Econometrics and Statistics 4 (2017) 31–38

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Non-identifiability of VMA and VARMA systems in the mixed frequency case

Manfred Deistler[a,*], Lukas Koelbl[b], Brian D.O. Anderson[c,d]

[a] Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria
[b] Accenture Digital, Accenture GmbH, Vienna, Austria
[c] Research School of Engineering, Australian National University, Canberra, Australia
[d] Canberra Research Laboratory, National ICT Australia Ltd., Australia

A R T I C L E I N F O

Article history:
Received 24 March 2016
Revised 21 November 2016
Accepted 23 November 2016
Available online 10 December 2016

Keywords:
VARMA
VMA
Mixed frequency
Non-identifiability

A B S T R A C T

Recently, identifiability results for VAR systems in the context of mixed frequency data have been shown in a number of papers. These results have been extended to VARMA systems, where the MA order is smaller than or equal to the AR order. Here, it is shown that in the VMA case and in the VARMA case, where the MA order exceeds the AR order, results are completely different. Then, for the case, where the innovation covariance matrix is non-singular, "typically" non-identifiability occurs – not even local identifiability. This is due to the fact that, e.g., in the VMA case, as opposed to the VAR case, the not directly observed autocovariances of the output can vary "freely". In the singular case, i.e., when the innovation covariance matrix is singular, things may be different.

1. Introduction

In many applications we encounter mixed frequency (MF) data in multivariate time series, i.e., the univariate component series often are available only at different sampling frequencies. A common and simple solution is to use the data only at the lowest sampling frequency and to run the statistical algorithms with these single frequency data. Clearly, such an approach is associated with a loss of information. For a number of reasons, such as the joint modeling of real and financial data, the problem of information extraction from mixed frequency data has attracted increasing attention (see Nijman, 1985; Ghysels et al., 2007; Wohlrabe, 2008; Marcellino and Schumacher, 2010; Foroni and Marcellino, 2013; Koelbl et al., 2016). There are several settings and different approaches in dealing with the problem of estimation from MF data, depending, for instance, on the model classes considered, the observational schemes producing the data and the estimation procedures used.

In this contribution we deal with linear stochastic models of ARMA type and mainly with stock data. Our focus is on identifiability of the underlying high frequency system generating the outputs at the highest frequency available and on the well-posedness of the relation between the second moments of the outputs which can be directly observed and the underlying parameters of the high frequency system. An analogous problem for time continuous systems has been analyzed in Chambers (2016).

* Corresponding author.
E-mail addresses: manfred.deistler@tuwien.ac.at (M. Deistler), lukas.koelbl@accenture.com (L. Koelbl), brian.anderson@anu.edu.au (B.D.O. Anderson).

http://dx.doi.org/10.1016/j.ecosta.2016.11.006

For the VAR case g(eneric)-identifiability results have been obtained using two different approaches: one approach uses the so-called extended Yule–Walker equations proposed by Chen and Zadrozny (1998). G-identifiability results using this approach have been derived in Anderson et al. (2012, 2016a). The second approach is based on blocked observations as described in Filler (2010), Anderson et al. (2016a), Ghysels (2016). An extension to the VARMA case has been given by Anderson et al. (2016b), Zadrozny (2016), where in Anderson et al. (2016b) g-identifiability has been shown for the case where the MA order is smaller than or equal to the AR order. We want to mention that condition 5 in Zadrozny (2016) excludes the case that the AR order is smaller than the MA order and thus the case considered in this paper. Here, we show that the situation is completely different in case of VMA systems and of VARMA systems with MA order exceeding the AR order. This is a consequence of the fact that non-observed autocovariances of the VMA output process can freely vary in a certain neighborhood without violating the assumptions. For the VARMA case considered here, this is true for special non-observed autocovariances. The central result of this paper is, in a certain sense, a negative result, as it shows that naive estimation approaches may fail in the case where the MA order exceeds the AR order.

In Section 2 we give a detailed analysis of non-identifiability of VMA systems from MF data for the case of stock as well as for the case of flow variables. In Section 3 we show non-identifiability for the VARMA case, where the MA order exceeds the AR order. Whereas Sections 2 and 3 deal with the case, where the innovation variance matrix is non-singular, Section 4 gives generic identifiability results for the case, where the innovation variance matrix is singular with rank being sufficiently small.

2. High frequency VMA systems and MF data: non-identifiability results

We consider (high frequency) multivariate MA(r) systems of the form

$$y_t = \nu_t + B_1\nu_{t-1} + \cdots + B_r\nu_{t-r}, \quad t \in \mathbb{Z}, \tag{2.1}$$

where r is a specified integer, $B_i \in \mathbb{R}^{n\times n}$ and $(\nu_t | t \in \mathbb{Z})$ is white noise with $\Sigma = \mathbb{E}\left(\nu_t\nu_t^T\right)$. In addition, we impose the strict miniphase condition

$$\det b(z) \neq 0, \quad |z| \leq 1, \tag{2.2}$$

where

$$b(z) = I_n + B_1 z + \cdots + B_r z^r, \tag{2.3}$$

is the corresponding VMA transfer function. We use z for a complex variable as well as for the backward shift on $\mathbb{Z}$. In this and in the next section we assume that

$$q = \text{rank}(\Sigma) = n, \tag{2.4}$$

holds. The VMA parameter space considered is of the form

$$\Theta_{\text{MA}} = \underbrace{\left\{\text{vec}(B_1,\ldots,B_r) \in \mathbb{R}^{n^2 r} \,|\, \det b(z) \neq 0,\ |z| \leq 1\right\}}_{S_{\text{MA}}} \times \underbrace{\left\{\text{vech}(\Sigma) \in \mathbb{R}^{\frac{n(n+1)}{2}} \,|\, \Sigma = \Sigma^T,\ \Sigma > 0\right\}}_{\underline{\Sigma}}.$$

Before going into the details let us outline the basic ideas in the flow of arguments behind Lemma 1, Theorem 1, Theorem 2 and Theorem 3. Consider Fig. 2.1 below. In this figure Γ_{MA} is the set of autocovariances corresponding to Θ_{MA}, ι_{F} is the mapping describing this correspondence, ι_{M} relates autocovariances to spectra and ι_{P} MA parameters to transfer functions. The basic idea is as follows: first we show for the MA(r) case that for every corresponding spectrum there is a neighborhood, which is contained in the set of MA(r) spectra, not violating the strict miniphase assumption. In the next step we demonstrate that by the bijective nature of the mapping between parameters and spectral densities in the high frequency case, we obtain that varying a particular autocovariance, that cannot be directly observed from mixed frequency data, leads to a corresponding variation of parameters; thus giving a (non-singleton) subset of a class of observationally equivalent parameters under mixed frequency observations.

Let Θ_{MA} be endowed with the usual Euclidean norm. As can be seen easily (as “det” is a continuous mapping) the set S_{MA} defined above is open in $\mathbb{R}^{n^2 r}$ and $\underline{\Sigma}$ is open in $\mathbb{R}^{\frac{n(n+1)}{2}}$. For every $\tau = \text{vec}(B_1,\ldots,B_r) \in S_{\text{MA}}$, of course, $\iota_P(\tau) = b(\mathrm{e}^{-i\lambda})$ is uniquely given by (2.3). Conversely, for given $b(\mathrm{e}^{-i\lambda})$, τ is uniquely given by

$$B_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} b\left(\mathrm{e}^{-i\lambda}\right)\mathrm{e}^{i\lambda j}d\lambda.$$

Thus, ι_P is a bijection between S_{MA} and $\iota_P(S_{\text{MA}})$ and in this sense we identify S_{MA} and $\iota_P(S_{\text{MA}})$. The following lemma is straightforward to show:

Lemma 1. *S_{MA} (endowed with the Euclidean metric) and $\iota_P(S_{\text{MA}})$ endowed with the supremum norm*

$$\left\|b\left(\mathrm{e}^{-i\lambda}\right)\right\|_{\sup} = \sup_{\lambda\in[-\pi,\pi]}\left\|b\left(\mathrm{e}^{-i\lambda}\right)\right\|_{\max}, \tag{2.5}$$

are homeomorphic. Here, $\|.\|_{\max}$ denotes the max norm for matrices with complex entries.

In other words, the above lemma says that S_{MA}, when endowed with the Euclidean metric and then endowed with the sup norm (2.5) has the same open sets, i.e., the two metrics are topologically equivalent.

Now consider the autocovariance sequence $\gamma : \mathbb{Z} \to \mathbb{R}^{n\times n}$ of the VMA(r) process (2.1), where

$$\gamma(j) = \mathbb{E}\left(y_j y_0^T\right). \tag{2.6}$$

The corresponding spectral density is given by

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi}\sum_{j=-r}^{r}\gamma(j)\mathrm{e}^{-i\lambda j} \\ &= b\left(\mathrm{e}^{-i\lambda}\right)\Sigma b^*\left(\mathrm{e}^{-i\lambda}\right), \end{aligned} \tag{2.7}$$

where * denotes the conjugate transpose. As is well known, there is a bijection between the covariances and the corresponding spectral density; let us write $\iota_{\mathrm{M}}(\gamma) = f$. The set of all autocovariance sequences has to be non-negative definite; in addition, for the MA case considered here $\gamma(j) = 0,\ \forall |j| > r$. Moreover, $\gamma(0)$ is symmetric and positive definite and due to (2.2) we have $f(\lambda) > 0,\ \forall\lambda \in [-\pi, \pi]$. Thus, for the VMA($r$) case, we can parameterize the set of all autocovariances by

$$\Gamma_{\mathrm{MA}} = \left\{\gamma = \left(\mathrm{vech}(\gamma(0))^T, \mathrm{vec}(\gamma(1), \ldots, \gamma(r))^T\right)^T \in \mathbb{R}^{\frac{n(n+1)}{2}+n^2 r} \middle| f(\lambda) = \iota_{\mathrm{M}}(\gamma)(\lambda) > 0,\ \forall\lambda \in [-\pi, \pi]\right\}.$$

Here, we have identified $\left(\mathrm{vech}(\gamma(0))^T, \mathrm{vec}(\gamma(1), \ldots, \gamma(r))^T\right)^T$ with γ. As above, we identify Γ_{MA} and $\iota_{\mathrm{M}}(\Gamma_{\mathrm{MA}})$ and we can endow Γ_{MA} with the Euclidean norm as well as with the sup norm for spectral densities and these norms are topologically equivalent. We have:

Theorem 1. *Under the assumptions (2.2) and (2.4) Γ_{MA} is open in $\mathbb{R}^{\frac{n(n+1)}{2}+n^2 r}$.*

Proof. Let $\gamma \in \Gamma_{\mathrm{MA}}$. Then, $f(\lambda) > 0\ \forall\lambda \in [-\pi, \pi]$ is equivalent to all leading principal minors of f being positive on $[-\pi, \pi]$. Since these minors are continuous functions in γ, they are bounded away from zero. Thus, there is a neighborhood of f (in the sup and therefore also in the Euclidean norm), where $f > 0$ and thus a neighborhood of γ which is contained in Γ_{MA}. $\square$

In the next step, let us consider the relation between $\theta \in \Theta_{\mathrm{MA}}$ and $\gamma \in \Gamma_{\mathrm{MA}}$, denoted by $\iota_{\mathrm{F}}(\theta) = \gamma$, defined by (2.7). Clearly ι_{F} is continuous. By the spectral factorization theorem (see Rozanov, 1967; Hannan, 1970; Hannan and Deistler, 2012), ι_{F} is bijective. In Anderson (1985), the continuity of the spectral factorization with the sup norm for spectra and the L_2 norm for transfer functions has been stated. As can be seen easily, on Θ_{MA} the L_2 and the Euclidean norm are the same up to normalization. Thus, we have:

Theorem 2. *The mapping ι_{F}: $\Theta_{\mathrm{MA}} \to \Gamma_{\mathrm{MA}}$ is a homeomorphism.*

Summarizing, this can be depicted in the following diagram: note that, strictly speaking, ι_{P} maps τ onto $b\left(\mathrm{e}^{-i\lambda}\right)$, but the extension of ι_{P} to Θ_{MA} is evident.

In this diagram the index M stands for moments, F stands for the factorization and P for parameters.

Now consider the MF case: let

$$y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix}, \tag{2.8}$$

where the n_f-dimensional fast component y_t^f is observed for every $t \in \mathbb{Z}$ and the n_s-dimensional slow component y_t^s is observed only for $t \in N\mathbb{Z}$ (N being an integer larger than one). Thus, we consider the MF stock case here. Throughout we assume that $n_f \geq 1$ and $n_s \geq 1$. Without loss of generality we restrict ourselves to the case $N = 2$ here. Let

$$\gamma(j) = \begin{pmatrix} \gamma^{ff}(j) & \gamma^{fs}(j) \\ \gamma^{sf}(j) & \gamma^{ss}(j) \end{pmatrix},$$

$$\begin{tikzcd} \theta \in \Theta_{\mathrm{MA}} \arrow[rr, leftrightarrow, bend left, "\iota_{\mathrm{F}}"] \arrow[d, leftrightarrow, "\iota_{\mathrm{P}}"'] & & \gamma \in \Gamma_{\mathrm{MA}} \arrow[d, leftrightarrow, "\iota_{\mathrm{M}}"] \\ (b(\mathrm{e}^{-i\lambda}), \Sigma) & & f \end{tikzcd}$$

Fig. 2.1. Diagram of the homeomorphisms.

denote the corresponding partitioned autocovariance matrices. Then, $\gamma^{ff}(j)$, $\gamma^{fs}(j)$, $\gamma^{sf}(j)$ and $\gamma^{ss}(2j)$, $j \in \mathbb{Z}$ can be directly observed, i.e., directly estimated from data, but this is not the case for $\gamma^{ss}(j)$, j odd. Consider, for instance, the case $r = 1$; then, under the mixed frequency scheme described above, *the autocovariances $\gamma^{ss}(1) \in \mathbb{R}^{n_s \times n_s}$ can be freely varied in a certain neighborhood corresponding to the "true"* $(\mathrm{vech}(\gamma(0))^T, \mathrm{vec}(\gamma(1))^T)^T$ *without leaving the set Γ_{MA} due to its openness.* This variation corresponds to a set $O \subseteq \Gamma_{\mathrm{MA}}$ of dimension $n_s \times n_s$ say, and, by what was stated before $\iota_{\mathrm{F}}^{-1}(O)$ is a subset of the class of *observationally equivalent* VMA systems of topological dimension $n_s \times n_s$. Thus, we have shown:

Theorem 3. *Under the assumptions (2.2) and (2.4) in the MF stock case, the VMA(r) parameters are not identifiable.*

As immediately can be seen we do not even have local identifiability in this case. A heuristic argument concerning the non-identifiability can be given by the counting condition, e.g., for the case $n = 2,\ r = 1,\ N = 2$ we can directly observe 6 autocovariances whereas the corresponding VMA(1) system is described by 7 free parameters. As shown in the following example, such an argument may be misleading:

Example. Consider the VMA(1) system, where $n = 2,\ b^{fs} = 0$ is fixed

$$B_1 = \begin{pmatrix} b^{ff} & 0 \\ b^{sf} & b^{ss} \end{pmatrix},\ \Sigma = \begin{pmatrix} \sigma^{ff} & \sigma^{fs} \\ \sigma^{sf} & \sigma^{ss} \end{pmatrix} > 0, \tag{2.9}$$

and $|b^{ff}|$, $|b^{ss}| < 1$. Now let us define an equivalent (in terms of directly observable autocovariances) VMA(1) system with

$$\bar{B}_1 = B_1 + \delta \begin{pmatrix} 0 & 0 \\ -\frac{\sigma^{fs}}{\sigma^{ff}} & 1 \end{pmatrix},\ \bar{\Sigma} = \begin{pmatrix} \sigma^{ff} & \sigma^{fs} \\ \sigma^{sf} & \bar{\sigma}^{ss} \end{pmatrix},$$

$$\bar{\sigma}^{ss} = \frac{\left(\left(1 + (b^{ss})^2\right)\sigma^{ss} + \left(2\delta b^{ss} + \delta^2\right)\frac{\left(\sigma^{sf}\right)^2}{\sigma^{ff}} \right)}{1 + (b^{ss} + \delta)^2} > 0,$$

where $\delta \in (-1 - b^{ss}, 1 - b^{ss})$ is the free variation. It follows that $\left(\mathrm{vec}(\bar{B}_1)^T, \mathrm{vech}(\bar{\Sigma})^T\right)^T \in \Theta_{\mathrm{MA}},\ \bar{\gamma}(0) = \gamma(0)$ but

$$\bar{\gamma}(1) = \bar{B}_1 \bar{\Sigma} = \gamma(1) + \delta \begin{pmatrix} 0 & 0 \\ 0 & \left(\bar{\sigma}^{ss} - \frac{(\sigma^{fs})^2}{\sigma^{ff}}\right) \end{pmatrix},$$

and therefore $\bar{\gamma}^{ss}(1) \neq \gamma^{ss}(1)$ for $\delta \neq 0$. If we assume that the class of equivalent VMA(1) systems satisfies $\bar{b}^{fs} = 0$, it can be shown that the formulas above, by varying δ, define a class of observationally equivalent systems, which is not a singleton. Note that in this case we have 6 unknown parameters (3 for B_1 and 3 for Σ) and 6 directly observable autocovariances and therefore the "counting condition" fails.

Until now, only the case where stock variables are observed, has been considered. Now we assume that the observed slow component w_t is obtained by the linear observation scheme

$$w_t = k_0 y_t^s + k_1 y_{t-1}^s + \cdots + k_{N-1} y_{t-N+1}^s, \quad t \in N\mathbb{Z}, \tag{2.10}$$

where $k_i \in \mathbb{R}^{n_s \times n_s}$ are known and k_0 is non-singular. Then, the corresponding autocovariances which can be directly observed, are $\gamma^{ff}(j) = \mathbb{E}(y_j^f (y_0^f)^T)$,

$$\gamma^{fw}(j) = \mathbb{E}\left(y_j^f w_0^T\right),$$

$$\gamma^{wf}(j) = \mathbb{E}\left(w_j \left(y_0^f\right)^T\right)$$

and

$$\gamma^{ww}(Nj) = \mathbb{E}\left(w_{Nj} w_0^T\right), \quad j \in \mathbb{Z}.$$

As shown in Anderson et al. (2016a), Section 5, $\gamma^{fs}(j)$ and $\gamma^{sf}(j)$ can be reconstructed from $\gamma^{ff}(j)$ and $\gamma^{fw}(j)$. Therefore, in a last step, we have to reconstruct $\gamma^{ss}(j)$ from

$$\gamma^{ww}(Nj) = \sum_{i,h=0}^{N-1} k_i \gamma^{ss}(Nj + i - h) k_h^T, \quad j \in \mathbb{Z}, \tag{2.11}$$

since $\gamma^{ww}(Nj)$ are the only directly observed autocovariances depending on $\gamma^{ss}(j)$. Note that $\gamma^{ww}(Nj) = 0$ for $|j| > \left\lfloor \frac{r-1}{N} + 1 \right\rfloor$, where $\lfloor \alpha \rfloor$ denotes the largest integer smaller than or equal to α, and therefore we only have $\left\lfloor \frac{r-1}{N} + 1 \right\rfloor + 1$ directly observed autocovariances from γ^{ww} and $r + 1$ unknown autocovariances from γ^{ss} (with non-negative lags) which are unequal to zero. For simplicity, we only consider the case $n_s = 1$, generalizations being straightforward. In this case

(2.11) changes to $\gamma^{ww}(Nj) = \sum_{i,h=0}^{N-1} k_i k_h \gamma^{ss}(Nj + i - h)$ and thus we obtain a system of equations with $\lfloor \frac{r-1}{N} + 1 \rfloor + 1$ equations and $r + 1$ unknown autocovariances γ^{ss}. Under the additional assumption that k_{N-1} is unequal to zero, we can uniquely reconstruct the γ^{ss} for the case $r = 1$ and therefore we obtain identifiability. Nevertheless, in the case $r > 1$ we have more unknown autocovariances γ^{ss} than equations and thus we do not obtain identifiability. For the case $r = 1$, $n_s \geq 1$ and $N \in \mathbb{N}$ we have to assume that k_{N-1} and $\sum_{i=0}^{N-1} k_i \otimes k_i$, where $\otimes$ denotes the Kronecker product, are non-singular to obtain identifiability. Note that this excludes the case of stock variables.

If we relax the strict miniphase assumption (2.2), in the sense that we allow roots on and outside the unit circle, we may have g(eneric)-identifiability in certain "extreme" cases, e.g., for $r = 1$, $n = 2$ and $b(z)$ has zeros at 1 and -1. Here, we say that a subset of a given set is generic, if it contains an open and dense subset of this set.

3. High frequency VARMA systems and MF data: non-identifiability results

In this section we consider VARMA(p, r) systems

$$y_t + A_1 y_{t-1} + \cdots + A_p y_{t-p} = \nu_t + B_1 \nu_{t-1} + \cdots + B_r \nu_{t-r}, \quad t \in \mathbb{Z}, \tag{3.1}$$

where $A_i, B_i \in \mathbb{R}^{n \times n}$ are parameter matrices, $p, r \in \mathbb{N}$ are specified integers and $(\nu_t | t \in \mathbb{Z})$ is white noise with $\Sigma = \mathbb{E}(\nu_t \nu_t^T)$. Let $a(z) = I_n + A_1 z + \cdots + A_p z^p$ and $b(z) = I_n + B_1 z + \cdots + B_r z^r$. We impose the usual stability condition

$$\det a(z) \neq 0, \quad |z| \leq 1, \tag{3.2}$$

as well as the strict miniphase condition (2.2). Throughout this section we assume that

$$r > p. \tag{3.3}$$

The case $r \leq p$ has been treated in Anderson et al. (2016b). The spectral density of (y_t) is given by

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} a^{-1}(e^{-i\lambda}) b(e^{-i\lambda}) \Sigma b^*(e^{-i\lambda}) a^{-*}(e^{-i\lambda}) \\ &= \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-i\lambda j}. \end{aligned} \tag{3.4}$$

The parameter space considered here is of the form

$$\begin{aligned} \Theta_{\text{ARMA}} = \{ &(A_1, \ldots, A_p, B_1, \ldots, B_r) | \det a(z) \neq 0,\ |z| \leq 1, \det b(z) \neq 0,\ |z| \leq 1 \\ &\text{rank}(A_p) = \text{rank}(B_r) = n, (a(z), b(z)) \text{ is left co-prime} \} \times \\ &\{ \Sigma | \Sigma = \Sigma^T, \Sigma > 0 \}. \end{aligned} \tag{3.5}$$

Note that in the high frequency case Θ_{ARMA} is identifiable (see Hannan, 1971; Hannan and Deistler, 2012) and that the conditions ($a(z)$, $b(z)$) is left co-prime and $\text{rank}(A_p) = \text{rank}(B_r) = n$ are generically fulfilled. In the next step we show that the AR parameters $(A_1, \ldots, A_p)$ are generically identifiable from MF stock data. If we postmultiply (3.1) by $\left(y_{t-j}^f\right)^T$, $j = r+1, \ldots, r+np$, and take the expectation, we obtain analogously to Chen and Zadrozny (1998)

$$\begin{aligned} &\mathbb{E}\left(y_t \left(\left(y_{t-r-1}^f\right)^T, \left(y_{t-r-2}^f\right)^T, \ldots, \left(y_{t-r-np}^f\right)^T\right)\right) \\ &= (-A_1, \ldots, -A_p) \underbrace{\mathbb{E}\left(\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} \left(\left(y_{t-r-1}^f\right)^T, \left(y_{t-r-2}^f\right)^T, \ldots, \left(y_{t-r-np}^f\right)^T\right)\right)}_{Z_{\text{ARMA}}}, \end{aligned} \tag{3.6}$$

where $Z_{\text{ARMA}} \in \mathbb{R}^{np \times npn_f}$. With the same techniques as in Anderson et al. (2016a, 2016b), it can be shown that Z_{ARMA} has generically full row rank. Thus, we get:

Theorem 4. *Under the assumptions (3.2), (2.2) and (2.4) the AR parameters of (3.1) are g-identifiable.*

Let $\mathcal{G} = (I_n, 0, \ldots, 0)$ and

$$\mathcal{A} = \begin{pmatrix} -A_1 & \cdots & -A_{p-1} & -A_p \\ I_n & & & \\ & \ddots & & \\ & & I_n & 0 \end{pmatrix},$$

be the companion form of the system parameters $\left(A_1,\dots,A_p\right)$ corresponding to $a(z)$. For $k=\left\lceil\frac{r-p+1}{N}\right\rceil$, where $\lceil\alpha\rceil$ denotes the smallest integer larger than or equal to α, we obtain the following system of equations

$$\underbrace{\begin{pmatrix}\gamma(kN)\\ \gamma((k+1)N)\\ \gamma((k+2)N)\\ \vdots\\ \gamma((k+np-1)N)\end{pmatrix}}_{\mathcal{O}_1}=\underbrace{\begin{pmatrix}\mathcal{G}\mathcal{A}^{kN-r}\\ \mathcal{G}\mathcal{A}^{kN-r}\mathcal{A}^{N}\\ \mathcal{G}\mathcal{A}^{kN-r}\mathcal{A}^{2N}\\ \vdots\\ \mathcal{G}\mathcal{A}^{kN-r}\mathcal{A}^{(np-1)N}\end{pmatrix}}_{\mathcal{O}}\begin{pmatrix}\gamma(r)\\ \vdots\\ \gamma(r-p+1)\end{pmatrix}.$$

Analogously to Koelbl (2015), Anderson et al. (2016b) it can be shown that $\mathcal{O}$ has generically full column rank. Thus, we can uniquely obtain $\gamma^{ss}(j)$, $r\geq|j|\geq r-p+1\geq 2$ from $\left(A_1,\dots,A_p\right)$ and the second moments contained in $\mathcal{O}_1$, which can be directly observed in the case of stock data, by using the formula

$$\begin{pmatrix}\gamma(r)\\ \vdots\\ \gamma(r-p+1)\end{pmatrix}=\left(\mathcal{O}^T\mathcal{O}\right)^{-1}\mathcal{O}^T\mathcal{O}_1. \tag{3.7}$$

We obtain $\gamma^{ss}(j)$, $|j|>r$ from $\gamma(j)=-\sum_{i=1}^{p}A_i\gamma(j-i)$. Obviously, we cannot reconstruct $\gamma^{ss}(1)$ from (3.7). To be more specific, there are still $n_s^2\left(r-p-\left\lfloor\frac{r-p}{N}\right\rfloor\right)$ autocovariances missing.

Theorem 5. *Under the assumptions (3.2), (2.2) and (2.4) in the MF stock case, for $r>p$, Θ_{ARMA} is not identifiable.*

Proof. Let the VARMA spectral density $f(\lambda)$ be given by (3.4). For simplicity of notation, we restrict ourselves to the case $n_s=1$. We define

$$\bar{f}(\lambda,\delta)=f(\lambda)+\begin{pmatrix}0&0\\0&\delta\left(e^{-i\lambda}+e^{i\lambda}\right)\end{pmatrix}, \tag{3.8}$$

and

$$a\left(e^{-i\lambda}\right)\bar{f}(\lambda,\delta)a^*\left(e^{-i\lambda}\right)=a\left(e^{-i\lambda}\right)f(\lambda)a^*\left(e^{-i\lambda}\right)+\ a\left(e^{-i\lambda}\right)\begin{pmatrix}0&0\\0&\delta\left(e^{-i\lambda}+e^{i\lambda}\right)\end{pmatrix}a^*\left(e^{-i\lambda}\right). \tag{3.9}$$

Note that the left hand side of (3.9) is a trigonometric polynomial matrix of degree r. As can be seen easily, Theorem 1 implies that for δ in a suitable neighborhood of zero this left hand side is a spectral density of an MA(r) system. Thus, (3.8) corresponds to a VARMA(p, r) spectral density where all second moments of (y_t) stay the same with the exception of $\gamma^{ss}(1)$ (and of course $\gamma^{ss}(-1)$), which has been disturbed by δ, where δ varies in a certain neighborhood of zero. Because left co-primeness of $(a(z), b(z))$ and the rank conditions $\operatorname{rank}\left(A_p\right)=\operatorname{rank}(B_r)=n$ are open properties, the system corresponding to $\bar{f}(\lambda,\delta)$ can be described in the parameter space Θ_{ARMA}. Since the mapping between the second moments and the parameters is homeomorphic, the variation of δ gives the corresponding variation in the parameter space. □

Note that the basic idea of the proof is not applicable for the case $p\geq r$. Now we consider the more general observation scheme (2.10). First of all, observe that again $\gamma^{fs}(j)$ and $\gamma^{sf}(j)$ can be reconstructed from $\gamma^{ff}(j)$ and $\gamma^{fw}(j)$. Therefore, we can g-identify the AR parameters, see (3.6). In a second step, we want to reconstruct the missing autocovariances $\gamma^{ss}(j)$, $j\in\mathbb{N}$ from (2.11). For simplicity, we just investigate the case $n_s=1$, generalizations are straightforward. For $k=\left\lceil\frac{r-p}{N}+1\right\rceil$ we can arrange the following system of equations

$$\underbrace{\begin{pmatrix}\gamma^{ww}(\alpha N)\\ \gamma^{ww}((\alpha+1)N)\\ \gamma^{ww}((\alpha+2)N)\\ \vdots\\ \gamma^{ww}((\alpha+np-1)N)\end{pmatrix}}_{\mathcal{O}_1^F}=\underbrace{\begin{pmatrix}(0,1)\mathcal{G}\left(\sum_{i,h=0}^{N-1}k_ik_h\mathcal{A}^{i-h+kN-r}\right)\\ (0,1)\mathcal{G}\left(\sum_{i,h=0}^{N-1}k_ik_h\mathcal{A}^{i-h+kN-r}\right)\mathcal{A}^{N}\\ (0,1)\mathcal{G}\left(\sum_{i,h=0}^{N-1}k_ik_h\mathcal{A}^{i-h+kN-r}\right)\mathcal{A}^{2N}\\ \vdots\\ (0,1)\mathcal{G}\left(\sum_{i,h=0}^{N-1}k_ik_h\mathcal{A}^{i-h+kN-r}\right)\mathcal{A}^{(np-1)N}\end{pmatrix}}_{\mathcal{O}^F}\begin{pmatrix}\begin{pmatrix}\gamma^{fs}(r)\\ \gamma^{ss}(r)\end{pmatrix}\\ \vdots\\ \begin{pmatrix}\gamma^{fs}(r-p+1)\\ \gamma^{ss}(r-p+1)\end{pmatrix}\end{pmatrix},$$

where $\mathcal{O}_1^F$ and $\mathcal{O}^F$ can be constructed from the observed autocovariances and the AR parameters. Again it can be shown that $\mathcal{O}^F$ has generically full column rank and therefore we obtain $\gamma^{ss}(j)$, $j\geq r-p+1\geq 2$. In a last step, we have to reconstruct the remaining unknown autocovariances $\gamma^{ss}(j)$, $j=0,\dots,r-p$ with the help of $\gamma^{ww}(Nj)$, $j=0,\dots,\left\lceil\frac{r-p}{N}\right\rceil$. Thus, we get a system of equations with $\left\lceil\frac{r-p}{N}\right\rceil+1$ equations and $r-p+1$ unknown autocovariances and therefore we cannot expect to uniquely reconstruct the remaining autocovariances except for the case $r=p+1$ (under the additional assumption that $k_{N-1}\neq 0$). Thus, VARMA $(p, p+1)$ systems are g-identifiable in the flow case, but not for $r>p+1$.

4. High frequency VARMA systems and MF data: the singular case

Here, we consider the singular case, i.e., the case $\mathrm{rank}(\Sigma) = q < n$. In this case, as is shown below, as opposed to the regular case $\mathrm{rank}(\Sigma) = n$, we may have identifiability for VMA and VARMA systems with $r > p$, if q is smaller than or equal to n_f.

In an evident notation we write the spectral density $f(\lambda)$ of the VARMA system as

$$f(\lambda) = \begin{pmatrix} f^{ff}(\lambda) & f^{fs}(\lambda) \\ f^{sf}(\lambda) & f^{ss}(\lambda) \end{pmatrix}. \tag{4.1}$$

Clearly, f has normal rank equal to q. Let

$$\Sigma = \begin{pmatrix} \Sigma^{ff} & \Sigma^{fs} \\ \Sigma^{sf} & \Sigma^{ss} \end{pmatrix},$$

be partitioned accordingly.

Lemma 2. *For $\Sigma^{ff} > 0$ it follows that $f^{ff}(\lambda)$ is non-singular almost everywhere.*

Proof. Let $\mathcal{H}(t) = \overline{\mathrm{span}}\{y_j : j \le t\}$ and $\mathcal{H}^f(t) = \overline{\mathrm{span}}\{y_j^f : j \le t\}$ be the Hilbert spaces spanned by all univariate components of y_j, $j \le t$ and y_j^f, $j \le t$, respectively. Furthermore, let $y_{t|t-1}^f$ and z_t be the projection of y_t^f onto $\mathcal{H}(t-1)$ and $\mathcal{H}^f(t-1)$, respectively. Then, it follows that $y_t^f = y_{t|t-1}^f + \nu_t^f = z_t + \eta_t$. Now $\mathcal{H}^f(t) \subseteq \mathcal{H}(t)$ implies $\mathbb{E}(\eta_t \eta_t^T) \ge \mathbb{E}(\nu_t^f (\nu_t^f)^T)$ and thus $\Sigma^{ff} > 0$ implies $f^{ff}(\lambda) > 0$ almost everywhere. □

If $q = n_f$ holds, then generically (in the parameter space of singular VARMA systems, where $q = \mathrm{rank}(\Sigma)$ has been prescribed) already $f^{ff}(\lambda)$ has normal rank equal to q. In the mixed frequency case $f^{ff}(\lambda)$, $f^{fs}(\lambda)$ and $f^{sf}(\lambda)$ are obtained from mixed frequency data. This is true for the stock, but also for the more general linear observation scheme (2.10). Then, we obtain $f^{ss}(\lambda)$ (according to the Wiener filtering formula) as

$$f^{ss}(\lambda) = f^{sf}(\lambda) f^{ff}(\lambda)^{-1} f^{fs}(\lambda), \tag{4.2}$$

and thus, in this case, i.e., if $f^{ff}(\lambda)$ has already normal rank equal to q, $f(\lambda)$ is generically unique from the second moments which can be directly observed or reconstructed in the case of more general linear observation schemes. As is easily seen, this argument can also be applied to the case $q < n_f$, if $f^{ff}(\lambda)$ has rank q. In this way we have directly reconstructed, generically, the high frequency spectral density and therefore the well known high frequency identifiability result holds. Note that this result hold for all p, r, not depending on the relation between p and r.

Theorem 6. *For $n_f \ge q$, the VARMA(p, r) classes are g-identifiable.*

5. Conclusions

This paper deals with the non-identifiability of VMA and VARMA(p, r) $r > p$ systems in the case of mixed frequency observations. The main result is that opposed to the classes of VAR and VARMA systems, where the MA order is smaller than or equal to the AR order, generic identifiability cannot be obtained in case of regular systems and stock variables. With some modifications, these results hold through for more general observation schemes. For the singular case $q = \mathrm{rank}(\Sigma) \le n_f$ generic identifiability can be obtained.

In a certain sense, the main contribution of the paper is a negative result, saying that the case considered here is substantially different from the VAR and the VARMA(p, r) $r \le p$ case. We think that the result is of some interest, in particular, since negative results in our opinion are also important and since they are not known up to now. The practical consequence of our results is, that naive approaches, e.g., using the EM algorithm in the case of mixed frequency VMA systems, will fail. The consequences of our results for model selection need further investigation.

Acknowledgments

Support by the Oesterreichische Nationalbank (Oesterreichische Nationalbank, Anniversary Fund, project number: 16546), the FWF (Austrian Science Fund under contract P24198/N18) and NICTA is gratefully acknowledged. NICTA is funded by the Australian Government.

References

Anderson, B.D.O., 1985. Continuity of the spectral factorization operation. J. Comput. Appl. Math. 4 (2), 139–156.

Anderson, B.D.O., Deistler, M., Felsenstein, E., Funovits, B., Koelbl, L., Zamani, M., 2016. Multivariate AR systems and mixed frequency data: g-identifiability and estimation. Econ. Theory 32 (4), 793–826. doi:10.1017/S0266466615000043.

Anderson, B. D. O., Deistler, M., Felsenstein, E., Funovits, B., Zadrozny, P. A., Eichler, M., Chen, W., Zamani, M., 2012. Identifiability of regular and singular multivariate autoregressive models. Proceedings of the Fifty-One IEEE Conference on Decision and Control (CDC) pp. 184–189. doi: 10.1109/CDC.2012.6426713.

Anderson, B.D.O., 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. J. Econ. 192 (2), 366–373. doi:10.1016/j.jeconom.2016.02.004.
Chambers, M.J., 2016. The estimation of continuous time models with mixed frequency data. J. Econ. 193 (2), 390–404. doi:10.1016/j.jeconom.2016.04.013.
Chen, B., Zadrozny, P.A., 1998. An extended Yule–Walker method for estimating a vector autoregressive model with mixed-frequency data. Adv. Econ. 13, 47–73.
Filler, A., 2010. Generalized Dynamic Factor Models – Structure Theory and Estimation for Single Frequency and Mixed Frequency Data. Vienna University of Technology Ph.D. Thesis.
Foroni, C., Marcellino, M. G., 2013. A Survey of Econometric Methods for Mixed-Frequency Data. Norges Bank Working Paper 2013-6.
Ghysels, E., 2016. Macroeconomics and the reality of mixed frequency data. J. Econ. 193 (2), 294–314.
Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: further results and new directions. Econ. Rev. 26 (1), 53–90.
Hannan, E.J., 1970. Multiple Time Series. John Wiley & Sons, New York.
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. SIAM Classics in Applied Mathematics, Philadelphia.
Koelbl, L., 2015. VAR Systems: g-Identifiability and Asymptotic Properties of Parameter Estimates for the Mixed-Frequency Case. Vienna University of Technology Ph.D. Thesis.
Koelbl, L., Braumann, A., Felsenstein, E., Deistler, M., 2016. Estimation of VAR systems from mixed-frequency data: the stock and the flow case. Adv. Econ. 35, 43–73. doi:10.1108/S0731-905320150000035002.
Marcellino, M., Schumacher, C., 2010. Factor MIDAS for nowcasting and forecasting with ragged-edge data: a model comparison for German GDP. Oxf. Bull. Econ. Stat. 72 (4), 518–550.
Nijman, T.E., 1985. Missing Observations in Dynamic Macroeconomic Modeling. VU University Amsterdam Ph.D. Thesis.
Rozanov, Y.A., 1967. Stationary Random Processes. Holden-Day, San Francisco.
Wohlrabe, K., 2008. Forecasting with mixed-frequency time series models Ludwig-Maximilians-Universität München, Ph.D. Thesis.
Zadrozny, P.A., 2016. Extended Yule–Walker identification of VARMA models with single- or mixed-frequency data. J. Econ. 193 (2), 438–446.