On the solution of the linear rational expectations model with multiple lags

Hugo Kruiniger

University College London, Department of Economics, Gower Street, London WC1E 6BT, UK

Abstract

In this paper the symmetric linear rational expectations model from Kollintzas (1985) is generalized by allowing for multiple lags. By using a convenient decomposition of the matrix lag polynomial of the Euler–Lagrange equations that encompasses that for the Kollintzas model, it is shown that the model admits a unique stable solution. The stability condition derived by Kollintzas turns out to be valid in the more general setting. Moreover a procedure is given to obtain a particularly useful representation of the solution that is as close as possible to a closed form representation. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: C61; C62; D92

Keywords: Rational expectations; Closed form solution; Unique solution; Time-to-build; Stability condition

1. Introduction

The modern macroeconomic literature includes many models which involve the solution to a stochastic dynamic optimization problem. The objective functions in these problems are often approximated by linear–quadratic specifications (e.g. see Christiano, 1990). There are now several methods available for solving such models. Generally, they involve the factorization of the spectral density-like matrix lag polynomial which appears in the first order conditions,

E-mail address: kruiniger@nber.org (H. Kruiniger)
the so-called Euler–Lagrange conditions (ELC). Hansen and Sargent (1981) outlined a relatively fast and revealing solution procedure for multivariate linear rational expectations (LRE) models. Although their procedures are applicable to a wide class of linear–quadratic models, in particular models with dynamic interrelation and multiple lags, their approach has some disadvantages. The main drawback of their procedures is that the solutions are expressed in terms of parameters from which the structural parameters cannot be uniquely identified.

Several other methods have been suggested in the literature for the estimation and solution of LRE models, that do not suffer from this drawback. Kollintzas (1985) proposes a solution method for a class of multivariate LRE models that satisfy a symmetry condition. When a model has this property, it is possible to obtain an attractive representation of the forward solution of the model by diagonalizing the matrices of the lag polynomial of the ELC simultaneously. Epstein and Yatchew (1985) outlined a simplified solution and estimation procedure for a sub-class of symmetric LRE models that involves a reparametrization of the model and yields a closed form solution. This procedure was extended by Madan and Prucha (1989) in order to estimate more general, possibly nonsymmetric models. Although all these procedures allow for identification of the structural parameters, they have been developed for more restrictive LRE models than the Hansen and Sargent procedures, i.e. for models that consist of second order difference equations.

In this paper we derive a convenient representation of the solution to a generalized version of Kollintzas’ model which allows for multiple lags in the transition equations of the endogenous state variables. This wider framework encompasses a model with both adjustment costs and a general time-to-build structure as in Kydland and Prescott (1982) as well as other more complicated dynamic models. The nonsymmetric models studied by Madan and Prucha (1989) and Cassing and Kollintzas (1991) only allowed for time-to-build lags of one period. This is quite restrictive since it prevents one to study variables that have different lag structures at the same time or to model the value-put-in-place pattern of the investment process.

In order to obtain a solution of the symmetric LRE model with multiple lags, which is as close as possible to a closed form solution (CFS) and still allows for identification of the structural parameters, we will follow Broze et al. (1995), and restate the ELC in a reduced form in terms of realizations and revision processes. Next, we will present a convenient decomposition of the matrix lag polynomial. This decomposition considerably facilitates the derivation of restrictions on the revision processes which restore the equivalence of both representations of the first order conditions to the LRE model. Moreover, this decomposition includes the matrix lag polynomial of the LRE model given in Kollintzas (1985) and this allows us to show that his stability condition is also sufficient for the uniqueness and the stability of the solution of the more general model: after adding the restrictions on the revision processes that are implied by
the stability condition to the ‘equivalence’ restrictions mentioned above, all revision processes can be written as unique functions of the exogenous processes. The ‘stability’ restrictions on the revision processes are such that after substitution of these functions in the reduced form, the nonstable roots of the matrix lag polynomials cancel out. A lemma provides a procedure for imposing the stability condition on the solutions, that is, for finding the factorization of the matrix lag polynomial of the exogenous part of the reduced form that includes the factor with the nonstable roots that has to cancel out.

The paper is organized as follows. In the next section the model is stated. Section 2 also gives other examples of LRE models that feature multiple lags. The third section contains the main results of this paper. Section 4 concludes.

2. The model

In this section the symmetric linear rational expectations model with multiple lags is described using an example with time-to-build. The choice of this example is motivated by the fact that it is not restrictive and empirically relevant: Rossi (1988), Johnson (1994), Oliner et al. (1995), and Peeters (1995) all find evidence favoring factor demand equations with time-to-build.

Consider a (representative) agent who maximizes the expected discounted stream of cash flows given technological constraints and information on the economic environment. As far as the agent is concerned, this environment only consists of product and factor markets. He is a price taker in both markets. Following Kollintzas (1985), the gross cash flows are given by

\[ g(x_t, \Delta x_{t+1}, \mu_t) = x_t'x_t + \mu_t x_t - \frac{1}{2}x_t'E x_t - x_t'H \Delta x_{t+1} - \frac{1}{2}\Delta x_t'G \Delta x_{t+1}, \] (2.1)

where \(x_t\) is a \((F \times 1)\) vector of stocks of production factors, decided upon at equidistant points in time. The \((F \times F)\) constant matrices \(E, G\) and \(H\) are symmetric, \(E\) is positive (semi-)definite (symmetric) (PSDS) and \(G\) is PSD; in addition \([\begin{bmatrix} E & H \\ H' & G \end{bmatrix}]\) is PSDS and \(G-H\) is nonsingular; \(x\) is a \((F \times 1)\) vector of constants, while \(\mu_t\) is a \((F \times 1)\) vector of exogenous technology shocks. Furthermore, let \(I_{f,t}, f = 1, \ldots, F,\) be the gross changes of the stocks of the inputs (e.g. gross investments), let \(p_{f,t}\) be the corresponding real factor prices, which are assumed to be exogenous as well, and let \(\gamma\) denote the real discount factor. Then the agent chooses a contingency plan for \(\{x_t\}\) that maximizes

\[ \lim_{T \to \infty} E_0 \sum_{t=0}^{T} \gamma^t \left[ g(x_t, \Delta x_{t+1}, \mu_t) - \sum_{f=1}^{F} p_{f,t} I_{f,t} \right] \] (2.2)

subject to an initial condition, e.g. \(x_0 = \bar{x}\), a terminal condition and the transition equations for \(x\). The agent has rational expectations with regard to the exogenous processes. Furthermore these processes are assumed to be linear.
In the generalized LRE model, the transition equations allow for multiple lags. Kydland and Prescott (1982) stressed that it takes time to build new capital (see Montgomery (1995) for some descriptive statistics) and suggested a specification for investment that takes this feature into account. Let \( s_{f,j,t} \) be the total size of an investment project of type \( f,j \) periods from completion, \( j = 0, \ldots, \theta_f \), where \( \theta_f \) is the gestation lag. Then the building process for inputs of type \( f \) can be formalized as

\[
\begin{align*}
    x_{f,t+1} &= (1 - \delta_f)x_{f,t} + s_{f,0,t+1}, \\
    s_{f,j,t+1} &= s_{f,j+1,t}, \quad j = 0, \ldots, \theta_f - 1,
\end{align*}
\]

where \( \delta_f \) is the depreciation rate of inputs of type \( f \). Furthermore, let \( \psi_{f,j} \) be the share of the resources spent on investment projects that are completed in \( j \) periods. Total investment of type \( f \) in period \( t \) equals

\[
I_{f,t} = \sum_{j=0}^{\theta_f} \psi_{f,j}s_{f,j,t} \quad \text{with} \quad \sum_{j=0}^{\theta_f} \psi_{f,j} = 1.
\]

From combining Eqs. (2.3a)–(2.3c), it follows that

\[
I_{f,t} = \sum_{j=0}^{\theta_f+1} \phi_{f,j} x_{f,t+j-1}, \quad f = 1, \ldots, F,
\]

where the \( \phi_{f,j} \)'s depend on the \( \psi_{f,j} \)'s and \( \delta_f \).

There are other examples which fit the framework described above. One can think of a firm which builds tunnels and bridges. Here the time-to-build lags do not pertain to the inputs but to the outputs. Moreover, as these projects require specific capital and skilled labour, the firm tries to minimize the fluctuations in the total size of its projects as well as in the composition of the portfolio of projects. Another example is investment in a foreign country where profits can only be repatriated after some time.

### 3. The existence, uniqueness and representation of the solution

Maximization of Eq. (2.2) subject to Eqs. (2.1) and (2.4) implies the following first order conditions (ELC) for the stocks:

\[
\begin{align*}
    E_t \cdot \theta_f \left\{ [G_{f,t} - H_{f,t}]x_{t+1} - E_{f,t+1} \frac{(1 + \gamma)}{\gamma} G_{f,t} - 2H_{f,t} \right\} x_t \\
    + \frac{1}{\gamma} [G_{f,t} - H_{f,t}]x_{t-1} \right\} + \alpha_f + E_t \cdot \theta_f \mu_{f,t} \\
    - \sum_{j=0}^{\theta_f+1} \phi_{f,j} \gamma^{-j+1} E_t \cdot \theta_f \left[ p_{f,t-j+1} \right] = 0, \quad f = 1, \ldots, F,
\end{align*}
\]
where $M_f$ is generic notation for the $f$th row of matrix $M$. Because of the gestation lags, the decisions concerning $x_{1,t+\theta_1}, x_{2,t+\theta_2}, \ldots, x_{F,t+\theta_F}$ are made at time $t$.

Defining $L$ as the lag operator\(^1\) we can write system (3.1) succinctly as

$$E_tL^{-J}\Pi(L)y_t + u_t = 0 \quad (3.2)$$

with $y_t = [x_{1,t+\theta_1}, x_{2,t+\theta_2}, \ldots, x_{F,t+\theta_F}]'$

$$J = \theta_1 - \theta_F + 1 \text{ and assuming } \theta_1 \geq \theta_2 \geq \cdots \geq \theta_F \geq 0,$$

$$u_t = \begin{bmatrix}
  \alpha_1 + E_t\mu_{1,t+\theta_1} - \gamma^{-\theta_1} \sum_{j=0}^{\theta_1+1} \phi_{1,\theta_1-j+1}\gamma^jE_t[p_{1,t+j}]
  \\
  \alpha_2 + E_t\mu_{2,t+\theta_2} - \gamma^{-\theta_2} \sum_{j=0}^{\theta_2+1} \phi_{2,\theta_2-j+1}\gamma^jE_t[p_{2,t+j}]
  \\
  \vdots
  \\
  \alpha_F + E_t\mu_{F,t+\theta_F} - \gamma^{-\theta_F} \sum_{j=0}^{\theta_F+1} \phi_{F,\theta_F-j+1}\gamma^jE_t[p_{F,t+j}]
\end{bmatrix} = d_t + U(L)v_t,$$

where $U(L) = U_0 + U_1L + \cdots + U_JL^J, v_t$ is a white noise process and $d_t$ is the deterministic part of $u_t$\(^2\)

$$\Pi(\lambda) = \sum_{i=-J}^{J} A_i\lambda^{J-i}. \quad (3.3)$$

In the definition of the $A_i$'s we use the definitions

$$C = [c_{jk}], D = [d_{jk}] (F \times F)$$

with

$$c_{jk} = e_{jk} + \left(\frac{1+\gamma}{\gamma}\right)g_{jk} - 2h_{jk}, \quad d_{jk} = g_{jk} - h_{jk}, \quad j,k = 1,2,\ldots,F$$

$$A_i = [a_{i,jk}] (F \times F), \quad -J \leq i \leq J \text{ have nonzero elements}^{3}$$

$$a_{\theta_i-\theta_0+1,jk} = d_{jk}, \quad a_{\theta_i-\theta_0,jk} = -c_{jk}, \quad a_{\theta_i-\theta_0-1,jk} = \frac{1}{\gamma}d_{jk}.$$

---

\(^1\)The inverted lag operator $L^{-1}$ is the forward operator.

\(^2\)For the moment we will assume that $u_t$ has a finite VMA representation of order $\tau$. A more general VARIMA model for $u_t$ can be handled as well, as we will see at the end of this section.

\(^3\)Matrices with the same subscripts should be added up.
Introducing the vectors of revision processes \( \varepsilon_i^j = E(y_{t+j}) - E_{t-1}(y_{t+j}) \), \( j = 0, \ldots, \theta_1 \), which are martingale differences, we obtain by replacing the expectations in Eq. (3.2) with realizations and revisions and after shifting the system \( J \) periods back in time (see Appendix A)

\[
\Pi(L)y_t = \sum_{i=0}^{J-1} \sum_{j=i+1}^{J} A_{ij} L^{j-i} \varepsilon_i^j - u_{t-j}.
\]

(3.4)

The effective (not multiplied by zero) revision processes in Eq. (3.4) are \( \varepsilon_1^0, \varepsilon_2^0, \ldots, \varepsilon_{\theta_1}^0, \ldots, \varepsilon_F^0, \ldots, \varepsilon_{\theta_1}^{\theta_1} \) where the subscript now indicates a production factor. Thus we have \((\theta_1 + 1)F - \sum_{i=1}^{F} \theta_i\) revision processes.\(^4\) They are subject to constraints as we will show in the sequence. \( \Pi(\lambda) \) can also be represented as

\[
\Pi(\lambda) = [\lambda^M q_{ij}(\lambda)]
\]

with \( q_{ij}(\lambda) = d_{ij} - c_{ij}\lambda + (d_{ij}/\gamma)\lambda^2 \), \( i, j = 1, 2, \ldots, F \) and

\[
M = \begin{bmatrix}
J - 1 & J + \theta_2 - \theta_1 - 1 & \cdots & J + \theta_F - \theta_1 - 1 \\
J + \theta_1 - \theta_2 - 1 & J - 1 & \cdots & J + \theta_F - \theta_2 - 1 \\
\vdots & \vdots & \ddots & \vdots \\
J + \theta_1 - \theta_F - 1 & J + \theta_2 - \theta_F - 1 & \cdots & J - 1
\end{bmatrix}.
\]

Then it is easily seen that \( \Pi(\lambda) \) can be decomposed as

\[
\Pi(\lambda) = \lambda^{J-1} P^{-1}(\lambda) Q(\lambda) P(\lambda),
\]

(3.5)

where \( P(\lambda) = \text{diag}(\lambda^0, \lambda^{\theta_1}, \ldots, \lambda^{\theta_1}) \)

\[
Q(\lambda) = [q_{ij}(\lambda)] = (G - H) - \left(E + \frac{(1 + \gamma)}{\gamma} G - 2H\right)\lambda + \frac{1}{\gamma}(G - H)\lambda^2
\]

\[
= S'\zeta(I + (\zeta^{-1})\lambda + (I/\gamma)\lambda^2)S,
\]

(3.6)

where \( \zeta \) is a diagonal matrix with the eigenvalues of \(- (C^{-1})D\) and \( S^{-1} \) is the matrix with the corresponding eigenvectors.

Eq. (3.4) will now be rewritten in a form that is suitable for deriving constraints on the revision processes (for derivation see also Appendix A)

\[
\Pi(L)y_t = \Pi(L) \left\{ \sum_{i=0}^{J-1} \varepsilon_i^{J-i} \right\} + \xi_{t-J} - u_{t-J}.
\]

(3.7)

\(^4\) Hereafter the subscript of the revision processes will refer to a production factor or indicate time.

\(^5\) This particular decomposition emerged quite naturally. Broze et al. (1995) propose two general decomposition methods: the adjoint operator method and the Smith form method.
where
\[
\eta_{t-j} = - \sum_{j=0}^{J-1} \sum_{h=1}^{J} A_h L^{j+h} \xi_{t-j} - \sum_{j=0}^{J-1} \sum_{h=0}^{J} A_h L^{j-h} \xi_{t-j}.
\]

Using the fact that \(\sum_{j=0}^{J-1} \sum_{h=1}^{J} j^h = y_t - E_{t-J} y_t\), we obtain
\[
\Pi(L) E_{t-J} y_t = \xi_{t-J} - u_{t-J}.
\]

Replacing \(\Pi(L)\) in the recursive equation (3.8) with its decomposition (3.5) and premultiplying both sides by \(L^{-\theta_1} P(L)\) yields
\[
L^{-\theta_1} Q(L) P(L) E_{t-J} y_t = L^{-\theta_1} P(L) \{\xi_{t-J} - u_{t-J}\}. \tag{3.9}
\]

The premultiplication shifts the equations forward in time. The LHS of the equation above contains only current and lagged values of \(y\). Note that \(P(L) y_t = x_t\). If \(\theta_1 > 0\) the components of the second term at the RHS are based on information after time \(t - J\). This observation is at the root of the following lemma.

\textbf{Lemma 1.} In order that \(y = \{y_t\}\) is a solution of Eq. (3.2), the revision processes \(\xi^i_t\) have to satisfy the constraints
\[
E\{L^{-\theta_1} P(L) [\xi_{t-J} - u_{t-J}] | \Omega_{t-i}\}
= E\{L^{-\theta_1} P(L) [\xi_{t-J} - u_{t-J}] | \Omega_{t-i-1}\}, \quad i = 0, \ldots, J - 1. \tag{3.10}
\]

\textbf{Proof.} Take expectations of both sides of Eq. (3.9) with respect to the information sets \(\Omega_{t-i}, i = 0, \ldots, J\), and calculate the difference between two adjacent expressions.\(^6\) See also Broze et al. (1995). \(\square\)

The dimension of the set of solutions that satisfy Eq. (3.1) depends on the number of free revision processes. The number of effective (univariate) martingale differences in Eq. (3.4) equals \((\theta_1 + 1) F - \sum_{i=1}^{F} \theta_i\). After imposing the constraints in Eq. (3.9), there are only \(F\) free revision processes left as we will prove now.

\textbf{Lemma 2.} The restrictions in Eq. (3.10) reduce the number of free revision processes in Eq. (3.4) by \(\theta_1 F - \sum_{i=1}^{F} \theta_i\).\(^7\)

\(^6\) Since \(\xi_{t-J}\) depends only on information up to \(t - J\), in some cases the equalities in (3.10) hold trivially and do not imply restrictions on the revision processes. For example, subtracting the expectation of the first equation in (3.7) with respect to \(\Omega_{t-i-1}\) from the expectation of the same equation with respect to \(\Omega_{t-i}\) for \(i = 0, \ldots, \theta_1 - \theta_F\) does not yield a single constraint.

\(^7\) This is in fact predicted by property (37) in the paper by Broze, Gouriéroux and Szafarz (1995). However we believe that our proof, as it is confined to the particular model above, is more revealing and complete than their proof.
Proof. See Appendix B.

So all the revision processes in Eq. (3.4) but $F$ are subject to restrictions. The space of candidate solutions to the optimal control problem that satisfy the Euler–Lagrange conditions (3.2) will be reduced even further by imposing the transversality condition or a slightly stronger condition like a stability condition. Such conditions guarantee uniqueness of the solution just as they do when there is no time-to-build. In the latter case $\Pi(\lambda) = Q(\lambda)$.

Recall the decompositions in Eqs. (3.5) and (3.6). Then premultiplying both sides of Eq. (3.4) by $(S'\xi)^{-1}L^{-\theta}P(L)$ leads to

$$
(I + (\xi^{-1})L + (I/\gamma)L^2)SL^{-\theta}P(L)y_i
$$

$$
= L^{-\theta}(S'\xi)^{-1}\left\{\sum_{i=0}^{J-1} \sum_{j=i+1}^{J} P(L)A_jL^{1-j+i}c_i^j - P(L)u_{t-1}\right\}. \tag{3.11}
$$

By virtue of Eq. (3.10), the number of free revision processes in Eq. (3.11) can be reduced to $F$, say $\hat{c}_0^0$. If we look for a solution $y_i$ that belongs to the class of VARIMA-processes, we can parameterize $\hat{c}_i^0$ as $\Phi^0 v_i$. Likewise, we can write $\hat{c}_i^j$ as $\Phi^j v_i$, $i = 1, \ldots, J - 1$, where the $\Phi^j$ are $F \times F$ matrices. Thus the RHS of Eq. (3.11) is a distributed lag of $v_i$, $R(L)v_i = (R_0 + R_1 L + \cdots + R_g L^g)v_i$ with $g = J + \tau$.

At this point let us introduce some notation for later. We denote the vectors comprising the effective revision processes by $\tilde{c}_i^j$. Further, we implicitly define the matrices $\tilde{\Phi}^j$ by $\tilde{c}_i^j \equiv \tilde{\Phi}^j v_i$. Now, the $\tilde{\Phi}^j$’s are subject to restrictions like Eq. (3.10). On the other hand, we note that knowledge of the rows of the $\Phi^j$ corresponding to the other revision processes in the $\hat{c}_i^j$ is not required because they cancel out in Eq. (3.11). Finally, we define the matrix $\tilde{\Phi} \equiv [(\tilde{\Phi}^0)' \ldots (\tilde{\Phi}^{J-1})']'$ and the vector $\tilde{e} \equiv [(\hat{e}^0)' \ldots (\hat{e}^{J-1})']'$.

A solution of the stochastic optimal control problem must also satisfy the transversality condition. Sufficient conditions for this condition to hold are that the exogenous stochastic process $v_i$ and a solution $y_i$ are of exponential order less than $\gamma^{-1/2}$ (see Sargent, 1987, pp. 200–201). The $u_t$ process has this property by assumption. To verify whether $y_i$ has this property, we investigate the LHS of Eq. (3.11). Let us define $\tilde{Q}(\lambda) \equiv (I + (\xi^{-1})\lambda + (I/\gamma)\lambda^2)$. Denote the roots of the $f$th equation in $\tilde{Q}(1/\lambda) = 0$ by $\lambda_{1f}$ and $\lambda_{2f}$. All pairs of roots $\{\lambda_{1f}, \lambda_{2f}\}$, $f = 1, \ldots, F$, satisfy $\lambda_{1f}/\lambda_{2f} = \gamma^{-1}$. Let $\lambda_{1f}$ be the smaller modulus root of each pair. Then we have $|\lambda_{1f}| \leq \gamma^{-1/2}$ and $|\lambda_{2f}| \geq \gamma^{-1/2}$. The stability condition given in Kollintzas

---

8 See Kollintzas (1985).

9 Generally, the matrix that results from leaving out the first $F$ columns from the matrix at the LHS of (A2.2) is nonsingular. See also footnote 14.
(1985) is necessary and sufficient for \(|\lambda_{1f}| < \gamma^{-1/2}\) to hold. For instance, this condition is satisfied when the adjustment costs are strongly separable \((H = 0)\). Notice that if \(|\lambda_{1f}| < \gamma^{-1/2}\), then \(\dot{\lambda}_1, \dot{\lambda}_2 \in \mathbb{R}\). But in order that the solution \(y_t\) is of exponential order less then \(\gamma^{-1/2}\), we also need that \(R(L)\) can be factorized as \((I - A^+L)\tilde{R}(L)\), where \(A^+ = \text{diag}(\lambda_{21}, \ldots, \lambda_{2F})\) and \(\tilde{R}(L)\) is a lag polynomial of order \(g - 1\), for otherwise we are not able to get rid of that part of the LHS of Eq. (3.11) that causes violation of the condition, i.e. the factor \((I - A^+L)\).

The lemma below provides a condition that is equivalent to the existence of the factorization but more easily verifiable.

**Lemma 3.** Let \(N(\lambda) = N_0 + N_1 \lambda + \cdots + N_w \lambda^w\) be a polynomial of order \(w\). This polynomial can be factorized as \((I - \Gamma \lambda)\bar{N}(\lambda)\) if and only if

\[
\sum_{i=0}^{w} \Gamma^{w-i}N_i = 0. \tag{3.12}
\]

**Proof.** See Appendix C.

Substituting \(R(L)\) for \(N(\lambda)\) and \(A^+\) for \(\Gamma\) gives the condition

\[
\sum_{i=0}^g (A^+)^g-iR_i = 0. \tag{3.13}
\]

The matrices \(R_0, \ldots, R_g\) depend on \(F \times F\) unknown entries of \(\Phi^0\). But Eq. (3.13) imposes \(F \times F\) restrictions on the \(\Phi^0_{ij}\)’s. In general, these restrictions uniquely identify \(\Phi^0\). To see this consider the RHS of Eq. (3.11) apart from the last term

\[
L^{-\theta}(S'\zeta)^{-1} \sum_{i=0}^{J-1} \sum_{j=1}^{J} P(L)A_jL^{1-j+i}\epsilon_i \equiv (\tilde{R}_0 + \tilde{R}_1L + \cdots + \tilde{R}_{J-1}L^{J-1})\epsilon_i, \tag{3.14}
\]

where \(\epsilon_i = (\epsilon_i^0 \cdots \epsilon_i^{J-1})\), \(\tilde{R}_i \equiv (S'\zeta)^{-1}[K_{i1} \ldots K_{iJ}], i = 0, \ldots, J - 1\); the \(\tilde{R}_i\)’s are \(F \times (J \times F)\) matrices and the \(K_{ij}\)’s are \(F \times F\) matrices with \(K_{ij} = 0, j > i + 1, K_{ij} = \sum_{k=1}^{\infty} A_{J+j-i-1,k}, 1 \leq j \leq i + 1\), \(^{10}\) where

\[
\tilde{A}_{J,1} = \begin{bmatrix} 0 & \ldots & 0 & \gamma d_{1F} \\ 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 \end{bmatrix}, \quad \tilde{A}_{J-1,1} = \begin{bmatrix} 0 & \ldots & 0 & -\tilde{c}_{1F} \\ 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 \end{bmatrix},
\]

\(^{10}\) \(\tilde{A}_{ij}\) matrices with the same indices should be added up.
\[\bar{A}_{J-2,1} = \begin{bmatrix} 0 & \ldots & 0 & \gamma d_{1F} \\ 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 \end{bmatrix},\]

\[\bar{A}_{J-3,1} = 0, \ldots, \bar{A}_{\theta_1 - \theta_{F-1} + 1,1} = \begin{bmatrix} 0 & \ldots & 0 & \gamma d_{1F-1} & 0 \\ 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & 0 \end{bmatrix},\]

\[\bar{A}_{\theta_1 - \theta_{F-1},1} = \begin{bmatrix} 0 & \ldots & 0 & -\bar{c}_{1F-1} & 0 \\ 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & 0 \end{bmatrix},\]

\[\bar{A}_{1,1} = \begin{bmatrix} \gamma d_{11} & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix},\]

\[\bar{A}_{J,2} = \begin{bmatrix} 0 & \ldots & 0 & 0 \\ 0 & \ldots & 0 & \gamma d_{2F} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 \end{bmatrix}, \quad \bar{A}_{J-1,2} = \begin{bmatrix} 0 & \ldots & 0 & 0 \\ 0 & \ldots & 0 & -\bar{c}_{2F} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 \end{bmatrix},\]

\[\bar{A}_{\theta_1 - \theta_2 + 1,2} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & \gamma d_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix},\]

\[\bar{A}_{J,F} = \begin{bmatrix} 0 & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & 0 \\ 0 & \ldots & 0 \end{bmatrix} \quad \text{with } \bar{c}_{ij} \equiv \gamma c_{ij}, \; i,j = 1,2,\ldots,F,\]
\( \bar{f}(i) = f \) for \( \theta_{f+1} - \theta_F < i \leq \theta_f - \theta_F \) with \( f = 1, \ldots, F, i = 0, \ldots, J - 1 \) and \( \theta_{F+1} \equiv \theta_F - 1 \).

Define

\[
B_i \equiv \sum_{j=1}^{J-1} (A^+)_{ij} (\gamma \zeta')^{-1} K_{ji+1}, \quad i = 0, \ldots, J - 1. \tag{3.15}
\]

Then condition (3.13) can be rewritten in the form

\[
\sum_{i=0}^{J-1} B_i \varepsilon_t^i + B \bar{v}_i = 0, \tag{3.16}
\]

where

\[
\bar{B} = -\sum_{i=0}^{J-1} (A^+)_{ij} (\gamma \zeta')^{-1} K_{ji+1} \quad \text{with} \quad K_{10} = \bar{U}_{0,F}, K_{20} = \bar{U}_{1,F}, \ldots, K_{\theta_f-\theta_F+1,0} = (\bar{U}_{\theta_f-1,0} + \bar{U}_{0,F-1}), \ldots, K_{\theta_0} = \bar{U}_{s,1}, \text{where} \quad \bar{U}_{s,1} \text{is a zero matrix with the} \quad s^{th} \text{row replaced by} \quad \bar{U}_{s,1}.
\]

Combining the restrictions on the revision processes in (3.10) and (3.16) gives

\[
\begin{bmatrix}
A_0 & A_1 & A_2 & \ldots & A_{J-2} & A_{J-1} & \varepsilon_t^0 \\
A_{J-1} & A_0 & A_1 & \ldots & A_{J-3} & A_{J-2} & \varepsilon_t^1 \\
A_{J-2} & A_{J-1} & A_0 & \ldots & A_{J-4} & A_{J-3} & \varepsilon_t^2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
A_{J-2} & A_{J-3} & A_{J-4} & \ldots & A_0 & A_1 & \varepsilon_t^{J-2} \\
B_0 & B_1 & B_2 & \ldots & B_{J-2} & B_{J-1} & \varepsilon_t^{J-1}
\end{bmatrix}
\begin{bmatrix}
U_0 \\
U_1 \\
\vdots \\
U_{\tau^*} \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
\varepsilon_t^0 \\
\varepsilon_t^1 \\
\varepsilon_t^2 \\
\vdots \\
\varepsilon_t^{J-2} \\
\varepsilon_t^{J-1}
\end{bmatrix} +
\begin{bmatrix}
\bar{B} \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
= 0.
\tag{3.17}
\]

where \( \tau^* = \min (\tau, J - 1) \).

We can write the system in (3.17) more compactly as

\[
W \varepsilon_t + Z \bar{v}_i = 0.
\tag{3.18}
\]

where \( W \) and \( Z \) are defined implicitly.

We note that, as mentioned in the proof of Lemma 2, \(-\theta_F + \sum_{i=1}^{F} \theta_i\) 'restrictions' in the system are false but they are included for convenience: it is easier to show the independence of the true restrictions by looking at the extended system as will become clear below. Furthermore in the true restrictions the revision processes in Eq. (3.17) that are not mentioned below Eq. (3.4) are multiplied by zeros and cancel out. For the last \( F \) rows this follows from the observation that Eq. (3.16) is obtained from Eq. (3.14) by first decomposing the \( A_j \) in terms, then premultiplying each term by \( A \) raised to some power, where the power depends on the lag of the corresponding revision process, and finally collecting terms (on the basis of the horizons of the processes). Thus the vectors of the revision processes in Eq. (3.16) (which differ in terms of their
horizons) are multiplied by the same rows of $A_j$ as in Eq. (3.14). The latter is obtained by shifting the rows in Eq. (3.4) back in time and by premultiplying the result by a constant matrix.

We write the system of the true restrictions contained in Eq. (3.18) as

$$
\bar{W} \bar{e}_t + \bar{Z} \bar{v}_t = 0.
$$

(3.18')

For uniqueness of the linear solution $\{y_t\}$ to Eq. (3.1) it is sufficient that $W$ is nonsingular. We will demonstrate that in general this condition is satisfied. In the model without time-to-build $W = A_1 = G$ is nonsingular.

**Theorem 1.** The matrix $W$ defined in Eqs. (3.17) and (3.18) is generally nonsingular, that is $\forall \gamma, M_{W,\gamma} = \{E,G,H|\det(W_\gamma) = 0\}$, the space of values of the entries of the matrices $E$, $G$, and $H$ for which $\det(W_\gamma) = 0$, has Lebesgue measure zero.

**Proof of Theorem 1.** See Appendix D.

The uniqueness of the solution to the model guarantees the existence of the closed form solution (CFS)

$$
x_t = A x_{t-1} + S^{-1} \tilde{R}(L)L^\theta v_t,
$$

(3.19)

where $A = S^{-1} A^-$, and $A^- \equiv \text{diag} (\lambda_{11}, \ldots, \lambda_{1F})$.

From Eq. (3.19) it is clear that the transversality condition can also be imposed on the solutions of the model with multiple lags.

Let $M = I - A$. If $|M| \neq 0$, then Eq. (3.19) admits the flexible accelerator form

$$
x_t - x_{t-1} = M(x_t^s - x_{t-1}^s) \quad \text{with} \quad x_t^s = M^{-1} S^{-1} \tilde{R}(L)L^\theta v_t.
$$

(3.20)

Notice that the autoregressive part of the closed form solution is invariant with respect to the order of the gestation lags. Our strategy for deriving expressions of the MA parameters $\tilde{R}_i$ in the CFS in terms of the structural parameters entails the following steps:

1. Obtain expressions for $\tilde{e}_t$: from Eq. (3.18') $\tilde{e}_t = - \bar{W}^{-1} \bar{Z} v_t$, i.e. $\tilde{\Phi} = - \bar{W}^{-1} \bar{Z}$.
2. Substitute these expressions in the RHS of Eq. (3.11) and find formulae for $R_0$ up to $R_g$, $g = J + \tau$. Make use of Eq. (3.14).
3. Use the procedure in the proof of Lemma 3 to obtain the factorization $R(L) = (1 - A^+ L)\tilde{R}(L)$, i.e. $\tilde{R}_0, \tilde{R}_1, \ldots, \tilde{R}_{g-1}$.

For nontrivial orders of the (gestation) lags, the MA parameters in Eq. (3.19) depend in a complicated way on the structural parameters and estimation of the latter is not straightforward when these restrictions are to be exploited.
To estimate the structural parameters in the CFS an iterative procedure can be used which entails the three steps mentioned above but carried out in reverse order. At each iteration one has to invert the (sparse) matrix \( \bar{W} \) in Eq. (3.18) to obtain \( \bar{\Phi} \). We note that the dimensions of \( \bar{W} \) (and also \( W \)) do not depend on \( \tau \). Furthermore, one can reduce the computational burden by replacing \( S \) and \( A^+ \) in Eq. (3.17) (or Eq. (3.18')) by estimates based on an initial consistent estimate of \( A \),\(^{11} \) which would result in a system that is linear in the technology parameters and the \( U_t \). Alternatively, one can of course replace the \( S \) and \( A^+ \) in Eq. (3.17) by new estimates at each iteration. So in practice we have found a representation of the solution that is as close as possible to a closed form representation.

The identification of the structural parameters hinges on whether the factor demand equations are estimated in a system together with the production (profit) function and (observable) exogenous processes or separately. If the latter are not included in the system, the MA parameters provide no information on the structural parameters. If the production (profit) function is not included, one can at best identify the parameters in \( C \) and \( D \) (up to a constant factor) but not those in \( E, G, \) and \( H \).

When one ignores the restrictions on the MA parameters, one still has a semi closed form representation of the solution. Epstein and Yatchew (1985) discuss an exhaustive set of properties of the accelerator matrix, \( M \), that are implied by the structure of the problem. By demanding that the estimates of \( M \) exhibit these properties and by estimating the factor demand equations together with the production function and the price equations, more precise estimates of the structural parameters can be obtained.

At the outset we assumed that \( u_t \sim \text{VMA}(\tau) \). However, prices sometimes follow more general VARIMA processes. In such cases the formula of the CFS should be amended by using the following procedure. The autoregressive lag polynomial of the exogenous processes \( \{u_t\} \) can be factorized such that one factor is a diagonal matrix lag polynomial with common roots on the diagonals: \( \bar{V}(L) \). For instance, if the prices are IMA processes this factor includes the unit roots. Then the VAR1 lag polynomial of the solution becomes \( \bar{V}(L)(I - AL) \). If the remainder of \( \{u_t\} \) is stationary, then that part can be represented as a VMA. The stability condition can still be imposed. However the procedure to obtain expressions for the MA parameters in the CFS in terms of the structural

\(^{11}\) First, notice that the eigenvalues of \( A \) are equal to the diagonal elements of \( A^{-} \) and recall that \( \gamma A^{-} A^{+} = I \). Therefore, one can obtain an estimate of \( A^{+} \) from an estimate of \( A \) (given \( \gamma \)). Second, \( A \) also provides an estimate of \( S \) up to a diagonal matrix, i.e. if \( S^0 \) is such that \( A = (S^0)^{-1} A^{-} S^0 \), then \( S^* = D^* S^0 \) also satisfies \( A = S^{-1} A^{-} S \). But notice that the diagonal matrix (and also \( \zeta^{-} \)) cancels out in the stability restrictions (the last \( F \) restrictions in Eq. (3.17)). Moreover, the (true) \( S(D) \) is identified by the RHS of Eq. (3.19) through \( C \) and \( D \).
parameters becomes unfeasible when the order of the MA is high ($\infty$). In practice the MA-part of the solution will often be approximated by a VARMA model.\textsuperscript{12}

In the model without adjustment costs, we do not need to impose stability to obtain a unique solution; the restrictions in Lemma 1 are sufficient.

Finally, similar results as those obtained above can also be derived for nonsymmetric LRE models ($H \neq H'$) with multiple lags. As examples of such models one can think of a model with time-to-build and gross adjustment costs or a model with time-to-build and net adjustment costs stemming from new investments, i.e. changes in the y’s.\textsuperscript{13} A decomposition of the second order matrix lag polynomial of the nonsymmetric model can be found in Kollintzas (1986). He also gives a stability condition for such models. Again, imposing the stability condition yields a unique solution to the model. However, the derivation of a closed form representation of the solution to the model (especially for the MA-part) is likely to be very complicated as the factor $A^+$ in the formulae above has to be replaced by a non-diagonal matrix. Nevertheless, the solution admits an accelerator formulation where the accelerator matrix $M$ has a structure very similar to that in Eqs. (3.19) and (3.20). Again, for (efficient) estimation one can exploit cross-equation restrictions between $M$ and the production function parameters and properties of $M$ (see Madan and Prucha, 1989).

4. Concluding comments

In this paper we have considered a generalized version of Kollintzas’ (1985) symmetric LRE model which allows for multiple lags in the transition equations of the state variables. We have shown that this model admits a unique stable solution and we have derived a representation of that solution that is as close as possible to a closed form. We have also seen that the stability condition given in Kollintzas (1985) still applies to the solutions of the generalized model. Kollintzas has shown that his condition is equivalent to a discrete time version of the stability conditions that can be found in Magill (1977) and Magill and Scheinkman (1979). These conditions impose a bound on the benefit-cost ratio of deviations from the equilibrium stock levels.

\textsuperscript{12} This does not undo the possibility to impose cross-equation restrictions though, when the CFS is estimated as part of a system of equations.

\textsuperscript{13} This adjustment cost function can be found in Madan and Prucha (1989). In the symmetric model presented in Section 2, the adjustment costs stem from completed investments, that is changes in the stocks of productive inputs ($x$). We notice that the so-called symmetric LRE with multiple lags is asymmetric in the decision variables unless the gestation lags (the $\theta_j$’s) are the same.
By allowing for multiple lags, the LRE model discussed in this paper yields solutions that follow VARMA processes with more general MA parts. Although the restrictions between the MA parameters of the CFS and the structural parameters can be too complicated to be exploited, the restrictions on the AR parameters can be imposed. Indeed, estimation of the demand equations together with the production function would allow imposition of cross-equation restrictions to obtain more precise estimates of the parameters.

Acknowledgements

This research was partly sponsored by the Economics Research Foundation, which is part of the Netherlands Organization for Scientific Research (NWO). The author is indebted to A. Kirman, A. Perea y Monsuwé, and S. Schim van der Loeff, and to two anonymous referees for helpful suggestions. All errors are solely mine.

Appendix A

A.1. Proof of Eq. (3.4)

Eq. (3.4) is the result of the following computations. The system we are investigating is

\[ E_t L^{-J} \Pi(L) y_t + u_t = 0, \quad (3.2) \]

substituting

\[ \Pi(L) = \sum_{j=-J}^{J} A_j L^{-j} \quad (3.3) \]

and lagging the system \( J \) periods gives

\[ E_{t-J} \sum_{j=-J}^{J} A_j L^{-j} y_{t-j} + u_{t-J} = 0 \]

\[ \Leftrightarrow \Pi(L)y_t + \left\{ E_{t-J} \sum_{j=1}^{J} A_j L^{-j} y_{t-j} \right\} - \sum_{j=1}^{J} A_j L^{-j} y_{t-j} + u_{t-J} = 0 \]

\[ \Leftrightarrow \Pi(L)y_t + \sum_{j=1}^{J} A_j \{ E_{t-J} y_{t-J+j} - y_{t-J+j} \} + u_{t-J} = 0 \]
recalling that \( e^j_t = E_t y_{t+j} - E_{t-1} y_{t+j} \), we obtain

\[
\Pi(L)y_t + \sum_{j=1}^{J} A_j \left\{ - \sum_{i=0}^{j-1} e^i_{t-j+i} \right\} + u_{t-j} = 0
\]

changing the order of summation yields

\[
\Pi(L) y_t = \sum_{i=0}^{J-1} \sum_{j=i+1}^{J} A_j L^{j-i} e^i_t - u_{t-j}. \quad (3.4)
\]

**A.2. Proof of Eq. (3.7).**

Starting with the result just shown, we have

\[
\Pi(L) y_t = \sum_{j=0}^{J-1} \sum_{h=j}^{J} A_h L^{j-h} e^j_t - u_{t-j}
\]

\[
\iff \Pi(L) y_t = \Pi(L) \left\{ \sum_{j=0}^{J-1} e^j_t \right\} - u_{t-j} - \sum_{j=0}^{J-1} \sum_{h=j}^{J} A_h L^{j-h} e^j_t
\]

\[
\iff \Pi(L) y_t = \Pi(L) \left\{ \sum_{j=0}^{J-1} e^j_t \right\} - u_{t-j} - \sum_{j=0}^{J-1} \sum_{h=j}^{J} A_h L^{j-h} e^j_t
\]

\[
\iff \Pi(L) y_t = \Pi(L) \left\{ \sum_{j=0}^{J-1} e^j_t \right\} - u_{t-j} - \sum_{j=0}^{J-1} \sum_{h=j}^{J} A_h L^{j-h} e^j_t
\]

\[
\iff \Pi(L) y_t = \Pi(L) \left\{ \sum_{j=0}^{J-1} e^j_t \right\} - u_{t-j} - \sum_{j=0}^{J-1} \sum_{h=j}^{J} A_h L^{j-h} e^j_t
\]

\[
\iff \Pi(L) y_t = \Pi(L) \left\{ \sum_{j=0}^{J-1} e^j_t \right\} - u_{t-j} - \sum_{j=0}^{J-1} \sum_{h=j}^{J} A_h L^{j-h} e^j_t
\]

\[
\iff \Pi(L) y_t = \Pi(L) \left\{ \sum_{j=0}^{J-1} e^j_t \right\} - u_{t-j} - \sum_{j=0}^{J-1} \sum_{h=j}^{J} A_h L^{j-h} e^j_t
\]
which can be written as

$$
\Pi(L)y_t = \Pi(L) \left\{ \sum_{j=0}^{J-1} \tilde{\xi}_{t-j}^j \right\} + \tilde{\xi}_{t-J} - u_{t-J}, \tag{3.7}
$$

where

$$
\tilde{\xi}_{t-J} = - \sum_{j=0}^{J-1} \sum_{h=1}^{J} A_{h} L^{j+h+j} \tilde{\zeta}_{t} - \sum_{j=0}^{J-1} \sum_{h=0}^{J} A_{h} L^{j-h+j} \tilde{\zeta}_{t}.
$$

### Appendix B. Proof of Lemma 2

The proof of Lemma 2 consists of three steps. First we will demonstrate that the restrictions of Lemma 1 do not involve revision processes that are different from those mentioned just below Eq. (3.4). Next we show that \((1-\theta_F)F + \sum_{i=1}^{F} \theta_i\) restrictions are redundant. Finally, we look at a system which includes the effective restrictions. To carry out the first step, we will concentrate on the restriction which corresponds to taking expectations with respect to \(\Omega_{t-J} + 1\) and \(\Omega_{t-J}\) and taking differences afterwards, since this restriction contains the revision processes with the most distant horizons and we could expect new revision processes to show up here, if anywhere.

Consider, without loss of generality, the \(i\)th row of \(L^{-\theta_F}P(L)\tilde{\xi}_{t-J} \),

$$
\tilde{\xi}_{i,t-1-\theta_F} = - \sum_{j=0}^{J-1} \sum_{h=1}^{J} A_{h,i} L^{1+0,-\theta_F + h} \tilde{\zeta}_{i,t} - \sum_{j=0}^{J-1} \sum_{h=0}^{J} A_{h,i} L^{1+0,-\theta_F - h} \tilde{\zeta}_{i,t}.
$$

(B.1)

The most recent observation of the \(k\)th production factor \(y_k\) that appears in (B.1) is \(y_{k,t-\theta_F} \). To show this we consider the following two cases.

If \(\theta_k > \theta_i + 1\) the most recent observation of \(y_k\) in (B.1) is part of the first double sum and corresponds to the smallest value of \(h\) for which \(A_{h,ik} \neq 0\), that is to \(h = \theta_k - \theta_i - 1\). Note that in this case \(A_{h,ik} = 0\) for \(h \in \mathbb{N}\). If \(\theta_k \leq \theta_i + 1\) the most recent observation of \(y_k\) in Eq. (B.1) is part of the second double sum and corresponds to the highest value of \(h\) for which \(A_{h,ik} \neq 0\), that is to \(h = \theta_i - \theta_k + 1\). Thus the revision processes with the most distant horizons are \(\tilde{\zeta}_{k,\theta_k - \theta_i}\), \(k = 1, \ldots, F\). But they are included in the set of those mentioned below Eq. (3.4). This observation completes the first step.

The redundancy is easily demonstrated. Again we look at the \(i\)th row of \(L^{-\theta_F}P(L)[\tilde{\xi}_{t-J} - u_{t-J}]\). Since \(\tilde{\xi}_{t-J} - u_{t-J}\) contains information only until time \(t - J\), taking expectations of the \(i\)th row of \(L^{-\theta_F}P(L)[\tilde{\xi}_{t-J} - u_{t-J}]\) with respect to \(\Omega_{t-\theta_F + \theta_i - 1}\) is equivalent to taking expectations of the \(i\)th row with respect to \(\Omega_{t-\theta_F + \theta_i - 1 - j}\) for each \(j \in \mathbb{N}\). So at least \((1-\theta_F)F + \sum_{i=1}^{F} \theta_i\) restrictions are redundant. The remaining restrictions of Eq. (3.10) are included in the following
system, where $\tau^* = \min(\tau, J - 1)$

\[
\begin{bmatrix}
    A_0 & A_1 & A_2 & \cdots & A_{J-2} & A_{J-1} \\
    A_{-1} & A_0 & A_1 & \cdots & A_{J-3} & A_{J-2} \\
    A_{-2} & A_{-1} & A_0 & \cdots & A_{J-4} & A_{J-3} \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    A_{-J+2} & A_{-J+3} & A_{-J+4} & \cdots & A_0 & A_1
\end{bmatrix}
\begin{bmatrix}
    e_t^0 \\
    e_t^1 \\
    e_t^2 \\
    \vdots \\
    e_{t-J}^{J-1}
\end{bmatrix}
\begin{bmatrix}
    U_0 \\
    U_1 \\
    \vdots \\
    0
\end{bmatrix}
+ U_{\tau^*} v_t = 0
\]

(B.2)

(B.2) is obtained by taking expectations of the $i$th row of $L^{-\theta_i} P(L)[z_{t-j} - u_{t-j}]$ with respect to $\Omega_{t-\theta_i + \theta_j - j}$ for $j = 1, \ldots, J$ and taking differences of the adjacent expressions, for $i = 1, \ldots, F$. Apart from the restrictions in Eq. (3.10), system (B.2) also includes $-\theta_1 F + \sum_{i=1}^{F} \theta_i$ incorrect but noninformative equations which therefore can be ignored.\(^{14}\)

In order to prove that the remaining restrictions from Eq. (3.10) are effective, it will now be shown that generally the matrix at the LHS of (B.2) has maximum rank. The Laplace development of the determinant of the matrix that is obtained by leaving out the last $F$ columns, includes the product of the diagonal entries of that matrix, which are the diagonal entries of all $A_0$ matrices. Only the diagonal entries of $A_0$ depend on the diagonal entries of $E$. Since the sum of their powers in other terms of the Laplace development is lower, the determinant differs from zero in general.\(^{15}\) But if all rows in (B.2) are independent, then the $\theta_1 F - \sum_{i=1}^{F} \theta_i$ true restrictions are also independent. Finally, from the first step of the proof, where it was shown that no revision processes are included in Eq. (3.10) other than those mentioned below Eq. (3.4), it is immediately clear that by taking expectations of Eq. (3.9) with respect to $\Omega_{t-j-j}$ with $j \in \mathbb{N}$ one does introduce

\(^{14}\)The vector $[(\omega)'] \cdots (\omega_{J-1})']$ in (B.2) also includes $-\theta_1 F + \sum_{i=1}^{F} \theta_i$ revision processes which are not mentioned below Eq. (3.4). They only enter the false ‘restrictions’, i.e. the restrictions which do not emanate from Eq. (3.10), so the number of additional ‘restrictions’ in (B.2) equals the number of (additional) revision processes introduced in these equations. It can easily be verified that the additional equations in Eq. (B.2) do not constrain the revision processes mentioned below Eq. (3.4), by looking at the diagonal of the matrix at the LHS of (B.2) which is comprised of $A_1$-matrices. Generally, all diagonal entries of $A_1$ are different from zero since $G$ is positive definite and it is reasonable to assume that $G_{ii} - H_{ii} \neq 0$ $\forall i$. Now, each of the $-\theta_1 F + \sum_{i=1}^{F} \theta_i$ additional revision processes is multiplied by such an entry. This also enables us to locate all the false ‘restrictions’.

\(^{15}\)One corollary to Lemma 4 below would say that the set of values of the diagonal of $E$ for which the determinant is zero has Lebesgue measure zero.
new revision processes but also that the number of free revision processes remains constant: one can derive a new restriction for each new revision process that is introduced. □

Appendix C. Proof of Lemma 3

Suppose this factorization is possible. Then
\[ N(\lambda) = N_0 + N_1 \lambda + \cdots + N_w \lambda^w = (1 - \Gamma \lambda) (\bar{N}_0 + \bar{N}_1 \lambda + \cdots + \bar{N}_{w-1} \lambda^{w-1}) = \bar{N}_0 + (N_1 - \Gamma \bar{N}_0) \lambda + (\bar{N}_2 - \Gamma \bar{N}_1) \lambda^2 + \cdots + (-\Gamma \bar{N}_{w-1}) \lambda^w. \]
Equating powers of \( \lambda \) gives
\[ N_0 = \bar{N}_0, \quad N_i = \bar{N}_i - \Gamma \bar{N}_{i-1}, \quad i = 1, \ldots, w-1, \quad N_w = -\Gamma \bar{N}_{w-1}. \]  
(C.1)

It follows that
\[ \sum_{i=0}^{w-1} \Gamma^{w-i} N_i = (\Gamma^w - \Gamma^{w-1} \Gamma) \bar{N}_0 + (\Gamma^{w-1} - \Gamma^{w-2} \Gamma) \bar{N}_1 + \cdots + (\Gamma - \Gamma I) \bar{N}_{w-1} = 0. \] But condition (3.12) is also sufficient. We propose a factorization and next verify its validity. Use the formulae for \( N_0 \) through \( N_w \) in (C.1) recursively to obtain \( \bar{N}_0, \bar{N}_1, \bar{N}_2, \ldots, \bar{N}_{w-1} \). Then \( \bar{N}_{w-1} = \sum_{i=0}^{w-1} \Gamma^{w-1-i} N_i \). Premultiplying both sides by \(-\Gamma\) and using condition (3.12) yields \(-\Gamma \bar{N}_{w-1} = N_w\) which is in agreement with the last equality in (C.1). □

Appendix D. Proof of Theorem 1

The proof of Theorem 1 proceeds as follows. First we will present and prove a lemma that would imply Theorem 1 immediately if the entries of \( W \) were linear functions of the entries of \( E, G \) and \( H \). Then we will show that Theorem 1 also holds more generally as it has been formulated. We will deal with nonlinearities by conditioning such that for given values of particular functions of the parameters, the matrix is linear in the free parameters. Then Lemma 4 applies once again. The proof is completed by showing that the lemma can be applied for arbitrary values of these functions.

Lemma 4.\(^{16}\) Let \( p: \mathbb{R}^n \to \mathbb{R} \) be a polynomial in \( n \) variables and let \( p \neq 0 \). Then \( M(p) = \{x \in \mathbb{R}^n \mid p(x) = 0\} \) has Lebesgue measure zero.

Proof of Lemma 4. The proof proceeds by induction over the degree of \( p \). Let the degree of \( p \) equal 0. Then \( p \) is a constant. As \( p \neq 0 \), \( p \) has no roots. Thus \( M(p) = 0 \), and has Lebesgue measure zero.

\(^{16}\) This proof of Lemma 4 was suggested to me by A. Perea y Monsuwé.
**Induction step:** Let the degree of \( p \) equal \( r \geq 1 \) and let the lemma be true for \( p \) with degree \( \leq r - 1 \).

Since the degree of \( p \) is at least 1, there is an \( i \in \{1, \ldots, n\} \) with \( \partial p / \partial x_i \neq 0 \). Assume w.l.o.g. that \( \partial p / \partial x_n \neq 0 \).

Define the function \( f(x) := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ p(x_1, \ldots, x_n) \end{bmatrix} \).

Because \( p(x) = 0 \) for all \( x \in M(p) \), it follows that \( f(M(p)) \subseteq \mathbb{R}^{n-1} \times \{0\} \). If \( f \) is a differentiable function, which is locally invertible almost everywhere, i.e. invertible for all \( x \in \mathbb{R}^n \setminus M_0 \) where \( M_0 \) is a set that has measure zero, then \( f \) is a local diffeomorphism almost everywhere. As will be shown below, we can cover \( M(p) \setminus M_0 \) by a countable union of open sets on each of which \( f \) is a (global) diffeomorphism. The image of the intersection of each of these open sets with \( M(p) \setminus M_0 \) under \( f \) is a subset of \( \mathbb{R}^{n-1} \times \{0\} \), which has measure zero in \( \mathbb{R}^n \). If \( f \) is a diffeomorphism on \( \mathbb{R}^n \setminus M_0 \), it follows that the intersection of each open set with \( M(p) \setminus M_0 \) has Lebesgue measure zero. Since \( M(p) \setminus M_0 \) is a countable union of sets with Lebesgue measure zero, \( M(p) \setminus M_0 \) has measure zero. From the fact that \( M_0 \) has measure zero, it follows that \( M(p) \) has measure zero.

Thus it remains to show that \( f \) is differentiable and locally invertible and that \( M(p) \subseteq \bigcup_{i \in I} B_i \), where the \( B_i \)'s are open sets on which \( f \) is a global diffeomorphism and \( I \) is a countable set.

It is easy to verify that \( f \) is differentiable. Next we focus on the property of local invertibility.

The derivative of \( f, D_x f \) is

\[
D_x f = \begin{bmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 0 \\ * & * & \ldots & * & \frac{\partial p}{\partial x_n} \end{bmatrix}
\]

Thus \( \det(D_x f) = \partial p / \partial x_n \). This is a function of \( x_1, \ldots, x_n \). We know that \( \partial p / \partial x_n \) is a polynomial in \( \mathbb{R}^n \) of degree \( \leq r - 1 \). From the induction assumption, it follows that \( M_0 = \{ x \mid \partial p / \partial x_n = 0 \} \) has measure zero. Thus \( \det(D_x f) \neq 0 \) \( \forall x \in \mathbb{R}^n \setminus M_0 \). Applying the implicit function theorem, it follows that \( f \) is locally invertible on \( \mathbb{R}^n \setminus M_0 \), where \( M_0 \) has Lebesgue measure zero.
To show that we can cover \( M(p) \) by a countable union of open sets on which \( f \) is a global diffeomorphism, we will first prove that \( M(p) \) is a closed set. \( M(p) \) is a closed set iff every convergent sequence \( x^k \in M(p) \) converges to a limit in \( M(p) \). Now let \( x^k \) be a sequence in \( M(p) \) converging to \( x \). Thus \( f(x^k) = 0 \forall k \). From the fact that \( f \) is continuous, it follows that \( f(x) = 0 \).

We can always cover \( M(p) \) by the union of open sets on each of which \( f \) is a diffeomorphism. We will show that we can find a countable number of such sets. Define \( M^k = M(p) \cap C^k \) where \( C^k = \{ x \in \mathbb{R}^n \mid \|x\|_\infty \leq k \} \).

Since \( C^k \) is compact and \( M(p) \) is closed, \( M^k \) is compact.

Thus we can find a cover \((B_i)_{i \in I_k}\) for \( M^k \), where \( I_k \) has a finite number of elements. It follows that \( M(p) = \bigcup_k M^k \subseteq \bigcup_k \bigcup_{i \in I_k} B_i = \bigcup_{i \in I_k} B_i \), where \( I = \bigcup_k I_k \) is a countable set.

\[ \square \]

**Proof of Theorem 1.** The determinant of a general matrix \( N \) is a polynomial, where the entries of the matrix are the variables. From Lemma 4 it follows that the space of values of the entries of \( N \) for which \( \det(N) = 0 \) has Lebesgue measure zero. The same lemma can be invoked when the entries are linear functions of (hyper-)parameters even if the number of parameters is smaller than the number of entries.

Since the entries of the last \( F \) rows of \( W \) are nonlinear functions of the entries of \( E, G, \) and \( H \), Theorem 1 has not been proven yet.

From Eq. (3.15) we know that the matrices \( B_i, i = 0, \ldots, J - 1 \), depend in a nonlinear fashion on \( A^+, S, \) and \( \zeta \). These matrices in turn depend on \((C^{-1})D\) (see the discussion following Eqs. (3.6) and (3.11)), where \( C \) and \( D \) are linear in (the entries of) \( E, G, \) and \( H \), given \( \gamma \) (see the discussion following Eqs. (2.1) and (3.3)). Thus \( W \) depends on \( C \) and \( D \). Define \( X \equiv - (C^{-1})D \). The conditions imposed on \( E, G, \) and \( H \) (see the discussion following Eq. (2.1)) imply that \( C \equiv E + G - 2H + (1/\gamma)G \) is PDS and \( D \equiv G - H \) is nonsingular and therefore \( X \) is nonsingular. By fixing two matrices from the set \( \{ C, D, X \} \) (i.e. by fixing the values of the entries in two matrices from the set \( \{ C, D, X \} \) without violating the conditions following Eq. (2.1), for instance \( C \) and \( X \), we fix the third matrix and we fix \( W \).

By fixing \( X \) we impose restrictions on \( C \) and \( D \) (\( E, G, \) and \( H \)) as \( CX + D = 0 \) must hold. Substitute these restrictions in Eq. (3.17) by replacing \( D \) by \( - CX \). Then given \( X \) (and \( \gamma \)), \( W \) is linear in \( C \). According to Lemma 4, \( M_{W,\gamma} \) has Lebesgue measure zero as long as \( \det(W) \) is not trivially zero. This is indeed not the case for almost all choices of \( X \), as will be shown now.

To this end we will develop the determinant of \( W \). Some terms in the Laplace development consist only of the product of the entries on the diagonals of the \( A_1 \)-matrices and \( F \) entries of \( B_0 \). The sum of all these terms will be shown to be unique, that is, its algebraic formula does not show up in another term of the Laplace development (LD) of the determinant of \( W \) and therefore cannot cancel out. Furthermore this sum is different from zero in general.
This can easily be checked in the special case where $E = 0$, $G = \tilde{G}$, and $H = 0$, where $\tilde{G}$ is an arbitrary PDS matrix. See Appendix E. From this simple example it follows that $\det(W)$ is not (always) the null function. Now let us return to the case of an arbitrary matrix $X$.

So given $X$, $\det(W)$ is not the null function unless a particular choice for $X$ nullifies it to start with. Only for a very exceptional choice for $X$ (a choice from a set of measure zero within the space of all possible choices for $X$) we might expect this to happen as will be shown now.

Observe that when we fix $X$, $C$ (or $D$) is still unrestricted; to be able to apply Lemma 4 given a particular value of $X$, it suffices to show that $\det(W) \neq 0$ for one particular choice of the value of $C$ (or $D$). As has already been indicated above, this will be done by looking for (algebraically) unique terms in the Laplace development of $\det(W)$. We will focus on the case where $C = \text{diag}(c_{11}, \ldots, c_{FF})$. Next we replace the $c_{ii}$’s by entries from $X$ and $D$. Note that the entries from $X$ and $D$ have to satisfy the constraints $X + \text{diag}(c_{11}, \ldots, c_{FF})^{-1}D = 0$. Note further that $X$ can still take any value by choosing $D$ appropriately and when the diagonal entries of $C$ are free but those of $D$ are not, by choosing the $c_{ii}$’s and the nondiagonal entries of $D$ appropriately.

In the remainder of the proof we will also exploit the fact that the $d_{ii}$’s only show up in the diagonals of $A_{-1}, A_0$, and of $A_{+1}$, and in $B_0$ but not in $B_i$, $\forall i \neq 0$. As the latter is not obvious, we will prove that $d_{ff}$, $f = 1, \ldots, F$ show up in $B_0$ but not in $B_i$, $\forall i \neq 0$, in Appendix F. Specifying only those terms that involve (one of) the $d_{ii}$’s, $B_0$ equals $\sum_{k=1}^{F} (A^+)^{g-\theta_{0}+\theta_{1}(\gamma S_{\gamma})^{-1}}A_{\theta_{1}-\theta_{0}+1,k} + \cdots$.

Now we will show that there exists a unique part in the Laplace development of $\det(W)$, LD-part for short, which is the sum of all LD-terms that include the factor $\prod_{k=1}^{F} d_{kk}^{1J_{kk}^{-1}}$ originating from the first $(J-1)F$ rows of $W$, and the factor $\prod_{k=1}^{F} d_{kk}$ originating from the last $F$ rows. This LD-part is meant to be unique in the sense that there are no other terms in the Laplace development of $W$ that include the factor $\prod_{k=1}^{F} d_{kk}$. If such a unique LD-part exists and if the values of the $d_{ii}$’s can be chosen freely given $X$, then $\det(W) \neq 0$.

Note that because the first $F(J-1)$ rows of $W$ only have $J-1$ rows which involve a specific $d_{ii}$, $i = 1, \ldots, F$, the highest power of $d_{ii}$ that could be encountered in the part of an LD-term that originates from the first $F(J-1)$ rows, is $J-1$. Furthermore the last $F$ rows of $W$ contribute at most $F$ $d_{ii}$’s to an LD-term, which will all be different and originate form the first $F$ columns, that is, from $B_0$.

Thus in order to obtain an LD-term that includes $\prod_{k=1}^{F} d_{kk}$ both parts of the $W$ matrix – the upper $F(J-1)$ rows and the lower $F$ rows – must contribute as many $d_{ii}$’s as possible (in fact $J-1$ and $1$ are upperbounds on the powers of the $d_{ii}$’s that are contributed to the LD-terms by the first $F(J-1)$ rows of $W$ and the last $F$ rows of $W$, respectively; however, these bounds are attainable as will become clear). The rows in the upper part of $W$ do not include
different $d_{ii}$’s, while in $B_0$, $d_{ii}$ shows up in the $i$th column in every row (entry), for $i = 1, \ldots, F$.

The $d_{ii}$’s, that are contributed by the last $F$ rows to the unique LD-part necessarily originate from the first $F$ columns of $W$, that is, from $B_0$. As a consequence the $J - 1$ $d_{ii}$’s, for $i = 1, \ldots, F$, all have to come from $W_M$, which is a $F(J - 1) \times F(J - 1)$ matrix. This in turn uniquely determines the origin of the $d_{ii}$’s in the first $F(J - 1)$ rows as they all have to come from the diagonals of the $A_1$ matrices in $W_M$. Given the unique origin of the $d_{ii}$’s in the LD-part, that come from the first $F(J - 1)$ rows, we are still left with many possible combinations of the entries in $B_0$ (which correspond to $d_{ii}$’s), as all possible combinations result in $\prod_{k=1}^{F} d_{kk}$. Because we cannot pursue our search for a unique LD-term further at this point, we add up all LD-terms that involve $\prod_{k=1}^{F} d_{kk}$. The next step is to show that these terms almost always, that is for almost all $X$, do not counterbalance each other. As the factor in these LD-terms that comes from the first $J(F - 1)$ rows is always the same, namely $y^{F(J-1)} \prod_{k=1}^{F} d_{kk}^{-1}$, we focus on the factors in the LD-terms which originate from $B_0$. Define the following matrix $Y = [y_{ij}], y_{ij} = [\lambda_{i} S' (\gamma)^{-1}]_{ij} (X_{i})^\theta_{i}$. This matrix comprises the “coefficients” of the $d_{ii}$’s in the LD-terms that originate from $B_0$. The sum of all coefficients of the $\prod_{k=1}^{F} d_{kk}$ factors in the LD-terms which originate from $B_0$, equals $\det(Y)$. Thus a necessary and sufficient condition for the existence of a unique LD-part that is characterized by the factor $\prod_{k=1}^{F} d_{kk}$, i.e. a condition for this part of the LD not to equal zero, is that $\det(Y) \neq 0$ (apart from the trivial condition that given $X$ the $d_{ii}$’s do not equal zero). We will now show that $\det(Y) \neq 0$ for almost all $X$. Notice that $Y$ is a function of $X$ and $X$ alone.

For each matrix $X$ we have a pair $\zeta, S^{-1}$, the diagonal matrix with the unique eigenvalues of $X$ and a matrix of eigenvectors of $X$, respectively, and vice versa. $\zeta_{kk}$ alone completely determines $\lambda_{kk}^+$ for each $k = 1, \ldots, F$. Since $X$ can take any value, $\zeta$ and $S$ can take any value and vice versa. If we have chosen the values of the $\zeta_{kk}$’s and thereby have fixed the $\lambda_{kk}^+$’s, $Y$ can still take any arbitrary value by choosing $S$ appropriately. In fact given (almost all) $\zeta$, there is an injective relation between $Y$ and $S$. Thus given $\zeta = \zeta^*$, for almost all $X$ with the eigenvalues equal to $\zeta_{kk}^*$, $k = 1, \ldots, F$, or equivalently for almost all $S$, $\det(Y) \neq 0$. For almost all values of $\zeta_{kk}^*$, $k = 1, \ldots, F$, this argument applies. Furthermore by letting the $\zeta_{kk}$’s take every possible value in $\mathbb{R}$, $k = 1, \ldots, F$, and by choosing $S$ given $\zeta^*$ appropriately, all $X$’s are covered (can be generated). Thus for almost all $X$, the condition $\det(Y) \neq 0$ is satisfied. As a consequence for almost all $X$ there is a unique part in the Laplace development of $W$ that depends on $\prod_{k=1}^{F} d_{kk}$. (Recall that it is unique because it is the only LD-part that involves $\prod_{k=1}^{F} d_{kk}$). Because the $d_{ii}$’s can be chosen freely, we can choose their values, given $X$, in such a way that $\det(W) \neq 0$. So for almost all $X$, $\det(W)$ is not trivially zero and by Lemma 4, $M_{w,\gamma}$ has Lebesgue measure zero. Since the space of values of
entries of $X$ for which \( \det(W) \) is the null function has Lebesgue measure zero, \( M_{W,\gamma} \) has Lebesgue measure zero unconditionally.

## Appendix E. The determinant of $W$

We examine \( \det(W) \) when \( E = 0, G = \tilde{G} \), and \( H = 0 \), where \( \tilde{G} \) is an arbitrary PDS matrix.

In this case \( C = [(1 + \gamma)/\gamma]\tilde{G} \) and \( D = \tilde{G} \). It follows that \( B_0 = -[(1 + \gamma)/\gamma]\text{diag}(\tilde{g}_{11}, \ldots, \tilde{g}_{FF}) \) and \( B_i = 0 \) for \( i = 1, \ldots, J - 1 \). Partition \( W \) as

\[
W = \begin{bmatrix}
W_1 & W_2 \\
B_0 & W_2
\end{bmatrix}.
\]

By the theorem on the partitioned inverse, (the absolute value of) the determinant of \( W \) equals

\[
|W| = |W_m|B_0 - W_2W_m^{-1}W_1|.
\]

Since \( B_i = 0 \) for \( i = 1, \ldots, J - 1 \) we have \( W_2 = 0 \). Thus \( |B_0 - W_2W_m^{-1}W_1| = |B_0| = \prod_{i=1}^{F} B_{0,ii} \neq 0 \) because \( D = \tilde{G} \) is nonsingular. It is easily seen from the definitions that the product of the diagonal entries of \( W_m \) is a unique term in the Laplace development of \( \det(W_m) \). Thus in general \( |W_m| \neq 0 \) as well. It follows that \( \det(W) \neq 0 \) for almost all choices of \( \tilde{G} \) when \( E = 0, G = \tilde{G} \), and \( H = 0 \).

## Appendix F

\( d_{ff}, f = 1, \ldots, F \), show up in \( B_0 \) but not in \( B_i, i \neq 0 \).

The proof proceeds as follows:

The \( d_{ff} \)'s (could only) enter the formulae of the \( B_i \)'s through \( K_{j+i,1}, j = i, \ldots, J - 1, i = 0, \ldots, J - 1 \), which are sums of \( \tilde{A}_{k1} \) matrices. Now \( \gamma d_{ff} \) only appears as the \( (f,f) \)-entry of \( \tilde{A}_{\theta_i-\theta_{i+1},1}, f = 1, \ldots, F \). We will show that these matrices are only part of \( B_0 \). Abstracting from factors that only depend on \( X \), the \( B_i \) matrices equal

\[
\sum_{j=i}^{J-i} K_{j+i,1} = \sum_{j=i}^{J-i} \sum_{k=1}^{f(j)} \tilde{A}_{j+i-j,k}.
\]

Suppose that \( \tilde{A}_{\theta_i-\theta_{i+1},1} \) enters \( B_i \), then \( J + i - j = J + \theta_F - \theta_f, k = f \). It follows that \( j = \theta_F - \theta_f + j \) and \( f(j) \geq f \). From the latter we have \( j \leq \theta_F - \theta_f \), and therefore \( i \leq 0 \). Thus \( d_{ff}, f = 1, \ldots, F \), only show up in \( B_0 \). 

\[\square\]
References


