Central European Journal of Economic Modelling and Econometrics

Identification and Estimation of Initial Conditions in Non-Minimal State-Space Models

Victor Bystrov*

Submitted: 25.03.2020, Accepted: 20.07.2020

Abstract

In this paper the identification problem is considered for initial conditions in a non-minimal state-space model that includes interpretable state variables generated by non-stationary stochastic processes. In order to solve the identification problem, structural restrictions are imposed on initial conditions in a state-space model with redundant state variables. The corresponding restricted maximum likelihood estimator of initial conditions is derived. The restricted estimator of initial conditions can be used in order to compute uniquely identified realizations of interpretable latent variables. The identification problem is illustrated analytically using a simple structural economic model.

Keywords: identification, latent variables, state-space model, redundancy

JEL Classification: C32, C51, C52

*University of Lodz; e-mail: victor.bystrov@uni.lodz.pl; ORCID: 0000-0003-0980-2790

1 Introduction

Various structural economic and financial models include latent indicators that are identified and estimated as state variables in corresponding state-space representations. Such unobservable variables include, among others, expected inflation (Burmeister et al., 1986), potential output and output gap (Watson, 1986; Kuttner, 1994), natural rate of interest (Laubach and Williams, 2003 and 2016; Holston et al. 2017), natural rate of unemployment (Gordon, 1997 and 2013; Staiger et al., 1997; Laubach, 2001), and stock market fundamentals (Balke and Wohar, 2002). These latent indicators have specific economic or financial interpretations and are often used as reference variables in policy analysis. In order to enable the use of these variables for policy advice, their values should be uniquely identified.

Latent economic and financial variables are often modelled as non-stationary integrated processes. A realization of an integrated process retains full memory of initial conditions. In the absence of prior information about initial conditions, they are either modelled as diffuse or identified and estimated using in-sample information. As diffuse priors imply high estimation uncertainty, the second approach is more often used for estimation of unobservable economic and financial indicators.

In many applications structural models of non-stationary processes are specified in such a way that the corresponding state-space representations are non-minimal (reducible): they include redundant state variables and can be reduced to models with fewer states. Redundancy implies that state variables cannot be uniquely identified using observable variables and system equations: there are various realizations of latent variables in identical models that generate the same moments for observable variables and the same value of the likelihood function. Various observationally equivalent realizations of the same latent variables are caused by unidentifiable initial conditions in the presence of redundant state variables.

If the primary objective of an application is the estimation of a state variable that has a specific economic or financial interpretation, then either an irreducible state-space model should be specified or a sufficient number of restrictions should be imposed onto initial conditions in order to identify the state variable. For example, if the objective is the estimation of a natural interest rate, then the corresponding structural model and initial conditions should be specified in such a way that there are no various observationally equivalent realizations of the natural rate in the same model. The structural model that admits alternative observationally equivalent realizations of the natural rate can have various policy implications, which depend on the choice of a specific realization of the natural rate.

The objective of this paper is to demonstrate the consequences of redundancy in state-space models including non-stationary processes and propose a solution to the identification problem in the presence of redundant state variables. Using the integrated likelihood function, it is shown that in the presence of redundancy various observationally-equivalent realizations of state variables can be obtained by changing initial conditions for these variables. The identification problem can be solved by imposing theoretically-motivated restrictions on the initial conditions in a model with redundant states. The methodological considerations of the paper are illustrated with a simple example of a structural economic model containing interpretable latent variables.

The paper aims to draw attention to an important problem that is often ignored in applied research. The policy analysis that uses latent indicators is well-grounded only if the latent indicators are interpretable and uniquely identified. The lack of identification, caused by redundancy, makes latent variables unsuitable for a policy advice. This issue should be solved at the stage of model specification.

The paper is organized as follows. Next section provides a brief literature review. A state-space model is specified in Section 3 and the identification of initial conditions is discussed in Section 4. The maximum likelihood estimation of initial conditions is described in Section 5. Latent variables are derived as functions of initial conditions in Section 6. An economic example illustrating the issue of identifiability is described in Section 7. Conclusions are given in Section 8.

2 Literature review

Hendry (1995, p. 36) considered three aspects of identification: correspondence to underlying economic behaviour, satisfaction of assumed interpretation, and uniqueness. These aspects of identification are usually applied to parameters of a structural econometric model. But in a state-space model implemented for identification and estimation of latent economic or financial variables, which can be used in policy analysis, these conditions for identification should also be applied to state variables. They should have desired interpretation and be uniquely identified.

The information-theoretical definition of identifiability for a general stochastic model is given in Rothenberg (1971). A method of identification in state-space, developed by Wall (1987), employed a blend of control theory and econometrics. Using the concept of minimal representation (a representation that includes no redundant state variables), Wall (1987) defined a class of observationally equivalent state-space model and gave an operational criterion of observational equivalence.

The identification method developed by Wall (1987) is implemented in few economic applications (Burmeister et al., 1986; Wall and Stoffer, 2002; McGrattan, 2010). Nevertheless, in many applications the minimality conditions, specified in Wall (1987), are not tested explicitly (Kuttner, 1994; Laubach, 2001; Laubach and Williams, 2003 and 2016; Holston et al., 2017) and in some of these applications state-space models include redundant states. For example, it can be demonstrated that the state-space model, implemented in Laubach and Williams (2003), includes redundant states (see Bystrov, 2019).

In practice, the identification of structural parameters is often analysed under assumed identification of latent variables. If, however, a state-space model includes redundant state variables, then state variables as well as structural parameters are generally not

identified.
The identification and estimation of state variables that are generated by non-stationary processes requires a specification of initial states. The identification theory, developed in Wall (1987), ignores initial conditions for state variables. For a state-space representation of a stationary process initial conditions can be specified as functions of parameters describing the state-space representation. Nevertheless, for a non-stationary process its state-space representation should be augmented by a model for initial states (see De Jong, 1988; Hamilton, 1994; Durbin and Koopman, 2012). In this paper it is demonstrated that the identification of latent variables generated by non-stationary processes in a non-minimal state-space model can be obtained by imposing restrictions onto initial conditions.

3 State-space model

In order to determine identification conditions for latent variables and to demonstrate the consequences of redundancy, the following state-space model is considered:

$$\xi_{t+1} = F\xi_t + Gv_{t+1}, \tag{1}$$

$$y_t = H\xi_t \quad + Jw_t, \tag{2}$$

where ξ_t is a $p \times 1$ vector of state variables, y_t is an $n \times 1$ vector of observed explained variables, v_{t+1} is a $q \times 1$ vector of structural shocks, $v_{t+1} \sim i.i.d.N(0, Q)$, and w_t is an $r \times 1$ vector of measurement errors, $w_t \sim i.i.d.N(0, R)$; F, G, H, J, Q, and R are system matrices of dimensions $p \times p$, $p \times q$, $n \times p$, $n \times r$, $q \times q$, and $r \times r$ correspondingly. Structural shocks v_t and measurement errors w_t are assumed to be mutually independent and independent of initial states ξ_0. Matrices G, J, Q, and R are assumed to satisfy the following rank conditions: $rank(G) = rank(Q) = q$, $rank(J) = rank(R) = r$ and $q + r \geq n$. All system matrices are functions of a parameter vector θ: $F = F(\theta)$, $G = G(\theta)$, $H = H(\theta)$, $J = J(\theta)$, $R = R(\theta)$, and $Q = Q(\theta)$. In what follows, it is assumed that at least one eigenvalue of F lies on the unit circle, but not outside of it: at least one state variable is integrated.
The specification of the model (1)–(2) is completed with initial conditions which are defined analogously to De Jong (1988): $\xi_0 \sim N(\mu, \Sigma)$, where Σ is assumed to be a full-rank matrix (unless $\Sigma \equiv 0$), and $\xi_1 = F\xi_0 + Gv_1$. The vector μ and the matrix Σ are not assumed to be functions of the parameter vector θ: they have to be estimated or chosen on the basis of prior information. Such specification of initial conditions is used when state variables are generated by non-stationary processes. For example, it might be used for modelling natural rates of interest and unemployment that are often assumed to be generated by non-stationary processes.
Exogenous variables are not included in the model (1)–(2) for ease of exposition. Nevertheless, the results will hold if exogenous regression effects are added either to transition equation (1) or measurement equation (2). Analogously to Durbin and

Koopman (2012), the model (1)–(2) can be represented in a general matrix form:

$$\xi = \mathbf{F}\xi_0 + \mathbf{\Gamma v}, \tag{3}$$
$$\mathbf{y} = \mathbf{H}\xi + \mathbf{Jw}, \tag{4}$$

where

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_T \end{bmatrix}, \ \mathbf{F} = \begin{bmatrix} F \\ F^2 \\ \vdots \\ F^T \end{bmatrix}, \ \mathbf{\Gamma} = \begin{bmatrix} G & 0 & \dots & 0 \\ FG & G & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ F^{T-1}G & F^{T-2}G & \dots & G \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_T \end{bmatrix},$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}, \ \mathbf{H} = \begin{bmatrix} H & 0 & \dots & 0 \\ 0 & H & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & H \end{bmatrix}, \ \mathbf{J} = \begin{bmatrix} J & 0 & \dots & 0 \\ 0 & J & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J \end{bmatrix}, \ \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_T \end{bmatrix}.$$

Substitution of (3) into (4) produces

$$\mathbf{y} = \mathbf{X}\xi_0 + \mathbf{u}, \ \mathbf{u} \sim N(0, \mathbf{\Omega}), \tag{5}$$

where $\mathbf{u} = \mathbf{H\Gamma v} + \mathbf{Jw}$ and $\mathbf{\Omega} = \mathbf{H\Gamma Q\Gamma' H'} + \mathbf{JRJ'}$. The system matrices in (5) are defined as follows:

$$\mathbf{X} = \begin{bmatrix} HF \\ HF^2 \\ \vdots \\ HF^T \end{bmatrix}, \ \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{bmatrix}, \ \mathbf{Q} = \begin{bmatrix} Q & 0 & \dots & 0 \\ 0 & Q & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q \end{bmatrix},$$

$$\text{and } \mathbf{R} = \begin{bmatrix} R & 0 & \dots & 0 \\ 0 & R & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R \end{bmatrix}.$$

By construction, the $Tr \times Tq$ matrix $\mathbf{\Gamma}$ has a full column rank: $rank(\mathbf{\Gamma}) = Tq$. The $Tn \times Tn$ covariance matrix $\mathbf{\Omega}$ is assumed to be invertible. (The rank conditions $rank(\mathbf{\Gamma}) = rank(\mathbf{Q}) = Tq$ and $rank(\mathbf{J}) = rank(\mathbf{R}) = Tr$ and $q + r \geq n$ are necessary but not sufficient for invertibility of $\mathbf{\Omega}$ as $rank(\mathbf{\Omega}) \leq rank(\mathbf{H\Gamma Q\Gamma' H'}) + rank(\mathbf{JRJ'})$). The first two moments of the observable variables $\mathbf{y}$ are

$$E\left[\mathbf{y}|\mu, \Sigma, \theta\right] = \mathbf{X}\mu, \tag{6}$$
$$Var\left[\mathbf{y}|\mu, \Sigma, \theta\right] = \mathbf{\Omega} + \mathbf{X}\Sigma\mathbf{X}', \tag{7}$$

where matrices $\mathbf{X}$ and $\mathbf{\Omega}$ depend on the parameter vector θ: $\mathbf{X} = \mathbf{X}(\theta)$ and $\mathbf{\Omega} = \mathbf{\Omega}(\theta)$.

4 Identification of initial conditions

Given the representation (5) and moment conditions (6)–(7), the minimal state-space model should satisfy the rank condition

$$rank(\mathbf{X}) = p \tag{8}$$

for any parameter vector θ from a set Θ of all possible vectors θ. The rank condition (8) implies that the system of expectations equations (6) has a unique solution with respect to μ.

For minimal state-space representations there is a well-defined class of observationally equivalent models, determined by invertible coordinate transformations of state vectors. The identification of parameters in a minimal model is achieved by imposing restrictions guaranteeing that the only admissible coordinate transformation is the identity transformation.

The rank condition (8) is satisfied if and only if columns of $\mathbf{X}$ are linearly independent for any admissible parameter vector $\theta \in \Theta$, which (by the definition of $\mathbf{X}$) is true if and only if the system of equations

$$HF^t c = 0 \;\text{ for }\; t = 1, \ldots, T \tag{9}$$

has the unique solution $c = 0$. From Cayley-Hamilton theorem it follows that the system (9) is equivalent to

$$HF^t c = 0 \;\text{ for }\; t = 1, \ldots, p,$$

that has the unique solution $c = 0$ if and only if

$$rank \begin{bmatrix} HF \\ HF^2 \\ \vdots \\ HF^p \end{bmatrix} = p. \tag{10}$$

The condition (10) above is an identifiability condition for μ and it is analogous to the observability condition in the theory of optimal control (for more details, see Youla, 1966).

If $rank(\mathbf{X}) = k < p$, which means that there are redundant state variables, the identification can be obtained by imposing restrictions on the initial conditions:

$$A\mu = b, \tag{11}$$

where b is a known $(p-k) \times 1$ vector and A is a known $(p-k) \times p$ matrix of the full row rank such that the linear sub-space spanned by the rows of A is complementary to the linear subspace spanned by the rows of $\mathbf{X}$:

$$span(A') \cap span(\mathbf{X}') = \{\mathbf{0}\}. \tag{12}$$

The condition (12), which should be satisfied for any admissible parameter vector $\theta \in \Theta$, guarantees that the matrix $[A' : \mathbf{X}']'$ formed by adjoining rows of A to $\mathbf{X}$ is of the full column rank p. It means that the restrictions (11) cannot be imposed independently from the specification of matrix $\mathbf{X}$. In order to maintain assumed interpretation of state variables, the choice of matrix A should be based on assumptions of the corresponding structural model.
Using (6) and (11), it is possible to specify the augmented system

$$\begin{bmatrix} b \\ E(\mathbf{y}|\mu, \Sigma, \theta) \end{bmatrix} = \begin{bmatrix} A \\ \mathbf{X} \end{bmatrix} \mu, \tag{13}$$

where the matrix $[A' : \mathbf{X}']'$ has the full column rank p. The system (13) has a unique solution with respect to μ for a given $\mathbf{X}$.
The restrictions (11) and and (12) allow specifying a class of observationally-equivalent non-minimal models such that the matrix $\mathbf{X}$ has a reduced column rank $k < p$ for any admissible parameter vector θ. Analougsly to minimal models, an equivalence class of non-minimal models subject to restrictions (11) and (12) is determined by invertible coordinate transformations of state vectors.

5 Estimation of initial conditions

In order to derive a resticted maximum likelihood estimator of initial conditions, it is necessary to specify the integrated likelihood function and describe a method of its evaluation. Using the moments (6) and (7) the integrated log-likelihood function can be written as

$$\log L(\mathbf{y}|\mu, \Sigma, \theta) = -\tfrac{Tn}{2}\log 2\pi - \tfrac{1}{2}\log|\mathbf{\Omega}| - \tfrac{1}{2}\log|\Sigma| - \tfrac{1}{2}\log|\Sigma^{-1} + \mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}| + \\ -\tfrac{1}{2}(\mathbf{y} - \mathbf{X}\mu)'(\mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}\mathbf{X}(\Sigma^{-1} + \mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1})(\mathbf{y} - \mathbf{X}\mu). \tag{14}$$

The first-order optimality conditions imply that the maximum likelihood estimator of $p \times 1$ vector of initial states, μ, should satisfy the normal equation:

$$\mathbf{X}'\Psi\mathbf{X}\mu = \mathbf{X}'\Psi\mathbf{y}, \tag{15}$$

where matrix $\Psi = \mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}\mathbf{X}(\Sigma^{-1} + \mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}$ is invertible by assumptions. If $rank(\mathbf{X}) = p$, then the vector μ is identified and there is a unique solution of the normal equation (15):

$$\hat{\mu} = [\mathbf{X}'\Psi\mathbf{X}]^{-1}\mathbf{X}'\Psi\mathbf{y}. \tag{16}$$

Nevertheless, if $rank(\mathbf{X}) < p$, then the matrix $[\mathbf{X}'\Psi\mathbf{X}]$ is not invertible, the parameter vector μ is not identified and the unique estimator (16) does not exist. In this case, a solution of the normal equation (15) exists if

$$[\mathbf{X}'\Psi\mathbf{X}]\,[\mathbf{X}'\Psi\mathbf{X}]^{\dagger}\,\mathbf{X}'\Psi\mathbf{y} = \mathbf{X}'\Psi\mathbf{y},$$

where $[\mathbf{X}'\Psi\mathbf{X}]^{\dagger}$ denotes the generalized (Moore-Penrose) inverse of matrix $[\mathbf{X}'\Psi\mathbf{X}]$. Any

$$\hat{\mu} = [\mathbf{X}'\Psi\mathbf{X}]^{\dagger}\,\mathbf{X}'\Psi\mathbf{y} + \left[I - [\mathbf{X}'\Psi\mathbf{X}]^{\dagger}[\mathbf{X}'\Psi\mathbf{X}]\right]\boldsymbol{\Gamma},\ \boldsymbol{\Gamma} \in \mathbb{R}^{p},$$

will satisfy the normal equation (15).
If the restrictions (11) and (12) are imposed, then the the first-order optimality conditions for constrained optimization imply that the restricted maximum likelihood estimator of p-dimensional vector of initial states should satisfy equations

$$\begin{bmatrix} \mathbf{X}'\Psi\mathbf{X} & A' \\ A & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\Psi\mathbf{y} \\ b \end{bmatrix}, \tag{17}$$

where λ is a $(p-k)\times 1$ vector of Lagrange multipliers. For an invertible symmetric Ψ the matrix $\mathbf{X}'\Psi\mathbf{X}$ spans the same row space as the matrix $\mathbf{X}$ and this space is complementary to the row space of A: $span(\mathbf{X}'\Psi\mathbf{X}) \cap span(A') = \{\mathbf{0}\}$. It implies that the matrix $[\mathbf{X}'\Psi\mathbf{X} : A']'$ has a full column rank equal to p. The matrix $[A : \mathbf{0}]'$ has a full column rank equal to $p-k$.
It is easy to demonstrate that matrices $[\mathbf{X}'\Psi\mathbf{X} : A']'$ and $[A : \mathbf{0}]'$ span complementary sub-spaces of $(2p-k)$-dimensional vector space:

$$span\left(\begin{bmatrix} \mathbf{X}'\Psi\mathbf{X} \\ A \end{bmatrix}\right) \cap span\left(\begin{bmatrix} A' \\ \mathbf{0} \end{bmatrix}\right) = \{\mathbf{0}\}.$$

This implies that the $(2p-k)\times(2p-k)$ matrix

$$\begin{bmatrix} \mathbf{X}'\Psi\mathbf{X} & A' \\ A & \mathbf{0} \end{bmatrix}$$

is invertible and there is a unique solution for the system (17):

$$\begin{bmatrix} \hat{\mu}_R \\ \hat{\lambda}_R \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\Psi\mathbf{X} & A' \\ A & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\Psi\mathbf{y} \\ b \end{bmatrix}, \tag{18}$$

where $\hat{\mu}_R$ is a restricted estimator of initial states.
The integrated likelihood function (14) can be evaluated using the Cholesky decomposition $\boldsymbol{\Omega} = \mathbf{L}^{-1}\mathbf{V}(\mathbf{L}')^{-1}$, where where $\mathbf{V}$ is a block-diagonal matrix and the

matrix $\mathbf{L}$ is a lower block triangular matrix,

$$
\begin{aligned}
\mathbf{V} &= \begin{bmatrix} V_1 & 0 & 0 & \dots & 0 \\ 0 & V_2 & 0 & \dots & 0 \\ 0 & 0 & V_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & V_T \end{bmatrix}, \\
\mathbf{L} &= \begin{bmatrix} I & 0 & 0 & \dots & 0 & 0 \\ L_1 & I & 0 & \dots & 0 & 0 \\ L_2 L_1 & L_2 & I & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ L_{T-1}\cdots L_1 & L_{T-1}\cdots L_2 & L_{T-1}\cdots L_3 & \dots & L_{T-1} & I \end{bmatrix},
\end{aligned}
$$

computed recursively by the Kalman filter. The recursions of the Kalman filter are

$$e_t = y_t - H\hat{\xi}_t,\ V_t = HP_tH' + JRJ',\ K_t = FP_tH'V_t^{-1},\ L_t = F - K_tH, \tag{19}$$

$$\hat{\xi}_{t+1} = F\hat{\xi}_t + K_te_t,\ P_{t+1} = FP_tF' - K_tV_tK_t' + GQG' \text{ for } t = 1, 2, \dots, T \tag{20}$$

with $\hat{\xi}_1 = F\hat{\xi}_0 = 0$ and $P_1 = GQG'$.

Using the Cholesky decomposition $\Omega = \mathbf{L}^{-1}\mathbf{V}(\mathbf{L}')^{-1}$, obtained by the Kalman filter, the integrated likelihood function (14) can be rewritten as

$$
\begin{aligned}
&\log L(\mathbf{y}|\mu, \Sigma, \theta) = -\tfrac{Tn}{2}\log 2\pi - \tfrac{1}{2}\log|\Sigma| - \tfrac{1}{2}\log|\mathbf{V}| - \tfrac{1}{2}\log\left|\Sigma^{-1} + \mathbf{S}\right| + \\
&-\tfrac{1}{2}\mu'\Sigma^{-1}\mu - \tfrac{1}{2}\mathbf{e}'\mathbf{V}^{-1}\mathbf{e} + \tfrac{1}{2}\left(\Sigma^{-1}\mu + \mathbf{s}\right)'\left(\Sigma^{-1} + \mathbf{S}\right)^{-1}\left(\Sigma^{-1}\mu + \mathbf{s}\right).
\end{aligned} \tag{21}
$$

(For detailed derivations, see de Jong, 1988; or Durbin and Koopman, 2012). The $Tn \times 1$ vector $\mathbf{e}$ is a vector of prediction errors, $\mathbf{e} = \mathbf{Ly}$. The $p \times p$ matrix $\mathbf{S}$ and $p \times 1$ vector $\mathbf{s}$ are defined by equations

$$\mathbf{S} = \mathbf{X}'\mathbf{L}'\mathbf{V}^{-1}\mathbf{LX} \text{ and } \mathbf{s} = \mathbf{X}'\mathbf{L}'\mathbf{V}^{-1}\mathbf{e} \tag{22}$$

and can be computed recursively:

$$\mathbf{s} = \mathbf{s} + X_t'Z_t'V_t^{-1}e_t,\quad \mathbf{S} = \mathbf{S} + X_t'Z_t'V_t^{-1}Z_tX_t,\quad Z_t = L_tZ_{t-1},\quad X_t = HF^t$$

with $\mathbf{s}$ and $\mathbf{S}$ initialized at 0, and $Z_0 = I$. The resulting matrix $\mathbf{S}$ and vector $\mathbf{s}$ can be used in order to compute the maximum likelihood estimator of expected initial states μ.

The first-order optimality conditions imply that the maximum likelihood estimator of $p \times 1$ vector of initial states, μ, should satisfy the equation

$$\mathbf{S}\mu = \mathbf{s} \tag{23}$$

that has a unique solution if and only if the matrix $\mathbf{S}$ is invertible. It is easily shown that $\mathbf{S}$ is invertible ($rank(\mathbf{S}) = p$) if and only if the rank condition (10) is satisfied. Then the solution is is

$$\hat{\mu} = \mathbf{S}^{-1}\mathbf{s}.$$

If $rank(\mathbf{S}) = k < p$ (there is redundancy), then for given restrictions (11) and (12) the first-order optimality conditions for constrained optimization imply that the restricted maximum likelihood estimator of p-dimensional vector of initial states should satisfy equations

$$\begin{bmatrix} \mathbf{S} & A' \\ A & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{s} \\ b \end{bmatrix}, \tag{24}$$

where λ is a $(p-k) \times 1$ vector of Lagrange multipliers. Analogously to the previous section, it can be shown that for given assumtions the matrix

$$\begin{bmatrix} \mathbf{S} & A' \\ A & \mathbf{0} \end{bmatrix}$$

is invertible and there is a unique solution for the system (24):

$$\begin{bmatrix} \hat{\mu}_R \\ \hat{\lambda}_R \end{bmatrix} = \begin{bmatrix} \mathbf{S} & A' \\ A & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{s} \\ b \end{bmatrix}. \tag{25}$$

6 Estimation of state variables

For given $\hat{\xi}_0$ and P_0 the Kalman filter produces a vector of filtered states

$$\hat{\xi} = \mathbf{LKy} + \mathbf{L}(\iota \otimes F\hat{\xi}_0),$$

where the matrix $\mathbf{L}$ is described above,

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_1 & 0 & \dots & 0 \\ 0 & K_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K_{T-1} \end{bmatrix} \text{ and } \iota = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where matrices K_t $(t = 1, 2 \dots, T-1)$ are defined in (19). If $\hat{\xi}_0 = \hat{\mu}$ and $P_0 = \hat{\Sigma}$, where $\hat{\mu}$ and $\hat{\Sigma}$ are the maximum likelihood estimators, then

$$\hat{\xi} = \mathbf{LKy} + \mathbf{L}(\iota \otimes F\hat{\mu}).$$

Conditionally on the data, the time series of filtered states $\hat{\xi}$ is an affine function of $\hat{\mu}$.
If the vector μ is unidentifiable (the rank condition (10) is not satisfied and matrix $\mathbf{S}$ is not invertible), then the state vector ξ is also unidentifiable: for a given parameter vector θ multiple realizations of process $\{\xi\}$, depending on the choice of initial conditions, will generate the same expected values of observable variables and the same value of the integrated likelihood (21).
In applications, the rank condition (10) is very rarely explicitly tested at the stage of model specification. If matrix $\mathbf{S}$ is found to be singular, it is treated as a computational issue and some modification in computations is introduced in order to guarantee matrix invertibility, calculate $\hat{\mu} = \mathbf{S}^{-1}\mathbf{s}$ and the corresponding realization of state variables $\hat{\xi}$. Nevertheless, if the objective of an application is to identify an interpretable state variable (e.g., expected inflation or natural rate of interest), then the failure of the rank condition (10) implies that this state variable is not uniquely identified and its multiple observationally equivalent realizations can be computed using the same dynamic model but different initial conditions. This is not a computational issue, which can be addressed at the estimation stage, but a modelling issue, which should be adressed at the stage of model specification.
If the restrictions (11) and (12) are imposed on the vector μ, then the state vector ξ is identified and the filtered states can be computed:

$$\hat{\xi} = \mathbf{LKy} + \mathbf{L}(\iota \otimes F\hat{\mu}_R),$$

where the estimator $\hat{\mu}_R$ is given by (25). By imposing structural restrictions on the vector of initial states, it is possible to proceed with a state-space representation that includes redundant state variables. Nevertheless, if the assumed interpretation of state variables is to be retained, the structural restrictions on initial states should be motivated by the corresponding theoretical model and, preferrably, exclude alternative theories from the class of observational equivalence.
The policy analysis that is based on the state variables, identified with the help of *ad hoc* restrictions is unreliable if a different observationally equivalent realization of the state variables can be obtained by a modification of *ad hoc* restrictions. Any *ad hoc* restrictions on initial states can be substituted with another *ad hoc* restrictions generating an observationally equivalent model with a different realization of state variables.

7 Illustrative example

In this section a state-space model of unobservable expected inflation is considered. The model is designed to demonstrate the identification problem in the presence of redundant states, using simple settings. There is no claim that the model can be directly applied in empirical research. Nevertheless, similar types of identification problems can be found in more complex economic and financial models (see, e.g., Fiorentini et al. 2018, Bystrov 2019). The relations between measured and unobservable variables are given by equations

$$\begin{aligned}\pi_t &= \pi_t^e + \varepsilon_{\pi t},\\ i_t &= r_t + \pi_t^e,\end{aligned}$$

where π_t is a measured inflation rate, π_t^e is an unobservable expected inflation rate, i_t is a measured nominal rate of interest, r_t is an unobservable real rate of interest, $\varepsilon_{\pi t}$ is an expectation error which is assumed to be independent and identically distributed over time: $\varepsilon_{\pi t} \sim i.i.d.N(0, \sigma_\pi^2)$. The dynamics of unobservable variables is given by

$$\begin{aligned}\pi_t^e &= \phi_1 \pi_{t-1}^e + \phi_2 \pi_{t-2}^e + \theta_1 r_{t-1} + \theta_2 r_{t-2} + \varepsilon_{\pi^e t},\\ r_t &= r_{t-1} + \varepsilon_{rt},\end{aligned}$$

where $\varepsilon_{\pi^e t}$ and ε_{rt} are independently and identically distributed: $\varepsilon_{\pi^e t} \sim i.i.d.N(0, \sigma_{\pi^e}^2)$ and $\varepsilon_{rt} \sim i.i.d.N(0, \sigma_r^2)$. Both observable and latent variables are assumed to be generated by non-stationary processes. The initial conditions for latent variables cannot be specified using parameters of dynamic equations: they have to be specified independently.

The model can be represented in the state-space form (1)–(2) with four state variables:

$$\begin{bmatrix} \pi_t^e \\ \pi_{t-1}^e \\ r_t \\ r_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \theta_1 & \theta_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \pi_{t-1}^e \\ \pi_{t-2}^e \\ r_{t-1} \\ r_{t-2} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{rt} \\ \varepsilon_{\pi^e t} \end{bmatrix}, \qquad (26)$$

$$\begin{bmatrix} \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \pi_t^e \\ \pi_{t-1}^e \\ r_t \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \varepsilon_{\pi t}, \qquad (27)$$

where

$$F = \begin{bmatrix} \phi_1 & \phi_2 & \theta_1 & \theta_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

As follows from the rank condition (10), the state-space representation (26)–(27) includes redundant states if the matrix

$$\begin{bmatrix} HF \\ HF^2 \\ HF^3 \\ HF^4 \end{bmatrix} = \begin{bmatrix} C^{(1)} & C^{(2)} & C^{(3)} & C^{(4)} \end{bmatrix}, \qquad (28)$$

where

$$C^{(1)} = \begin{bmatrix} \phi_1 \\ \phi_1 \\ \phi_1^2 + \phi_2 \\ \phi_1^2 + \phi_2 \\ \phi_1(\phi_1^2 + 2\phi_2) \\ \phi_1(\phi_1^2 + 2\phi_2) \\ \phi_1^2\phi_2 + (\phi_1^2 + \phi_2)^2 \\ \phi_1^2\phi_2 + (\phi_1^2 + \phi_2)^2 \end{bmatrix}, \ C^{(2)} = \begin{bmatrix} \phi_2 \\ \phi_2 \\ \phi_2\phi_1 \\ \phi_2\phi_1 \\ \phi_2(\phi_1^2 + \phi_2) \\ \phi_2(\phi_1^2 + \phi_2) \\ \phi_2\phi_1(\phi_1^2 + 2\phi_2) \\ \phi_2\phi_1(\phi_1^2 + 2\phi_2) \end{bmatrix},$$

$$C^{(3)} = \begin{bmatrix} \theta_1 \\ \theta_1 + 1 \\ \theta_1(1+\phi_1) + \theta_2 \\ \theta_1(1+\phi_1) + \theta_2 \\ \theta_1(1+\phi_1+\phi_1^2+\phi_2) + \theta_2(1+\phi_1) \\ \theta_1(1+\phi_1+\phi_1^2+\phi_2) + \theta_2(1+\phi_1) + 1 \\ \theta_1(1+\phi_1+\phi_1^2+\phi_1^3+2\phi_1\phi_2+\phi_2) + \theta_2(1+\phi_1+\phi_1\phi_2+\phi_1^2) \\ \theta_1(1+\phi_1+\phi_1^2+\phi_1^3+2\phi_1\phi_2+\phi_2) + \theta_2(1+\phi_1+\phi_1\phi_2+\phi_1^2) + 1 \end{bmatrix},$$

$$C^{(4)} = \begin{bmatrix} \theta_2 \\ \theta_2 \\ \theta_2\phi_1 \\ \theta_2\phi_1 \\ \theta_2(\phi_1^2 + \phi_2) \\ \theta_2(\phi_1^2 + \phi_2) \\ \theta_2\phi_1(\phi_1^2 + 2\phi_2) \\ \theta_2\phi 1(\phi_1^2 + 2\phi_2) \end{bmatrix},$$

is of reduced column rank, which is true for any $\phi_2 \neq 0$ and $\theta_2 \neq 0$ because the fourth column of this matrix, $C^{(4)}$, is proportional to the second column, $C^{(2)}$: $C^{(4)} = \delta C^{(2)}$, where $\delta = \theta_2/\phi_2$. (If either $\phi_2 = 0$ or $\theta_2 = 0$, then the matrix (28) is also of reduced rank, because either $C^{(2)}$ or $C^{(4)}$ is a zero column.)

The failure of the rank condition (10) implies that the vector of expected initial states $E\left[\pi_0^e,\, \pi_{-1}^e,\, r_0,\, r_{-1}\right]'$ is not uniquely identified in the model (26)–(27). For given non-zero parameters ϕ_2 and θ_2 and initial values $\bar{\pi}_0^e = E\left[\pi_0^e\right]$ and $\bar{r}_0 = E\left[r_0\right]$ any choice of $\bar{\pi}_{-1}^e = E\left[\pi_{-1}^e\right]$ and $\bar{r}_{-1} = E\left[r_{-1}\right]$, such that $(\bar{\pi}_{-1}^e + \frac{\theta_2}{\phi_2}\bar{r}_{-1})$ is fixed, will generate the same expected values for observable variables and the same value of the integrated likelihood function (21).

If the non-identifiability of initial conditions were ignored, then for observable time series $\{\pi_t\}_{t=1}^T$ and $\{i_t\}_{t=1}^T$ the maximum likelihood estimation of the structural parameters $[\phi_1, \phi_2, \theta_1, \theta_2, \sigma_\pi, \sigma_{\pi^e}, \sigma_r]'$ and initial conditions $E\left[\pi_0^e, \pi_{-1}^e, r_0, r_{-1}\right]'$,

based on the integrated likelihood function, would encounter computational problems as matrix **S** defined in (22) would not be invertible.
The issue of redundancy can be solved by changing the model specification. If the dynamics of the real interest were be described by an autoregressive process with two lags, $r_t = \varphi_1 r_{t-1} + \varphi_2 r_{t-2} + \varepsilon_{rt}$, then the model with four states would be minimal. Alternatively, restrictions can be imposed onto initial values in the non-minimal model.
Any non-zero linear combination of the rows of matrix (28) has the fourth component equal to the second component multiplied by $\delta = \theta_2/\phi_2$ (assuming that $\phi_2 \neq 0$ and $\theta_2 \neq 0$). If a restriction

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix} \begin{bmatrix} E\pi_0^e \\ E\pi_{-1}^e \\ Er_0 \\ Er_{-1} \end{bmatrix} = b, \tag{29}$$

such that $a_4 \neq \delta a_2$ for non-zero ϕ_2 and θ_2, is imposed, then the vector $[a_1, a_2, a_3, a_4]$ cannot be written as a linear combination of the rows of matrix (28). The augmented matrix obtained by adjoining the row $[a_1, a_2, a_3, a_4]$ to the matrix (28) has a full column rank and the initial conditions $E\left[\pi_0^e,\, \pi_{-1}^e,\, r_0,\, r_{-1}\right]'$ are uniquely identified as a solution of the augmented system defined in (13). In particular, the restriction

$$\begin{bmatrix} 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} E\pi_0^e \\ E\pi_{-1}^e \\ Er_0 \\ Er_{-1} \end{bmatrix} = 0. \tag{30}$$

implies unique identification of the initial conditions for non-zero ϕ_2 and θ_2.
In a more complex model the choice of restrictions would involve more structural assumptions. For example, Holston et al. (2017) use a non-minimal state-space model in order to estimate natural interest rates and output gaps in four advanced economies. The natural interest rate (r_t^*) is modelled as a sum of two latent variables generated by random walks: potential output growth rate (g_t) and a non-growth component (z_t). The model specification requires that initial conditions for these variables should include their lags: $E\left[g_0, g_{-1}, z_0, z_{-1}\right]'$. These initial conditions are not identified, because the model is non-minimal (for a proof, see Bystrov, 2019). Nevertheless, the identification can be obtained if restrictions are imposed on the initial conditions determining the natural interest rate.

8 Conclusions

Policy makers often use model-based latent variables, estimated by economists, as reference indicators in their decision making. The model-based reference indicators are usually compared with actual economic variables (e.g., the potential growth rate is compared with the measured growth rate or the natural interest rate is compared with the measured interest rate).

If a latent variable is not uniquely identified, then its use in the policy analysis is unjustified, as various observationally-equivalent realizations of this variable are possible. It often happens in empirical research that a specific realization of the unidentified latent variable is obtained by making auxiliary *ad hoc* assumptions at the estimation stage. But the identification problem has to be solved at the stage of the model specification.

This paper has focused on the identification of latent variables generated by non-stationary processes in a state-space model with redundant states. The redundancy in the specification of a state-space model causes unidentifiability of state variables. Nevertheless, state variables can be identified in the presence of redundancy if structural restrictions are imposed onto initial conditions for the state variables. In this paper, structural restrictions on initial conditions are specified in a state-space model with redundant states, the restricted maximum likelihood estimator of intitial conditions is derived and the estimated state variables are presented as functions of the system matrices and the restricted initial states.

The paper demonstrates that the identification of latent indicators that are generated by non-stationary processes can be obtained by a complete specification of the corresponding state-space model, and the problem of redundant state variables can be solved by imposing structural restrictions onto initial conditions. For the ease of exposition, the state-space model, considered in this paper, does not include exogenous variables. Nevertheless, the results can be extended to a more general case, although at the cost of more extensive derivations that are left for future research.

Acknowledgements

The author is grateful to two anonymous referees for their helpful comments and suggestions.

References

[1] Balke N. S., Wohar M. E., (2002), Low-Frequency Movements in Stock Prices: A State-Space Decomposition, *Review of Economics and Statistics* 84(4), 649–667.

[2] Burmeister E., Wall K. D., Hamilton J. D., (1986), Estimation of unobserved expected monthly inflation using Kalman filtering, *Journal of Business and Economic Statistics* 4(2), 147–160.

[3] Bystrov V., (2019), The observational equivalence of natural and unnatural rates of interest, *Przegląd Statystyczny* 66(3), 183–199.

[4] De Jong P., (1988), The likelihood for a state-space model, *Biometrika* 75, 165–169.

[5] Durbin J., Koopman S. J., (2012), *Time Series Analysis by State-Space Methods: Second Edition*, Oxford University Press.

[6] Fiorentini G., Galesi A., Pérez-Quirós G., Sentana E., (2018), The rise and fall of the natural interest rate, *Banco de Espana Working Papers*, 1822.

[7] Francke M. K., Koopman S. J., de Vos A. F., (2010), Likelihood functions for state-space models with diffuse initial conditions, *Journal of Time Series Analysis* 31, 407–414.

[8] Gordon R. J., (1997), The Time-Varying NAIRU and its Implications for Economic Policy, *Journal of Economic Perspectives* 11(1), 11–32.

[9] Gordon R. J., (2013), The Phillips Curve is Alive and Well: Inflation and the NAIRU during the Slow Recovery, *NBER Working Paper* 19390.

[10] Hamilton J. D., (1994), *Time Series Analysis*, Princeton University Press.

[11] Hendry D. F., (1995), *Dynamic Econometrics*, Oxford: Oxford University Press.

[12] Holston, K., Laubach, T., Williams, J. C., (2017), Measuring the natural rate of interest: international trends and determinants, *Journal of International Economics* 108(S1), 59–75.

[13] Kuttner K. N., (1994), Estimating Potential Output as a Latent Variable, *Journal of Business and Economic Statistics* 12(3), 361–368.

[14] Laubach T, (2001), Measuring the Nairu: Evidence from Seven Economies, *Review of Economics and Statistics* 83(2), 218–231.

[15] Laubach T., Williams J. C., (2003), Measuring the natural rate of interest, *Review of Economics and Statistics* 85(4), 1063–1070.

[16] Laubach T., Williams J. C., (2016), Measuring the natural rate of interest redux, *Business Economics* 51(2), 57–67.

[17] McGrattan E. R., (2010), Measurement with minimal theory, *Quarterly Review, Federal Reserve Bank of Minneapolis*, July, 2–13.

[18] Rothenberg T. J., (1971), Identification in Parametric Models, *Econometrica* 39(3), 577–591.

[19] Staiger D., Stock J. H., Watson, M. W., (1997), How Precise are Estimates of the Natural Rate of Unemployment?, NBER Working Paper No. 5477.

[20] Wall K. D., (1987), Identification theory for varying coefficient regression models, *Journal of Time Series Analysis* 8(3), 359–371.

[21] Wall K. D., Stoffer D., (2002), A state space approach to bootstrapping conditional forecasts in ARMA models, *Journal of Time Series Analysis* 23(6), 733–751.

[22] Watson M. W., (1986), Univariate Detrending Methods with Stochastic Trends, *Journal of Monetary Economics* 18, 49–75.

[23] Youla D. C., (1966), The synthesis of linear dynamical systems form prescribed weighting patterns, *SIAM Journal of Applied Mathematics* 14(3), 527–549.