^2(1 + (n+1)\kappa)$.

Proof. We claim that $\left[\begin{smallmatrix} P'\otimes I_n \\ R \end{smallmatrix}\right]$ is of full column rank if and only if $R_{DS}(D' \otimes I_n)$ is of full column rank. Let

$$\zeta = \operatorname{vec}\left(\begin{bmatrix} Y_0 & \cdots & Y_\kappa & X_0 & \cdots & X_\kappa \end{bmatrix}\right).$$

Then $(P' \otimes I_n)\zeta = 0$ if and only if

$$\begin{bmatrix} Y_0 & \cdots & Y_\kappa \end{bmatrix}\begin{bmatrix} T & H \end{bmatrix} = \underbrace{\begin{bmatrix} X_0 & \cdots & X_\kappa & 0_{n\times m} & \cdots & 0_{n\times m} \end{bmatrix}}_{1+(n+1)\kappa \text{ blocks}}.$$

This is equivalent to

$$\begin{bmatrix} Y_0 & \cdots & Y_{(n+1)\kappa} \end{bmatrix} \underbrace{\begin{bmatrix} C_0 & C_1 & \cdots & \cdots & C_{n(\kappa+1)} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_1 \\ 0 & \cdots & \cdots & 0 & C_0 \end{bmatrix}}_{T_{DS}} = \begin{bmatrix} X_0 & \cdots & X_{(n+1)\kappa} \end{bmatrix}$$

$$E_\perp \underbrace{\operatorname{vec}\left(\begin{bmatrix} Y_0 & \cdots & Y_{(n+1)\kappa} & X_0 & \cdots & X_{(n+1)\kappa} \end{bmatrix}\right)}_{\zeta_{DS}} = 0_{n^2(n+m)\kappa\times 1}.$$

Thus, the kernel of $\left[\begin{smallmatrix} P'\otimes I_n \\ R \end{smallmatrix}\right]$ is $\{0\}$ if and only if the kernel of $\left[\begin{smallmatrix} P'_{DS}\otimes I_n \\ R_{DS} \end{smallmatrix}\right]$ is $\{0\}$, where

$$P_{DS} = \begin{bmatrix} -T_{DS} \\ I_{m(1+(n+1)\kappa)} \end{bmatrix}.$$

Since $A = BC$, D is an orthogonal complement to P'_{DS}. Thus, $(P'_{DS}\otimes I_n)\zeta_{DS} = 0$ if and only if $\zeta_{DS} = (D'\otimes I_n)\xi_{DS}$ for some $\xi_{DS} \in \mathbb{R}^{n^2(1+(n+1)\kappa)}$. It follows that $\left[\begin{smallmatrix} P'_{DS}\otimes I_n \\ R_{DS} \end{smallmatrix}\right]\zeta_{DS} = 0$ if and only if $\zeta_{DS} = (D'\otimes I_n)\xi_{DS}$ and $R_{DS}(D'\otimes I_n)\xi_{DS} = 0$. In other words, the kernel of $\left[\begin{smallmatrix} P'_{DS}\otimes I_n \\ R_{DS} \end{smallmatrix}\right]$ is $\{0\}$ if and only if the kernel of $R_{DS}(S'\otimes I_n)$ is $\{0\}$. □

Corollary A.11 due to Deistler & Schrader (1979) is evidently equivalent to Theorem A.10. The main difference between the two formulations is that our determining matrix is populated by impulse responses, whereas Deistler and Schrader's matrix is populated by structural parameters.

Suppose now that we are interested in identifying just the i-th equation of (13). Let Ω^R_{VARMA} be as before and let $(B, A) \in \Omega^R_{VARMA}$. We say that the i-th equation of (13) is *identified at* (B, A) *in* Ω^R_{VARMA} if every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in Ω^R_{VARMA} has the same i-th equation as (B, A). We say that the i-th equation of (13) is *locally identified at* (B, A) *in* Ω^R_{VARMA} if (B, A) has a neighbourhood N in Ω^R_{VARMA} such that every $(\tilde{B}, \tilde{A}) \sim (B, A)$ in N has the same i-th equation as (B, A). Again, if all equations are (locally) identified at (B, A) in Ω^R_{VARMA}, then (B, A) is (locally) identified in Ω^R_{VARMA}.

Theorem A.12. *Let* Ω^R_{VARMA} *be the set of* $(B, A) \in \Omega_{VARMA}$ *satisfying*

$$R_i \operatorname{vec}\left(e'_i \begin{bmatrix} B_0 & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right) = u_i, \tag{17}$$

where $R_i \in \mathbb{R}^{r\times(n+m)(\kappa+1)}$, $u_i \in \mathbb{R}^r$, *and* $e_i \in \mathbb{R}^n$ *is the* i*-th standard unit vector. If* $(B, A) \in \Omega^R_{VARMA}$ *and* P *is defined as in Theorem A.9, then the* i*-th equation of* (13) *is identified at* (B, A) *in* Ω^R_{VARMA} *if and only if*

$$M_i = \begin{bmatrix} P' \\ R_i \end{bmatrix}$$

is of full column rank $(n + m)(\kappa + 1)$*.*

Proof. Let $\zeta = \text{vec}\left(\begin{bmatrix} B_0 & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right)$. If M_i is of full column rank, then $\text{vec}\left(e_i'\begin{bmatrix} B_0 & \cdots & B_\kappa & A_0 & \cdots & A_\kappa \end{bmatrix}\right) = (I_{(n+m)(\kappa+1)} \otimes e_i')\zeta$ is the only point in $\ker(P')$ that satisfies (17), since $P'(I_{(n+m)(\kappa+1)} \otimes e_i') = (I_{m(1+(n+1)\kappa)} \otimes e_i')(P' \otimes I_n)$. Theorem A.9 (i) then implies that any parameter in Ω^R_{VARMA} that is observationally equivalent to (B, A) must have the same i-th equation as (B, A). Thus the i-th equation is identified at (B, A) in Ω^R_{VARMA}. If M_i is not of full rank, then there exists $0 \neq \xi_i \in \ker(P') \cap \ker(R_i)$. If $c > 0$ is sufficiently small, Theorem A.9 (ii) implies that $(B, A) \sim \text{mat}(\zeta + c\xi_i \otimes e_i)$ and since $\text{mat}(\zeta + c\xi_i \otimes e_i)$ satisfies (17) but has a different i-th equation than (B, A), the i-th equation is not identified at (B, A) in Ω^R_{VARMA}. □

References

Al-Sadoon, M. M. (2018). The Linear Systems Approach to Linear Rational Expectations Models. *Econometric Theory*, *34*(03), 628–658.

Anderson, B. D., Deistler, M., Chen, W., & Filler, A. (2012). Autoregressive Models of Singular Spectral Matrices. *Automatica*, *48*(11), 2843 – 2849.

Anderson, B. D., Deistler, M., Felsenstein, E., & Koelbl, L. (2016). The Structure of Multivariate AR and ARMA Systems: Regular and Singular Systems; the Single and the Mixed Frequency Case. *Journal of Econometrics*, *192*(2), 366 – 373.

Bacchiocchi, E. & Kitagawa, T. (2019). The Dark Side of the SVAR: a Trip Into the Local Identification World. Technical report.

Blanchard, O. (2018). On the Future of Macroeconomic Models. *Oxford Review of Economic Policy*, *34*(1-2), 43–54.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York. A Wiley-Interscience Publication.

Canova, F. & Sala, L. (2009). Back to Square One: Identification Issues in DSGE Models. *Journal of Monetary Economics*, *56*(4), 431–449.

Clancey, K. F. & Gohberg, I. (1981). *Factorization of Matrix Functions and Singular Integral Operators.* Operator Theory: Advances and Applications (Vol. 3). Boston, USA: Birkhäuser Verlag Basel.

Deistler, M. & Schrader, J. (1979). Linear Models with Autocorrelated Errors: Structural Identifiability in the Absence of Minimality Assumptions. *Econometrica*, *47*(2), 495–504.

Dufour, J.-M. & Renault, E. (1998). Short Run and Long Run Causality in Time Series: Theory. *Econometrica*, *66*(5), 1099–1125.

Fuller, W. A. (1987). *Measurement Error Models.* Wiley Series in Probability and Mathematical Statistics. New York, USA: John Wiley & Sons, Inc.

Gohberg, I., Lancaster, P., & Rodman, L. (2006). *Invariant Subspaces of Matrices with Applications.* Classics in Applied Mathematics. Philadelphia, USA: SIAM.

Hall, B. C. (2003). *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.* Graduate Texts in Mathematics. NewYork, USA: Springer.

Hannan, E. J. (1971). The Identification Problem for Multiple Equation Systems with Moving Average Errors. *Econometrica*, *39*(5), 751–765.

Hannan, E. J. & Deistler, M. (2012). *The Statistical Theory of Linear Systems.* Classics in Applied Mathematics. Philadelphia, PA, USA: SIAM.

Hansen, L. P. & Sargent, T. J. (1981). Formulating and Estimating Dynamic Linear Rational Expectations Models. In R. E. Lucas & T. J. Sargent (Eds.), *Rational Expectations and Econometric Practice* (pp. 91–125). Minneapolis: The University of Minnesota Press.

Herbst, E. P. & Schorfheide, F. (2016). *Bayesian Estimation of DSGE Models.* Princeton, USA: Princeton University Press.

Horn, R. A. & Johnson, C. R. (1985). *Matrix Analysis.* Cambridge, United Kingdom: Cambridge University Press.

Horn, R. A. & Johnson, C. R. (1991). *Topics in Matrix Analysis.* Cambridge, United Kingdom: Cambridge University Press.

Hsiao, C. (1983). Identification. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of Econometrics*, volume 1 of *Handbook of Econometrics* chapter 4, (pp. 223–283). Elsevier.

Iskrev, N. (2010). Local Identification in DSGE Models. *Journal of Monetary Economics*, *57*(2), 189–202.

Kociecki, A. & Kolasa, M. (2018). Global Identification of Linearized DSGE Models. *Quantitative Economics*, *9*(3), 1243–1263.

Komunjer, I. & Ng, S. (2011). Dynamic Identification of Dynamic Stochastic General Equilibrium Models. *Econometrica*, *79*(6), 1995–2032.

Krantz, S. G. & Parks, H. R. (2002). *A Primer of Real Analytic Functions* (2 ed.). Birkhäuser Advanced Texts: Basler Lehrbücher. Boston, USA: Birkhäuser Boston.

Lindquist, A. & Picci, G. (2015). *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation, and Identification.* Series in Contemporary Mathematics 1. Berlin Heidelberg: Springer-Verlag.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis.* Berlin, Germany: Springer.

Muth, J. F. (1981). Estimation of Economic Relationships Containing Latent Expectations Variables. In R. E. Lucas & T. J. Sargent (Eds.), *Rational Expectations and Econometric Practice* (pp. 321–328). Minneapolis: The University of Minnesota Press.

Onatski, A. (2006). Winding Number Criterion for Existence and Uniqueness of Equilibrium in Linear Rational Expectations Models. *Journal of Economic Dynamics and Control*, *30*(2), 323–345.

Pagan, A. & Hannan, E. J. (1985). The ET Interview: Professor E. J. Hannan. *Econometric Theory*, *1*(2), 263–289.

Pesaran, M. (1981). Identification of Rational Expectations Models. *Journal of Econometrics*, *16*(3), 375 – 398.

Pesaran, M. H. & Smith, R. P. (2011). Beyond the DSGE Straitjacket. *The Manchester School*, *79*(s2), 5–16.

Qu, Z. & Tkachenko, D. (2017). Global Identification in DSGE Models Allowing for Indeterminacy. *The Review of Economic Studies*, *84*(3), 1306–1345.

Romer, P. (2016). The Trouble with Macroeconomics. *The American Economist*, *20*, 1–20.

Rosenblatt, M. (2000). *Gaussian and Non-Gaussian Linear Time Series and Random Fields.* Springer Series in Statistics. New York, USA: Springer Verlag.

Rubio-Ramírez, J. F., Waggoner, D. F., & Zha, T. (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *The Review of Economic Studies*, *77*(2), 665–696.

Rudin, W. (1976). *Principles of Mathematical Analysis* (3 ed.). New York, USA: McGraw Hill, Inc.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*(1), 1–48.

Wallis, K. F. (1980). Econometric Implications of the Rational Expectations Hypothesis. *Econometrica*, *48*(1), 49–73.