Federal Reserve Bank of Dallas
Globalization and Monetary Policy Institute
Working Paper No. 285
http://www.dallasfed.org/assets/documents/institute/wpapers/2016/0285.pdf

System Reduction and Finite-Order VAR Solution Methods for Linear Rational Expectations Models*

Enrique Martínez-García
Federal Reserve Bank of Dallas

September 2016

Abstract

This paper considers the solution of a large class of linear rational expectations (LRE) models and its characterization via finite-order VARs. The solution of the canonical LRE model can be cast in state-space form and solved for by the method of undetermined coefficients. In this paper I propose an approach that simplifies the systematic characterization of the solution into finite-order VAR form and checks existence and uniqueness based on the solution of a companion Sylvester equation. Solving LRE models with a finite-order VAR representation via the Sylvester equation is straightforward to implement, efficient, and can be handled easily with standard matrix algebra. An application to the workhorse New Keynesian model with accompanying Matlab codes is provided to illustrate the implementation of the procedure in practice.

JEL codes: C32, C62, C63, E37

* Enrique Martínez-García, Federal Reserve Bank of Dallas, Research Department, 2200 N. Pearl Street, Dallas, TX 75201. 214-922-5762. enrique.martinez-garcia@dal.frb.org. I would like to thank Nathan Balke, Jesús Fernández-Villaverde, Andrés Giraldo, Ayse Kabukcuoglu, María Teresa Martínez-García, Mike Plante, Michael Sposi, Mark Wynne, Carlos Zarazaga and seminar and conference participants for helpful suggestions and comments. All codes for this paper are publicly available in the following website: https://sites.google.com/site/emg07uw/. The codes can be downloaded using the following link: https://sites.google.com/site/emg07uw/econfiles/LRE_model_solution.zip?attredirects=0. I acknowledge the excellent research assistance provided by Valerie Grossman, and the help of Arthur Hinojosa. All remaining errors are mine alone. The views expressed in this paper are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 Introduction

Many rational expectations macro models can be cast as a linear system of expectational difference equations. The linearity of the system may be a feature of the model itself but, often, is attained by means of an approximation. In either case, the solution of linear or linearized rational expectations (LRE) models is an important part of modern macroeconomics. Blanchard and Kahn (1980) established the conditions for existence and uniqueness of a solution to the canonical LRE model (see also, among others, the related methodological contributions of King and Watson (1998), Uhlig (1999), and Klein (2000)). More recently, Fernández-Villaverde *et al.* (2007), Ravenna (2007), and Franchi and Paruolo (2015) have explored the conditions under which the canonical LRE model solution permits a VAR representation which facilitates the recovery of fundamental shocks.

The structural shocks of the LRE model cannot always be recovered from a VAR specification due to lack of fundamentalness, though. A model is said to be fundamental if there exists a (possibly infinite-order) VAR representation for the observables with respect to the structural shocks (Hansen and Sargent (1980)). When the number of shocks is equal to the number of observables, the fundamentalness property of the solution can be checked with the 'poor man's invertibility condition' of Fernández-Villaverde *et al.* (2007). Ravenna (2007) proposes a 'unimodularity condition' to check when the VAR representation of the LRE solution is of finite order. Franchi and Paruolo (2015), then, come to show that if the linear state-space form of the LRE solution is minimal, the 'poor man's invertibility condition' and the 'unimodularity condition' are necessary and sufficient to ensure both fundamentalness and the existence of a finite-order VAR representation for the solution to the LRE model.

In this paper, I complement this literature with a novel approach to characterize—for a large class of LRE models—the unique solution in finite-order VAR form via a companion quadratic matrix equation and a companion Sylvester equation. I also propose a set of straightforward conditions to check the properties—of existence and uniqueness—of a finite-order VAR solution. LRE models that include backward-looking and forward-looking features with one or more lags and leads can be reduced to the canonical form given by an expectational first-order difference equation without backward-looking terms. After system reduction is achieved after solving the companion quadratic matrix equation, the method of undetermined coefficients can be used to solve the canonical forward-looking part of the LRE model. Conditions under which a finite-order VAR representation of the LRE model solution can be obtained and a simple (yet efficient) algorithm to compute it can then be derived from a companion Sylvester equation.

The system reduction involves the solution of a quadratic matrix equation, and permits generalizing the approach (and the algorithms) I propose in this paper to cover a wider range of LRE models (Binder and Pesaran (1995, 1997)). One of the important contribution of the paper, though, is that the characterization of the finite-order VAR solution of the forward-looking part of the model arises from an associated matrix equation—the well-known Sylvester matrix equation. I propose a simple approach based on this companion Sylvester equation to check for and identify LRE solutions in finite-order VAR form as well as a related algorithm to compute such solutions.

The solution method proposed in this paper based on the companion Sylvester equation—augmented with a companion quadratic matrix equation to handle backward-looking components—equips economists with a straightforward way to efficiently solve a large class of multivariate LRE models with multiple leads and lags. In this paper, I illustrate this approach based on the companion Sylvester equation with the workhorse three-equation New Keynesian model—showing how the procedure is used to derive the finite-order VAR representation of the LRE solution, to establish its existence and uniqueness, and to make economically-relevant inferences about the New Keynesian transmission mechanism for (and the identification of) structural shocks.

The rest of the paper proceeds as follows: Section 2 describes the system reduction method to decouple the backward-looking and forward-looking parts and shows how to use the method of undetermined coefficients to characterize the linear state-space form solution of the forward-looking part of the canonical LRE model. Section 3 describes the mapping of the LRE model solution into a finite-order VAR via a companion Sylvester equation. This section also discusses the conditions under which a finite-order VAR solution can be attained from the companion Sylvester equation as well as the algorithms available to compute it. Section 4 applies the method to an economically-interesting example on the effects of monetary policy on inflation determination based on the workhorse three-equation New Keynesian model with which I also illustrate the computational efficacy of the approach. Section 5 then concludes.

The Appendix provides additional technical details on the system reduction approach used in this paper to isolate the forward-looking part of the LRE model—including a generalized eigenvalue problem algorithm to implement it—and also discusses related technical features of the linear state-space form of the solution and the general form of the LRE models for which the procedure can be utilized.

2 The Canonical LRE Model

Going from the structural relationships implied by the LRE model to a reduced-form solution requires explicit assumptions about the formation of expectations and the stochastic process of the exogenous driving variables. The structural relationships implied by the LRE model are always true according to theory, but the reduced-form solution of the model may depend on those assumptions. Here, I consider LRE models where expectations are fully rational and the exogenous driving variables are assumed to follow a VAR process. Often, a link arises between the reduced-form solution and a VAR representation for the endogenous variables under those assumptions that can explain why VARs appear to fit the data well, while providing researchers with a better tool for policy analysis than that of an unrestricted VAR.

Under rational expectations, agents understand the structure of the economy and formulate expectations optimally incorporating all available information.[1] Hence, the LRE model solution generally varies across alternative policy regimes and also as a result of variations in expectations about the future induced by policy changes. Ultimately, the assessment of an LRE model must be based upon its ability to capture the relevant aspects of the macro data. However, reduced-form econometric models fitted to historical data (unrestricted VARs) may describe the data well and still be of limited use for policy analysis without the structure imposed by theory—as stressed by Lucas (1976). In this paper, I explore the connection between the reduced-form solution of LRE models and VARs to help bridge the gap between theoretical and applied work towards a more unified approach (for testing the cross-equation restrictions imposed by theory and for policy analysis).

A large class of LRE models can be cast into a canonical system of expectational difference equations,

[1] The idea of rational expectations can be traced to the seminal work by Muth (1961). Lucas (1976) and Sargent (1980) were among the leading economists that rejected *ad hoc* assumptions on the formation of expectations and advocated the adoption of rational expectations that became prevalent in modern macroeconomics since the 80s.

featuring forward- and backward-looking dynamics. The expectational difference equations capture the structural relationships between a set of k endogenous variables $W_t = (w_{1t}, w_{2t}, ..., w_{kt})^T$ and k exogenous driving variables $X_t = (x_{1t}, x_{2t}, ..., x_{kt})^T$ as follows:

$$W_t = \Phi_1 W_{t-1} + \Phi_2 \mathbb{E}_t [W_{t+1}] + \Phi_3 X_t, \tag{1}$$

where Φ_1, Φ_2, and Φ_3 are conforming $k \times k$ square matrices.[2] The remaining p endogenous variables of the LRE model $\widetilde{W}_t = (\widetilde{w}_{1t}, \widetilde{w}_{2t}, ..., \widetilde{w}_{pt})^T$ can then be expressed as functions of W_t and X_t, possibly including lags and expectations of their leads too.

I complete the specification of the canonical form of the LRE model given by (1) with a standard VAR(1) specification for the vector of k driving variables X_t of the following form:

$$X_t = AX_{t-1} + B\epsilon_t, \tag{2}$$

where A is a $k \times k$ matrix that has all its eigenvalues inside the unit circle, B is a $k \times k$ matrix that describes the matrix of variances and covariances, and ϵ_t is the corresponding vector of innovations of dimension k.

The conventional method to characterize a reduced-form solution for a forward-looking LRE model is laid out in Blanchard and Kahn (1980), which also provide conditions to check the existence and uniqueness of the solution. Blanchard and Kahn's (1980) method was further refined by Broze *et al.* (1985, 1990), King and Watson (1998), Uhlig (1999), and Klein (2000) among others, to obtain solutions in more general settings. Other popular solution methods applied to LRE models include the method of undetermined coefficients of Christiano (2002) and the method of rational expectational errors of Sims (2002) (see also Lubik and Schorfheide (2003)).

2.1 Decoupling Backward- and Forward-Looking Terms

The quadratic determinantal equation (QDE) method of Binder and Pesaran (1995, 1997) builds on Broze *et al.* (1985, 1990), but deals explicitly with the simultaneous dependence of W_t on its past and on its future expected values given by the canonical form in (1). I adopt a key aspect of the QDE method implementing a simple transformation of (1) that achieves a system reduction excluding all backward-looking terms from the expectational difference system and then I work out the LRE model solution by parts.

For a given $k \times k$ matrix Θ, the proposed transformation of the endogenous variables given by $W_t = Z_t + \Theta W_{t-1}$ implies that the expectational difference system in (1) can be rewritten as

$$\begin{aligned} Z_t + \Theta W_{t-1} &= \Phi_1 W_{t-1} + \Phi_2 \mathbb{E}_t [Z_{t+1} + \Theta W_t] + \Phi_3 X_t \\ &= \Phi_1 W_{t-1} + \Phi_2 [\mathbb{E}_t (Z_{t+1}) + \Theta (Z_t + \Theta W_{t-1})] + \Phi_3 X_t, \end{aligned} \tag{3}$$

which then becomes

$$(I_k - \Phi_2 \Theta) Z_t = \Phi_2 \mathbb{E}_t (Z_{t+1}) + \left(\Phi_2 \Theta^2 - \Theta + \Phi_1\right) W_{t-1} + \Phi_3 X_t. \tag{4}$$

[2]Here, I consider the case where the number of endogenous and exogenous variables in (1) − (2) is the same—hence, if the solution can be represented in VAR form, it can be exploited to recover the fundamental economic shocks included in the model. Notice that the canonical form in (1) − (2) can be generalized to capture LRE models including more than one lead and lag in the specification, as explained in the Appendix.

From here, this condition follows:

Condition 1 *A system reduction that excludes the backward-looking terms in* (1) *can be attained by choosing a $k \times k$ matrix Θ to satisfy that*

$$P(\Theta) = \Phi_2 \Theta^2 - \Theta + \Phi_1 = 0_k, \tag{5}$$

where 0_k is a square $k \times k$ matrix of zeroes.

This transformation uncouples the solution of (1) into a backward-looking part, $W_{tb} = \Theta W_{t-1}$, and a forward-looking part, $W_{tf} = Z_t$, such that $W_t = W_{tb} + W_{tf}$. Hence, to solve the canonical LRE model in (1), one first needs to determine the matrix Θ solving the quadratic matrix equation in (5) to characterize the backward-looking part of the solution and reduce the canonical system to its purely forward-looking expectational difference part. Binder and Pesaran (1995, 1997) establish the necessary and sufficient conditions under which real-valued solutions for Θ satisfying (5) exist and provide a straightforward iterative algorithm to compute its stable solution. A discussion of an alternative algorithm to characterize the solution Θ based on the generalized eigenvalue problem can be found in the Appendix.

After decoupling, the vector of k transformed variables $Z_t = W_t - \Theta W_{t-1}$ follows a canonical first-order forward-looking expectational difference system of the following form:

$$\Gamma_0 Z_t = \Gamma_1 \mathbb{E}_t [Z_{t+1}] + \Gamma_2 X_t, \tag{6}$$

where $\Gamma_0 \equiv (I_k - \Phi_2 \Theta)$, $\Gamma_1 \equiv \Phi_2$, and $\Gamma_2 \equiv \Phi_3$ are conforming $k \times k$ matrices. Whenever Γ_0 is nonsingular, the system of structural relationships implied by (6) can be rewritten as

$$Z_t = F \mathbb{E}_t [Z_{t+1}] + G X_t, \tag{7}$$

where $F \equiv (\Gamma_0)^{-1} \Gamma_1$ and $G \equiv (\Gamma_0)^{-1} \Gamma_2$.

The invertibility of Γ_0 required to go from (6) to (7) depends on the choice of matrix Θ. Binder and Pesaran (Proposition 2, 1997) discuss conditions under which the solution can be characterized analytically and then provide sufficient conditions under which $(I_k - \Phi_2 \Theta)$ would be nonsingular. Whenever the Binder and Pesaran (1997) conditions are satisfied, the matrix Γ_0 can be shown to be nonsingular and invertible. Those conditions are only sufficient (not necessary), but I find that most well-specified economic models indeed produce a matrix Γ_0 that is nonsingular. Hence, I take the specification in (7) to constitute the relevant benchmark to describe the forward-looking part of the LRE model solution in the remainder of this paper.

2.2 The Method of Undetermined Coefficients for Solving the LRE Model

Assuming that the Blanchard and Kahn (1980) conditions for existence and uniqueness are satisfied, the solution to the forward-looking part of the LRE model can be written in linear state-space form as follows:

$$X_t = A X_{t-1} + B \epsilon_t, \tag{8}$$

$$Z_t = C X_{t-1} + D \epsilon_t, \tag{9}$$

where A, B, C, and D are conforming $k \times k$ matrices. Equation (8) simply re-states the assumption made in (2) about the dynamics of the exogenous driving variables, while (9) indicates that current innovations and lagged exogenous driving variables are mapped into the current endogenous variables in the reduced-form solution of the LRE model.

In order to pin down this reduced-form solution, I need to relate the unknown reduced-form coefficients C and D to the composite matrices that describe the structural relationships of the LRE model (F, G) and those that describe the shock process of the driving variables (A, B). Using the method of undetermined coefficients, such a solution can be fully characterized via the solution of a Sylvester equation and, under additional constraints, it can be shown to have a finite-order VAR representation.

Step 1. Using (9) shifted one period ahead to replace Z_{t+1} in the purely forward-looking system of equations given in (7) implies that:

$$\begin{aligned} Z_t &= [FC + G] X_t && (10) \\ &= [FC + G] AX_{t-1} + [FC + G] B\epsilon_t, && (11) \end{aligned}$$

where the second equality from from replacing X_t out using (8). The general form of the forward-looking LRE model solution given in (9) can be matched with (11) to link the unknown reduced-form coefficients C and D to the composite matrices that describe the LRE model (F, G, A, B). Hence, by the method of undetermined coefficients, it follows that the conforming square matrices C and D that characterize (9) in the solution must satisfy the following conditions:

$$\begin{aligned} C &= [FC + G] A, && (12) \\ D &= CA^{-1}B. && (13) \end{aligned}$$

The eigenvalues of matrix A must be inside the unit circle by construction for the VAR(1) process of the driving variables to be stationary. I also assume that zero is not an eigenvalue of matrix A as that ensures the inverse matrix A^{-1} in (13) exists and is well-defined according to the Invertible Matrix Theorem.

As a result, the existence and uniqueness of a solution to C that satisfies (12) also pins down D through (13) ensuring the existence and uniqueness of the full solution to the forward-looking part of the LRE model given by the state-space form in (8) $-$ (9). Hence, assuming Γ_0 is invertible as indicated before, solving the companion Sylvester equation given by (12) to obtain C turns out to be enough to characterize the solution to the forward-looking part of the LRE model in (8) $-$ (9).

Step 2. Using (8) and the invertibility of A, I can write X_{t-1} as $X_{t-1} = A^{-1} (X_t - B\epsilon_t)$. Replacing this expression in (9), I infer that:

$$Z_t = CA^{-1}X_t + \left[D - CA^{-1}B\right] \epsilon_t. \tag{14}$$

Then, it follows from the condition (13) that characterizes the matrix D using the method of undetermined coefficients that the term related to the vector of innovations ϵ_t must drop out from (14). As a result, the solution of the forward-looking part of the LRE model implies a straightforward linear mapping from the vector of exogenous driving variables X_t to the vector of endogenous variables Z_t where

$$Z_t = CA^{-1}X_t, \tag{15}$$

if a matrix C exists that solves the Sylvester equation implied by condition (13).

Whenever a unique C exists which can also be shown to be invertible, equation (15) implies that $X_t = AC^{-1}Z_t$ given that A is invertible by construction. Shifting this expression one period back and replacing it in (9), I obtain the following VAR(1) specification that characterizes the solution of the forward-looking part of the LRE model:

$$Z_t = CAC^{-1}Z_{t-1} + CA^{-1}B\epsilon_t, \tag{16}$$

where I replaced D using condition (13). The existence and uniqueness of the matrix D follows naturally under condition (13) from the existence and uniqueness of a matrix C that solves the Sylvester equation given by condition (12). However, I also find that the solution of the forward-looking part of the LRE model has a finite-order VAR representation given by (16) whenever a unique matrix C solving the Sylvester equation exists which can also be shown to be invertible. Therefore, the characterization of the finite-order VAR solution in (16) depends on the invertibility of Γ_0 as indicated before, but also on the existence, uniqueness, and invertibility of the solution C arising from condition (12).

Rewriting condition (12), the characterization, existence, and uniqueness of a finite-order VAR solution can be summarized as follows:

Lemma 1 *If Γ_0 is invertible, a VAR(1) representation of the solution to the first-order expectational difference system of equations in (7) can be obtained by solving a companion Sylvester equation in C*

$$FCA - C = H, \tag{17}$$

where

$$F \equiv (\Gamma_0)^{-1}\Gamma_1,\ H \equiv -GA = -(\Gamma_0)^{-1}\Gamma_2 A. \tag{18}$$

If a unique matrix C exists and is invertible, the VAR(1) representation of the solution is given by (16).

The proof of this lemma follows directly from the derivation of conditions (12) – (13) by the method of undetermined coefficients, as discussed above.

Step 3. Then, the solution of the full LRE model can be obtained by combining its backward- and forward-looking parts—i.e., $W_t = W_{tb} + W_{tf}$ where the backward-looking part, $W_{tb} = \Theta W_{t-1}$, follows from the solution to the quadratic matrix equation in (5) and the forward-looking part, $W_{tf} = Z_t = W_t - \Theta W_{t-1}$, is defined by the state-space solution in (8) – (9) given conditions (12) – (13) whenever a solution C to the companion Sylvester equation exists. Under the additional condition that C be invertible (stated in Lemma 1), the implication is that the solution of the full model must follow a VAR(2) process of the following form

Corollary 1 *If a unique matrix C exists that solves the companion Sylvester equation of Lemma 1 and this matrix is also invertible, the VAR(2) representation of the canonical LRE model solution for the vector of endogenous variables W_t is given by*

$$W_t = \Psi_1 W_{t-1} + \Psi_2 W_{t-2} + \Psi_3\epsilon_t, \tag{19}$$

where the corresponding coefficient matrices are $\Psi_1 \equiv \left(\Theta + CAC^{-1}\right)$, $\Psi_2 \equiv -CAC^{-1}\Theta$, *and* $\Psi_3 \equiv CA^{-1}B$.

The derivation of this corollary follows naturally from the definition of the forward-looking part of the solution Z_t as $Z_t = W_t - \Theta W_{t-1}$ and the derivation of a VAR(1) representation for Z_t implied under the terms of Lemma 1. Notice that the matrix Ψ_3 is not necessarily going to be positive-semidefinite and symmetric, so it does not have the standard interpretation of a variance-covariance matrix unlike the matrix B.[3] However, what follows from the definition of Ψ_3 in Corollary 1 is that Ψ_3 is a linear transformation of the variance-covariance matrix of the shock process B where the mapping is determined by CA^{-1} (which depends on the solution C of the companion Sylvester equation).

3 The Solution to the Companion Sylvester Equation

If a solution C exists for the companion Sylvester equation given by (17), then the solution of the forward-looking part of the LRE model given in state-space form by (8) – (9) exists. According to Lemma 1, such a solution can be represented with a finite-order VAR whenever C is also shown to be invertible. Hence, the main methodological contribution of the paper is to solve a large class of LRE models that fit into the canonical form given by (1) – (2) via a companion Sylvester equation. The corollary of the approach I propose is that under some additional conditions on the form of the solution to the companion Sylvester equation, the LRE model solution admits a finite-order VAR representation. Hence, in this section I aim to provide some additional discussion on characterizing the solution to the Sylvester equation in (17) and provide an overview of the efficient algorithms to compute it numerically.

3.1 Characterization of the Sylvester Equation Solution

Equation (17) proposes a companion Sylvester equation—i.e., $FCA - C = H$ with $F, A, H \in \mathbb{R}^{k\times k}$ given and $C \in \mathbb{R}^{k\times k}$ to be determined—as an alternative to characterize the solution of the forward-looking part of the LRE model. The Sylvester equation is one of the better known matrix equations in stability and control theory and its applications.[4] Using the Kronecker (tensor) product notation and the properties of the vectorization operator *vec*, I can re-write Sylvester's equation in its standard form as a linear system of equations:

$$\mathcal{A} vec\left(C\right) = vec\left(H\right), \tag{20}$$

$$\mathcal{A} := \left[\left(A^T \otimes F\right) - I_{k^2}\right], \tag{21}$$

[3] To test whether Ψ_3 has the properties of a variance-covariance matrix, I can use the *chol* function in Matlab. If *chol* returns a second argument that is zero from $[R, p] = chol\left(\Psi_3\right)$, then the matrix is confirmed to be symmetric and—in this case—also positive definite.

[4] The Sylvester matrix equation is extensively discussed in Lancaster and Tismenetsky (1985, Chapter 12). Additional useful references on the characterization of the solution to the Sylvester equation include Horn and Johnson (1991) and Jiang and Wei (2003).

where $\otimes$ denotes the Kronecker product.[5] In this way, the Sylvester equation is represented by a linear system of dimension $k^2 \times k^2$ conformed by k^2 equations in k^2 unknown variables (where the unknowns correspond to the elements of the matrix C).

Having transformed the Sylvester equation into the linear system given by $(20)-(21)$, well-known matrix algebra results suffice to determine the following criteria for the existence and uniqueness of a solution C to the companion Sylvester equation:

Proposition 1 *Let $F, A, H \in \mathbb{R}^{k\times k}$. Then, it follows that:*

(a) (Existence) The Sylvester equation in (17) *has at least one solution $C \in \mathbb{R}^{k\times k}$ if and only if $rank\,[\mathcal{A}\ vec\,(H)] = rank\,[\mathcal{A}]$.*

(b) (Uniqueness) The Sylvester equation in (17) *has a unique solution $C \in \mathbb{R}^{k\times k}$ if and only if $rank\,[\mathcal{A}] = k^2$; that is, the solution is unique if and only if $\mathcal{A}$ has full rank. Then, $\mathcal{A}$ is nonsingular and invertible implying that the unique solution to the Sylvester equation can be recovered as $vec\,(C) = \mathcal{A}^{-1} vec\,(H)$.*

Proof. (a) Trivially it follows that $rank\,[\mathcal{A}\ vec\,(H)] \geq rank\,[\mathcal{A}]$. If there is a solution $vec\,(C) = [\bar{c}_1\ \bar{c}_2\ ...\ \bar{c}_{k^2}]^T$ for the linear system given by $(20)-(21)$, then $\sum_{i=1}^{k^2} \mathcal{A}_{*i}\bar{c}_i = vec\,(H)$ where $\mathcal{A}_{*1}, \mathcal{A}_{*2}, ..., \mathcal{A}_{*k^2}$ denote the corresponding columns of the matrix $\mathcal{A}$. Hence, $vec\,(H)$ is a linear combination of the columns of $\mathcal{A}$ and, as a result, the rank of the augmented matrix $[\mathcal{A}\ vec\,(H)]$ cannot be different than the rank of $\mathcal{A}$—because for $rank\,[\mathcal{A}\ vec\,(H)] > rank\,[\mathcal{A}]$ to be true, $vec\,(H)$ needs to be linearly independent from the columns of $\mathcal{A}$ and that contradicts $vec\,(C)$ being a solution.

(b) If the square matrix $\mathcal{A}$ has full rank—and, therefore, is nonsingular and invertible—the linear system in $(20)-(21)$ has a unique solution given by $vec\,(C) = \mathcal{A}^{-1} vec\,(H)$. The converse statement follows naturally as well. If the linear system has a unique solution, then $\mathcal{A}$ must have full rank and be nonsingular. Otherwise, at least one column in $\mathcal{A}$ is not linearly independent from the rest of the columns and can be written as a linear combination of them. Hence, for any given solution $vec\,(C)$ defined over the linearly independent columns, another different solution exists including non-trivially the linearly dependent columns of $\mathcal{A}$ as well. The existence of more than one solution then contradicts the uniqueness assumption. ■

Proposition 1 characterizes the solution C to the Sylvester equation in (17) and, by extension, determines conditions for the existence and uniqueness of the solution to the forward-looking part of the LRE model. The two rank conditions stated in this proposition depend solely on the properties of the matrices $F, A, H \in \mathbb{R}^{k\times k}$

[5] For a given matrix $X \in \mathbb{R}^{k\times k}$, express $X = [X_{*1}\ X_{*2}\ ...\ X_{*k}]$ where $X_{*j} \in \mathbb{R}^k$, $j = 1, 2, ..., k$. Then, the vectorization associated to matrix X defines the following vector-valued function

$$\begin{bmatrix} X_{*1} \\ X_{*2} \\ ... \\ X_{*k} \end{bmatrix} \in \mathbb{R}^{k^2}$$

which is denoted $vec\,(X)$. The vectorization operation is linear, i.e. $vec\,(\alpha X + \beta Y) = \alpha vec\,(X) + \beta vec\,(Y)$ for any $X, Y \in \mathbb{R}^{k\times k}$ and $\alpha, \beta \in \mathbb{R}^k$. If $X = [x_{ij}]_{i,j=1}^k \in \mathbb{R}^{k\times k}$ and $Y \in \mathbb{R}^{k\times k}$, then the Kronecker (tensor) product of X and Y, written $X \otimes Y$, is defined to be the partitioned matrix

$$X \otimes Y = [x_{ij}Y] = \begin{bmatrix} x_{11}Y & x_{12}Y & ... & x_{1k}Y \\ x_{21}Y & x_{22}Y & ... & x_{2k}Y \\ \vdots & \vdots & & \vdots \\ x_{k1}Y & x_{k2}Y & & x_{kk}Y \end{bmatrix} \in \mathbb{R}^{k^2\times k^2}.$$

Proposition 4 in Chapter 12.2 of Lancaster and Tismenetsky (1985) shows that the vectorization operation is closely related to the Kronecker product as follows: If $X, Y, Z \in \mathbb{R}^{k\times k}$, then $vec\,(XYZ) = \left(Z^T \otimes X\right) vec\,(Y)$.

that describe the structural relationships of the LRE model. The uniqueness rank condition implies a solution of the form $vec\,(C) = \mathcal{A}^{-1} vec\,(H)$ for the Sylvester equation and, naturally, that proves existence as well. If the uniqueness rank condition is violated, the existence rank condition determines whether there is no solution to the companion Sylvester equation—if $rank\,[\mathcal{A}\ vec\,(H)] \neq rank\,[\mathcal{A}]$—or whether multiple solutions exist—if $rank\,[\mathcal{A}\ vec\,(H)] = rank\,[\mathcal{A}] < k^2$. In the latter case, it can be shown that the number of linearly independent solutions is determined by the dimension of the kernel of $\mathcal{A}$.

The characterization of the linearly independent solutions of the companion Sylvester equation whenever a solution exists but is not unique can be found elsewhere in Theorem 12.5.1, Theorem 12.5.2 and Corollary 12.5.1 of Lancaster and Tismenetsky (1985). Focusing on the case of interest for this paper where a solution for the Sylvester equation in (17) exists and is unique, the full rank condition on $\mathcal{A}$ can be expressed in terms of the eigenvalues of F and A as follows:

Proposition 2 *Let $F, A \in \mathbb{R}^{k \times k}$ be given. Let $\lambda_1, ..., \lambda_k$ be the eigenvalues of F and $\mu_1, ..., \mu_k$ the eigenvalues of A. Then, for any matrix $H \in \mathbb{R}^{k \times k}$, it follows that the companion Sylvester equation in* (17) *has a unique solution if and only if $\lambda_i \mu_j \neq 1$ for all $i, j = 1, ..., k$. In other words, the Sylvester equation has a unique solution $C \in \mathbb{R}^{k \times k}$ if and only if the matrices F and A^{-1} have no eigenvalues in common.*

Proof. The eigenvalues of A are the same as those of its transpose A^T. Given that and the properties of the Kronecker product, the eigenvalues of $\left(A^T \otimes F\right)$ are the k^2 numbers defined by the product between the eigenvalues of F and A, i.e., $\lambda_i \mu_j$ for all $i, j = 1, ..., k$. Then, the eigenvalues of $\mathcal{A} := \left[\left(A^T \otimes F\right) - I_{k^2}\right]$ are simply the k^2 numbers $\lambda_i \mu_j - 1$ for all $i, j = 1, ..., k$. By Proposition 1, the existence and uniqueness of a solution to the Sylvester equation requires $\mathcal{A}$ to be nonsingular (and have full rank). The matrix $\mathcal{A}$ is nonsingular if and only if all its eigenvalues are nonzero, i.e. if and only if $\lambda_i \mu_j - 1 \neq 0$ for all $i, j = 1, ..., k$. Re-arranging the nonzero conditions on the eigenvalues, it follows that $\lambda_i \neq \frac{1}{\mu_j}$ for all $i, j = 1, ..., k$. Given that the eigenvalues of A^{-1} are $\frac{1}{\mu_1}, ..., \frac{1}{\mu_k}$ while those of F are $\lambda_1, ..., \lambda_k$, a unique solution is said to exist if and only if indeed the matrices F and A^{-1} have no eigenvalues in common. ■

According to Proposition 2, the companion Sylvester equation for the forward-looking part of the LRE model has a unique solution C for each matrix H if and only if F and A^{-1} have no eigenvalues in common. The Sylvester operator $S : \mathbb{R}^{k \times k} \to \mathbb{R}^{k \times k}$ can be defined as follows:

$$S\,(C) = FCA - C, \tag{22}$$

where $F, A \in \mathbb{R}^{k \times k}$ are given and $C \in \mathbb{R}^{k \times k}$ is the solution to be identified. Then, the Sylvester equation can simply be written as $S\,(C) = H$ for any given $H \in \mathbb{R}^{k \times k}$.

The k^2 eigenvalues of the Sylvester operator $S\,(C)$ are $\lambda_i \mu_j - 1$, for all $i, j = 1, ..., k$, where $\lambda_1, ..., \lambda_k$ are the eigenvalues of F, and $\mu_1, ..., \mu_k$ are the eigenvalues of A. Let v_i be the right eigenvector of F associated with the eigenvalue λ_i such that $F v_i = \lambda_i v_i$ for all $i = 1, ..., k$. Let w_j be the left eigenvector of A associated with the eigenvalue μ_j such that $w_j^T A = \mu_j w_j^T$ for all $j = 1, ..., k$. Then, for any $i, j = 1, ..., k$, $C = v_i w_j^T$ is an eigenvector matrix of the Sylvester operator $S\,(C)$ associated with its eigenvalue $\lambda_i \mu_j - 1$. It follows from here that the Sylvester operator can be expressed as:

$$
\begin{aligned}
S(C) = FCA - C &= F\left(v_i w_j^T\right) A - v_i w_j^T \\
&= (F v_i)\left(w_j^T A\right) - v_i w_j^T \\
&= (\lambda_i v_i)\left(\mu_j w_j^T\right) - v_i w_j^T \\
&= (\lambda_i \mu_j - 1)\, v_i w_j^T \\
&= (\lambda_i \mu_j - 1)\, C.
\end{aligned} \tag{23}
$$

Hence, it can be shown that the Sylvester operator $S(C)$ must be nonsingular whenever $\lambda_i \mu_j \neq 1$ for all $i, j = 1, ..., k$; that is, $S(C)$ is nonsingular if and only if the solution to the Sylvester equation exists and is unique.

The eigenvalues of A, $\mu_1, ..., \mu_k$, are all inside the unit circle by construction to ensure the stochastic process for the driving variables is stationary and none of those eigenvalues is 0 so A is invertible. However, the conditions for existence and uniqueness of the Sylvester equation solution stated in Proposition 2 do not require the eigenvalues of F, $\lambda_1, ..., \lambda_k$, to be inside the unit circle. Imposing additional restrictions on the eigenvalues of F, an explicit form for the solution C of the Sylvester equation in (17) can be obtained as follows:

Proposition 3 *Let $F, A \in \mathbb{R}^{k\times k}$ where $\lambda_1, ..., \lambda_k$ are the eigenvalues of F and $\mu_1, ..., \mu_k$ are the eigenvalues of A. Then, for any matrix $H \in \mathbb{R}^{k\times k}$, it follows that the companion Sylvester equation in* (17) *has a unique solution whenever $\lambda_i \mu_j < 1$ for all $i, j = 1, ..., k$ and this solution is given by*

$$
C = -\sum\nolimits_{s=0}^{\infty} F^s H A^s. \tag{24}
$$

Proof. Let me define the following recursion: $FC_{r-1}A - C_r = H$ for iterations $r = 1, 2, 3, ...$ with the initial condition $C_0 = 0_k$ and 0_k is a $k \times k$ matrix of zeros. If this recursion converges as r goes to infinity, then by construction the limit characterizes the solution of the Sylvester equation $\lim\limits_{r\to\infty} C_r = -\sum\nolimits_{s=0}^{\infty} F^s H A^s = C$. The convergence condition is equivalent to $\lim\limits_{s\to\infty} F^s H A^s = 0$. It naturally follows that any eigenvalue of $F^s H A^s$ must be proportional to the $s-$power of the product between the eigenvalues of F and A, i.e., $(\lambda_i \mu_i)^s$ for any $i = 1, ..., k$. Hence, if all cross products between the eigenvalues of F and A are strictly less than one, the corresponding eigenvalues for $F^s H A^s$ must go to zero in the limit as $s \to \infty$ and this suffices to show that the recursion must converge as stated (Lancaster and Tismenetsky (1985, Chapter 12.3)). ■

Proposition 3 implies that whenever the product of the spectral radii of matrices A and F is strictly less than one, a unique solution exists that takes the special form of an infinite sum. Furthermore, this special case permits the straightforward computation of the solution to the companion Sylvester equation via a recursion on a convergent sequence as suggested by the proof of the proposition. I explore later on a related numerical algorithm and other alternative efficient algorithms to compute the unique solution of the Sylvester equation in (17). Often a numerical solution rather than one in explicit form is all that is needed to recover the forward-looking part of the LRE model.

However, even when a solution C to the companion Sylvester equation in (17) exists and is unique, characterizing the solution of the forward-looking part of the LRE model with a finite-order VAR form requires C also to be nonsingular. If C can be shown to be nonsingular, its inverse C^{-1} can be computed in order to obtain the finite-order VAR representation of the LRE model given by Lemma 1 and Corollary 1.

Proving the existence of C and its uniqueness does not suffice to ensure the solution is also invertible. For some matrices $H \in \mathbb{R}^{k \times k}$, the unique solution C that exists can be singular. For example, for any matrix H the solution C of the Sylvester equation exists and is unique—given by $vec\,(C) = \mathcal{A}^{-1} vec\,(H)$—if and only if $\mathcal{A} := \left[\left(A^T \otimes F\right) - I_{k^2}\right]$ is invertible (Proposition 1). Hence, it immediately follows that the unique solution is $C = 0_k$ whenever $H = 0_k$ and this solution is clearly singular. The following condition must hold in order to ensure that C is invertible:

Condition 2 *Assume the conditions stated in Proposition 1 and Proposition 2 on $F, A \in \mathbb{R}^{k \times k}$ are satisfied and a unique solution C exists for the companion Sylvester equation in* (17)*. Then, for a given matrix $H \in \mathbb{R}^{k \times k}$, the solution C is said to be nonsingular and invertible if and only if C has full rank; that is, if and only if $rank\,(C) = k$.*

Condition 2 is straightforward and follows directly from the terms of the Invertible Matrix Theorem. Related to this rank condition, there is evidence that ties the properties of the matrices $F, A, H \in \mathbb{R}^{k \times k}$ that underpin the Sylvester equation to a rank minimization condition. By Roth's Removal Theorem (Lancaster and Tismenetsky (1985, Chapter 12.5)), the Sylvester equation in (17) has a solution $C \in \mathbb{R}^{k \times k}$ if and only if there exists a nonsingular matrix $P \in \mathbb{R}^{k \times k}$ such that:

$$P \begin{pmatrix} F & HA^{-1} \\ 0 & A^{-1} \end{pmatrix} = \begin{pmatrix} F & 0 \\ 0 & A^{-1} \end{pmatrix} P. \tag{25}$$

Then, the following rank identity has been noted elsewhere (Lin and Wimmer (2011)):

$$\begin{aligned} &\min \left\{ rank\left(FC - CA^{-1} - HA^{-1}\right) \mid C \in \mathbb{R}^{k \times k} \right\} \\ &= \min \left\{ rank\left[P \begin{pmatrix} F & HA^{-1} \\ 0 & A^{-1} \end{pmatrix} - \begin{pmatrix} F & 0 \\ 0 & A^{-1} \end{pmatrix} P \right] \mid \forall P \in \mathbb{R}^{k \times k} \text{ s.t. } rank\,(P) = k \right\}. \end{aligned} \tag{26}$$

These and related results in the mathematical literature provide ways to connect the properties of the matrices $F, A, H \in \mathbb{R}^{k \times k}$ to the rank condition on the invertibility of the unique solution C to the Sylvester equation (Condition 2).

However, I prefer at this point to postpone for future research the full exploration of those connections. The reason for this is purely practical. If a solution exists and is unique according to the conditions stated in Proposition 1 and Proposition 2, then computing the matrix C is all that is needed to describe the solution to the forward-looking part of the LRE model in state-space form as given by equations (8) − (9) under conditions (12) − (13). Then, it is straightforward to check the rank condition for the invertibility of the computed matrix C (Condition 2). If that rank condition is violated, then one can conclude that the solution to the forward-looking part of the LRE model in state-space form cannot be transformed into a finite-order VAR form. If that rank condition is satisfied, it follows that a finite-order VAR representation exists given by equation (16) as indicated in Lemma 1. In this case, the VAR specification should permit the recovery of the true economic shocks underlying the model from the observed data.

3.2 Numerical Methods and Algorithms

It follows from Proposition 1 that a unique solution C of the companion Sylvester equation in (17) exists if and only if the $k^2 \times k^2$ square matrix $\mathcal{A}$ is invertible. If so, the unique solution takes the form of a linear system with k^2 equations and k^2 unknown variables given by $vec(C) = \mathcal{A}^{-1} vec(H)$ which can be solved in $O(k^6)$ operations. Although obtaining the solution in this way is straightforward, there are algorithms and methods that can improve efficiency in the numerical computation of C. I base my approach on three steps: First, a linear transformation of the companion Sylvester equation based on Schur's triangulation; second, the transformed equation is solved (in this case, Schur's triangulation permits a recursive implementation); and, finally, the inverse transformation of Schur's triangulation is applied to obtain the solution to the original form of the companion Sylvester equation.

Step 1. The first step of the approach is to implement the generalized Schur triangulation to re-write the companion Sylvester equation in (17). I find the real Schur decompositions $F = UKU^T$ and $A = VQV^T$ where $U, V \in \mathbb{R}^{k \times k}$ are unitary matrices of dimension k such that $UU^T = U^T U = I_k$ and $VV^T = V^T V = I_k$. The matrices $K, Q \in \mathbb{R}^{k \times k}$, referred to as the Schur forms corresponding to $F, A \in \mathbb{R}^{k \times k}$ respectively, are both upper triangular.[6] The eigenvalues of F and A are then the diagonal entries of the (upper) triangular matrices K and Q, respectively. Hence, I can re-write the companion Sylvester equation—i.e., $FCA - C = H$ with $F, A, H \in \mathbb{R}^{k \times k}$ given and $C \in \mathbb{R}^{k \times k}$ to be determined—as follows

$$KYQ - Y = R, \tag{27}$$

where $Y = U^T CV$ and $R = U^T HV$.

Step 2. The second step of the approach is to solve the transformed Sylvester equation in (17). The transformed Sylvester equation can be vectorized to obtain,

$$\overline{\mathcal{A}} := \left[\left(Q^T \otimes K \right) - I_{k^2} \right], \tag{28}$$

$$\overline{\mathcal{A}} vec(Y) = vec(R). \tag{29}$$

Then, this can be solved directly by calculating the inverse of $\overline{\mathcal{A}}$ and using standard matrix algebra to solve the linear system $vec(Y) = \overline{\mathcal{A}}^{-1} vec(R)$ for Y.

Step 3. The last step of the approach is to recover the solution C to the Sylvester equation. For that, I simply un-do Schur's triangulation as follows $C = UYV^T$.

Recursive Implementation of the Proposed Solution Method. Although the three-step approach laid out before works in general, the solution to the transformed system in (28)−(29) can be further optimized under additional assumptions on the matrix F and, particularly, on the matrix A.

(a) The solution to the transformed Sylvester equation given by (28) − (29) permits a more efficient recursive implementation, if A is diagonalizable. The Diagonalization Theorem indicates that the $k \times k$ matrix A is diagonalizable if and only if A has k linearly independent eigenvectors. Then, if A is diagonalizable, then the matrix S of its eigenvectors is invertible and $S^{-1} AS = M = diag(\mu_1, ..., \mu_k)$ is the diagonal matrix of its eigenvalues. A sufficient (but not necessary) condition for A to be diagonalizable is that all its k

[6] Since K is similar to F, they both have the same eigenvalues; similarly, since Q is similar to A, their eigenvalues are the same.

eigenvalues be distinct.[7] By the Principal Axis Theorem, it follows that if A is a real matrix (i.e., all entries of A are real numbers) and a symmetric matrix (i.e., $A^T = A$), then A is diagonalizable as well. Hence, the additional assumption that A be a diagonalizable matrix does not appear to be too restrictive for most practical applications—given that the stochastic process for the driving variables is often assumed symmetric—i.e., A is a real symmetric square matrix—and, even when the symmetry assumption is relaxed, generally the eigenvalues appear as distinct.

Assuming from now on that the matrix A is diagonalizable, I can re-write the companion Sylvester equation—i.e., $FCA - C = H$ with $F, A, H \in \mathbb{R}^{k\times k}$ given and $C \in \mathbb{R}^{k\times k}$ to be determined—with the Schur triangulation of F as before but using the diagonalization of A to obtain:

$$K\widehat{Y}M - \widehat{Y} = \widehat{R}, \tag{30}$$

where $\widehat{Y} = U^T CS$ and $\widehat{R} = U^T HS$. Then, the transformed Sylvester equation can be vectorized as:

$$\widehat{\mathcal{A}} := \left[\left(M^T \otimes K\right) - I_{k^2}\right], \tag{31}$$

$$\widehat{\mathcal{A}} vec\left(\widehat{Y}\right) = vec\left(\widehat{R}\right). \tag{32}$$

The matrices M^T and I_{k^2} are diagonal, while K is an upper triangular matrix. As a result, it follows that $\widehat{\mathcal{A}}$ itself must be an upper triangular matrix. The resulting linear system can be solved directly by calculating the inverse of $\widehat{\mathcal{A}}$ and using standard matrix algebra to solve $vec\left(\widehat{Y}\right) = \overline{\mathcal{A}}^{-1} vec\left(\widehat{R}\right)$ for the vector of unknowns $\widehat{Y}$. The inverse of an upper triangular matrix is also upper triangular, so the diagonalization of A can help reduce the number of calculations needed to compute the solution to the transformed Sylvester equation.

Moreover, the resulting linear system lends itself to a recursive implementation that does not require the computation of the inverse of $\widehat{\mathcal{A}}$ explicitly. Let me define $\widehat{\mathcal{A}} = [\widehat{a}_{i,j}]_{i,j=1}^{k^2} \in \mathbb{R}^{k^2\times k^2}$ as well as the column vectors $vec\left(\widehat{Y}\right) = [\widehat{y}_i]_{i=1}^{k^2}$ and $vec\left(\widehat{R}\right) = [\widehat{r}_i]_{i=1}^{k^2}$. Then, for any given $j = 1, ..., k^2$, it holds true that $\widehat{a}_{i,j} = 0$ for all $i = 1, ..., k^2$ and $i < j$. It follows from here that $\widehat{a}_{k^2,k^2}\widehat{y}_{k^2} = \widehat{r}_{k^2}$ pins down $\widehat{y}_{k^2}$. Given $\widehat{y}_{k^2}$, the expression for $\widehat{y}_{k^2-1}$ can be easily obtained from $\widehat{a}_{k^2-1,k^2-1}\widehat{y}_{k^2-1} + \widehat{a}_{k^2-1,k^2}\widehat{y}_{k^2} = \widehat{r}_{k^2-1}$. Given $\widehat{y}_{k^2}$ and $\widehat{y}_{k^2-1}$, the expression for $\widehat{y}_{k^2-2}$ is derived from $\widehat{a}_{k^2-2,k^2-2}\widehat{y}_{k^2-2} + \widehat{a}_{k^2-2,k^2-1}\widehat{y}_{k^2-1} + \widehat{a}_{k^2-2,k^2}\widehat{y}_{k^2} = \widehat{r}_{k^2-2}$. And so on and so forth. Then, once the matrix $\widehat{Y}$ is completed in this recursive way, the last step of the procedure is to recover the solution C to the Sylvester equation. For that, I simply un-do the transformation as follows $C = U\widehat{Y}S^{-1}$.

(b) The computation of matrix C can be further improved whenever F and A are both diagonalizable matrices. The diagonalization theorem implies that the $k \times k$ square matrices F and A are diagonalizable if and only if each of these matrices has k linearly independent eigenvectors, i.e., if and only if the rank of the matrix formed by the eigenvectors is k. I also know that if both matrices are real symmetric, they would be diagonalizable. Furthermore, if the eigenvalues of each matrix are distinct, this is sufficient (albeit not necessary) for each matrix to be diagonalizable as well. Assuming matrices F and A can be diagonalized—i.e., using $F = T\Lambda T^{-1}$ and $A = SMS^{-1}$ where $\Lambda = diag\left(\lambda_1, ..., \lambda_k\right)$ and $M = diag\left(\mu_1, ..., \mu_k\right)$—I obtain the

[7] A matrix A can be diagonalizable, yet have repeated eigenvalues. For example, the identity matrix I_k is diagonal (hence diagonalizable), but has one eigenvalue repeated k times (i.e., $\mu_i = 1$ for all $i = 1, ..., k$).

following transformation of the companion Sylvester equation:

$$\left(T\Lambda T^{-1}\right) C \left(SMS^{-1}\right) - C = H. \tag{33}$$

Multiplying the left-hand side of this matrix equation by T^{-1} and the right-hand side by S, it follows that

$$\Lambda T^{-1}CSM - T^{-1}CS = T^{-1}HS. \tag{34}$$

Let $\widetilde{Y} = T^{-1}CS$ and $\widetilde{R} = T^{-1}HS$. Then, it follows that

$$\Lambda\widetilde{Y}M - \widetilde{Y} = \widetilde{R}. \tag{35}$$

Denoting the (i,j)-th entry of $\widetilde{Y}$ as $\widetilde{y}_{ij}$ and the (i,j)-th entry of $\widetilde{R}$ as $\widetilde{r}_{ij}$, the diagonalized Sylvester equation can be rewritten simple as:

$$\lambda_i\mu_j\widetilde{y}_{ij} - \widetilde{y}_{ij} = \widetilde{r}_{ij},\ \forall i,j = 1,...,k, \tag{36}$$

which means that

$$\widetilde{y}_{ij} = \frac{\widetilde{r}_{ij}}{\lambda_i\mu_j - 1}. \tag{37}$$

Since the eigenvectors and eigenvalues of a diagonalizable matrix can be found with only $O\left(k^3\right)$ operations, the transformed Sylvester equation can be solved more efficiently in this way. Then, the matrix C can be immediately recovered un-doing the transformation as $C = T\widetilde{Y}S^{-1}$.

Other Numerical Algorithms to Solve the Sylvester Equation. (a) Bartels-Stewart Approach. A classical numerical algorithm for solving the Sylvester equation include the Bartels-Stewart algorithm which makes use of the Schur decompositions of F and G to obtain a more efficient algorithm to compute the solution C (Bartels and Stewart (1972)). Using a Schur decomposition as before, the companion Sylvester equation—i.e., $FCA-C = H$ with $F, A, H \in \mathbb{R}^{k\times k}$ given and $C \in \mathbb{R}^{k\times k}$ to be determined—can be re-written as in equation (27). Let Q_{ij} denote a block of the upper triangular matrix Q, and let Y and R be partitioned according to a column partitioning of Q. The key step is to exploit these facts to decompose the transformed Sylvester equation in (27) into smaller Sylvester equations by blocks as follows:

$$KY_1Q_{11} - Y_1 = R_1, \tag{38}$$

$$KY_jQ_{jj} - Y_j = R_j - K\sum\nolimits_{i=1}^{j-1} Y_jQ_{ij},\ \forall j = 2,...,k. \tag{39}$$

Each of these block equations—i.e., (38) – (39)—takes the form of the transformed Sylvester equation in (27) as the sum that appears in (39) is known.[8]

An improved modification of the Bartels-Stewart algorithm, known as the Hessenberg-Schur algorithm, was proposed in Golub *et al.* (1979). This algorithm uses the Hessenberg decomposition instead of the Schur decomposition to transform the companion Sylvester matrix equation (see, e.g., Golub and van Loan (1996), Anderson *et al.* (1996)). The Hessenberg decomposition implies $F = UKU^H$ and $A = VQV^H$ where $U, V \in \mathbb{R}^{k\times k}$ are unitary matrices of dimension k where U^H, V^H denote the corresponding conjugate

[8] If Q is a real matrix from the Schur decomposition, then Q_{jj} for all $j = 1, 2, ..., k$ must be either a scalar or a 2×2 matrix.

transpose and the matrices $K, Q \in \mathbb{R}^{k\times k}$ are the corresponding Hessenberg matrices.

The Sylvester equation is a special case of the Lyapunov equation. Hence, the *dlyap* function in the Control Systems Toolbox which solves the discrete-time Lyapunov equation can be used to solve the companion Sylvester equation in (17) as follows: $C = dlyap\,(F, A, -H)$. This function uses the SLICOT (Subroutine Library In COntrol Theory) library—with a routine that implements the Hessenberg-Schur algorithm. Starting with R2014a, the matrix C can also be computed in Matlab via the *sylvester* function as $C = sylvester\,(F, -inv(A), H * inv(A))$.[9]

(b) Doubling Algorithm. The doubling algorithm exploits the convergence result posited in Proposition 3 which establishes that, for any matrix $H \in \mathbb{R}^{k\times k}$, the companion Sylvester equation in (17) has a unique exact solution given by (24) whenever $\lambda_i \mu_j < 1$ for all $i, j = 1, ..., k$ where $\lambda_1, ..., \lambda_k$ are the eigenvalues of F and $\mu_1, ..., \mu_k$ are the eigenvalues of A. The doubling algorithm defines the following sequence:

$$\begin{aligned} \Lambda_{k+1} &= \Lambda_k \Lambda_k, \\ \Psi_{k+1} &= \Psi_k \Psi_k, \\ C_{k+1} &= C_k + \Lambda_k C_k \Psi_k, \end{aligned} \tag{40}$$

where $\Lambda_0 = F$, $\Psi_0 = A$, and $C_0 = -H$, which converges to the solution of the companion Sylvester equation C. By repeated substitution, it can be shown that each iteration doubles the number of terms in the sum—hence the name of the algorithm—such that

$$C_r = -\sum\nolimits_{s=0}^{2^r-1} F^s H A^s \tag{41}$$

becomes arbitrarily close to the solution $C = -\sum\nolimits_{s=0}^{\infty} F^s H A^s$ as r gets arbitrarily large. Further discussion on this popular algorithm and Matlab codes to implement it can be found in, e.g., Anderson *et al.* (1996).

4 An Application to the Workhorse New Keynesian Model

A Univariate Model of Inflation: The Hybrid Phillips Curve. The hybrid Phillips curve with backward- and forward-looking components features prominently in the New Keynesian literature, arising naturally—for instance—from the well-known Calvo (1983)-type model of price-setting behavior with indexation developed by Yun (1996). The hybrid Phillips curve can be specified generically as

$$\pi_t = \gamma_f \mathbb{E}_t\left(\pi_{t+1}\right) + \gamma_b \pi_{t-1} + e_t, \tag{42}$$

where π_t is the inflation rate, and $\mathbb{E}_t\left(\pi_{t+1}\right)$ is the expected inflation rate next period. The parameters $\gamma_f, \gamma_b > 0$ determine the sensitivity of current inflation to inflation expectations (the forward-looking part) and lagged inflation (the backward-looking part of the model) and satisfy that $\gamma_f + \gamma_b \leq 1$. The variable e_t refers to the exogenous real marginal cost which is assumed to evolve according to a given first-order

[9] Further details on standard implementation methods using Matlab can be found in Sima and Benner (2015). For further references on the Matlab function *dlyap*, see: http://www.mathworks.com/help/control/ref/dlyap.html and http://slicot.org/matlab-toolboxes/basic-control/basic-control-fortran-subroutines. For reference on the Matlab function *sylvester*, see: http://www.mathworks.com/help/matlab/ref/sylvester.html

autoregressive process, i.e.,

$$e_t = \eta e_{t-1} + \sigma_\delta \delta_t, \tag{43}$$

where δ_t is white noise with mean zero and variance of one. The persistence parameter $-1 < \eta < 1$ is expected to be less than one in absolute value to ensure the stationarity of the process, while the parameter $\sigma_\delta > 0$ pins down the real marginal cost shock volatility.

The simple inflation model given by the system in (42) – (43) consists of just one endogenous variable, π_t, and one driving variable, e_t. Hence, it is not difficult to obtain a closed-form solution for inflation in this case and to characterize it in autoregressive form analytically. Using the notation introduced in the previous section, the model-implied relationship between the vector of endogenous variables $W_t = (\pi_t)$ and the vector of driving variables $X_t = (e_t)$ can be cast in the LRE model's canonical form given by (1) – (2) with 1×1 composite matrices $\Phi_1 = (\gamma_b)$, $\Phi_2 = (\gamma_f)$, $\Phi_3 = (1)$, $A = (\eta)$ and $B = (\sigma_\delta)$.

To start, I consider a system reduction to split the solution of the model given in (42) – (43) into a backward-looking part and a forward-looking part. From the quadratic matrix equation (5) in Condition 1 applied to this example, I find that the decoupling depends on the roots of the following characteristic equation:

$$\theta^2 - \frac{1}{\gamma_f}\theta + \frac{\gamma_b}{\gamma_f} = 0, \tag{44}$$

i.e., $\theta_1 = \frac{1 - \sqrt[2]{1-4\gamma_f\gamma_b}}{2\gamma_f}$ and $\theta_2 = \frac{1 + \sqrt[2]{1-4\gamma_f\gamma_b}}{2\gamma_f}$. The solution Θ that permits splitting the backward- and forward-looking parts of the model as in Binder and Pesaran (1995, 1997) requires a stable eigenvalue that lies within the unit circle to exist, i.e. $\Theta = (\theta_1)$ if and only if $|\theta_1| < 1$. As can be easily seen, the existence of the solution Θ depends solely on the parameters γ_f and γ_b. If such a solution exists, then the transformed endogenous variable $Z_t = W_t - \Theta W_{t-1} = (\pi'_t)$ would take the following form: $\pi'_t = \pi_t - \theta_1 \pi_{t-1}$ where θ_1 is the corresponding stable root of the quadratic equation.

Then, as indicated in general form by equation (6) before, the forward-looking part of the hybrid Phillips curve-based model for $Z_t = W_t - \Theta W_{t-1} = (\pi'_t)$ becomes:

$$\Gamma_0 \pi'_t = \Gamma_1 \mathbb{E}_t\left[\pi'_{t+1}\right] + \Gamma_2 e_t, \tag{45}$$

where $\Gamma_0 \equiv (1 - \gamma_f\theta_1)$, $\Gamma_1 \equiv (\gamma_f)$, and $\Gamma_2 \equiv (1)$ are conforming 1×1 matrices and the driving variable e_t remains untransformed. It follows from the properties of the quadratic equation roots that $1 - \gamma_f\theta_1 = \gamma_f\theta_2$. Hence, so long as θ_2 is different from zero, the 1×1 matrix Γ_0 is invertible and the forward-looking part of the LRE model can be re-expressed as in equation (7), i.e.,

$$\pi'_t = F\mathbb{E}_t\left[\pi'_{t+1}\right] + Ge_t, \tag{46}$$

where $F \equiv \left((\gamma_f\theta_2)^{-1}\gamma_f\right) = \left((\theta_2)^{-1}\right)$ and $G \equiv \left((\gamma_f\theta_2)^{-1}\right)$. From the Blanchard and Kahn (1980) conditions applied to the system in (43) and (46), it is straightforward to show that a solution to the forward-looking part of the LRE model exists and is unique if and only if $|\theta_2| > 1$.

Hence, all of this ultimately implies that the full-fledged LRE model in (42) – (43) can be split into a backward- and a forward-looking part and solved uniquely if and only if the roots of the quadratic equation in (44) satisfy that $|\theta_1| < 1$ and $|\theta_2| > 1$. Then, given equation (17) of Lemma 1, the companion Sylvester equation in this case is $FCA - C = H$, where $F = \left(\frac{1}{\theta_2}\right)$, $A = (\eta)$ and $H \equiv -GA = \left(-\left(\frac{\eta}{\gamma_f \theta_2}\right)\right)$. The solution to this equation gives $C = \left(\frac{1}{\gamma_f}\left(\frac{\eta}{\theta_2 - \eta}\right)\right)$ which is well-conditioned and invertible if and only if $\eta \neq 0$ and $\theta_2 \neq \eta$. Hence, the closed-form solution of the forward-looking part of the LRE model under rational expectations maps the shocks into the transformed endogenous variables as in equation (15) above and can be expressed as:

$$\pi'_t = \pi_t - \theta_1 \pi_{t-1} = CA^{-1} e_t, \tag{47}$$

$$CA^{-1} \equiv \left(\frac{1}{\gamma_f}\left(\frac{1}{\theta_2 - \eta}\right)\right), \tag{48}$$

which, together with the autoregressive process specification give in (43), fully describes the inflation dynamics implied by the model. Then, I can immediately infer from this a representation of the dynamics of the transformed inflation rate in autoregressive form as in (16) given by:

$$(\pi_t - \theta_1 \pi_{t-1}) = CAC^{-1}(\pi_{t-1} - \theta_1 \pi_{t-2}) + CA^{-1} B \delta_t, \tag{49}$$

$$CAC^{-1} \equiv (\eta), \; CA^{-1}B \equiv \left(\frac{1}{\gamma_f}\left(\frac{1}{\theta_2 - \eta}\right)\sigma_\delta\right). \tag{50}$$

From an economic point of view, this solution highlights the importance of the backward-looking component of the hybrid Phillips curve in the dynamics of inflation. The persistence of the inflation process is not solely determined by the persistence of the exogenous real marginal cost shock, η, but it also depends on the root θ_1 which is a composite of the backward-looking and forward-looking coefficients of the hybrid Phillips curve (that is, of $\gamma_f > 0$ and $\gamma_b > 0$).[10]

A Bivariate Model of Inflation. The closed-form solution of the univariate hybrid Phillips curve model in (49) shows that it is possible to characterize the solution to an LRE model in finite-order autoregressive form. That, in turn, may permit the identification of the fundamental economic shock driving inflation, δ_t. The method proposed in this paper provides the tools to generalize the logic behind this result to a more general setting with the same number—but multiple—endogenous and driving variables. The approach suggested in the paper helps characterize a finite-order VAR solution for a large class of LRE models by solving a well-known quadratic matrix equation and a companion Sylvester matrix equation—checking the existence, uniqueness, and invertibility of its solution.

To highlight the practical implementation of the method, I start augmenting the hybrid Phillips curve-based model given by (43) and (46) with the following variant of the Taylor (1993) rule with inertia to introduce monetary policy explicitly:

$$i_t = \rho_i i_{t-1} + (1 - \rho_i)[\psi_\pi \pi_t] + m_t, \tag{51}$$

[10] After reversing the transformation, the dynamics of inflation implied by the model can be easily represented with a second-order autoregressive process as in (19) given by: $\pi_t = (\theta_1 + \eta)\pi_{t-1} - \eta\theta_1 \pi_{t-2} + \frac{1}{\gamma_f}\left(\frac{1}{\theta_2 - \eta}\right)\sigma_\delta \delta_t$. The characteristic quadratic equation associated with this second-order autoregressive process is $\phi^2 - (\theta_1 + \eta)\phi - (-\eta\theta_1) = 0$, with roots given by $\phi_1 = \frac{1}{2}(\theta_1 + \eta) - \frac{1}{2}\sqrt[2]{(\theta_1 + \eta)^2 - 4\eta\theta_1}$ and $\phi_1 = \frac{1}{2}(\theta_1 + \eta) + \frac{1}{2}\sqrt[2]{(\theta_1 + \eta)^2 - 4\eta\theta_1}$. Notice that the properties of the roots of the quadratic equation imply that $\phi_1 + \phi_2 = (\theta_1 + \eta)$ and $\phi_1\phi_2 = \eta\theta_1$.

where the policy rate is denoted i_t and the policy inertia is modelled with the parameter $0 < \rho_i < 1$. The policy rule responds to deviations of inflation under the Taylor principle with the parameter ψ_π set to $\psi_\pi > 1$. Here, the associated monetary policy shock m_t follows an exogenously given first-order autoregressive process of the following form:

$$m_t = \rho_m m_{t-1} + \sigma_\xi \xi_t, \tag{52}$$

where ξ_t is assumed to be white noise with mean zero and variance of one, and also uncorrelated at all leads and lags with δ_t. The persistence parameter $-1 < \rho_m < 1$ is expected to be less than one in absolute value to ensure the stationarity of the process, while the parameter $\sigma_\xi > 0$ pins down the monetary shock volatility.

Let me define the vector of endogenous variables as $W_t = (\pi_t, i_t)^T$, the vector of driving variables as $X_t = (e_t, m_t)^T$, and the vector of innovations as $\epsilon_t = (\delta_t, \xi_t)^T$. The augmented model of inflation given by (42) and (51) in matrix form, i.e.,

$$\begin{pmatrix} 1 & 0 \\ -(1-\rho_i)\psi_\pi & 1 \end{pmatrix} \begin{pmatrix} \pi_t \\ i_t \end{pmatrix} = \begin{pmatrix} \gamma_b & 0 \\ 0 & \rho_i \end{pmatrix} \begin{pmatrix} \pi_{t-1} \\ i_{t-1} \end{pmatrix} + \begin{pmatrix} \gamma_f & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbb{E}_t[\pi_{t+1}] \\ \mathbb{E}_t[i_{t+1}] \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e_t \\ m_t \end{pmatrix}, \tag{53}$$

can be expressed in the form of (1) as follows:

$$W_t = \Phi_1 W_{t-1} + \Phi_2 \mathbb{E}_t[W_{t+1}] + \Phi_3 X_t, \tag{54}$$

$$\Phi_1 = \begin{pmatrix} \gamma_b & 0 \\ \gamma_b(1-\rho_i)\psi_\pi & \rho_i \end{pmatrix}, \; \Phi_2 = \begin{pmatrix} \gamma_f & 0 \\ \gamma_f(1-\rho_i)\psi_\pi & 0 \end{pmatrix}, \; \Phi_3 = \begin{pmatrix} 1 & 0 \\ (1-\rho_i)\psi_\pi & 1 \end{pmatrix}. \tag{55}$$

The shock processes in (43) and (52) can be cast in the form indicated by the matrix equation (8) with conforming matrices A and B given by

$$A = \begin{pmatrix} \eta & 0 \\ 0 & \rho_m \end{pmatrix}, \; B = \begin{pmatrix} \sigma_\delta & 0 \\ 0 & \sigma_\xi \end{pmatrix}. \tag{56}$$

This constitutes the canonical form (equation (1)) of the augmented LRE model.

In order to solve the full-fledged model, I split the backward- and forward-looking parts of the model as indicated in Condition 1. In order to solve the quadratic matrix equation in (5), I construct the following two companion matrices

$$D = \begin{bmatrix} 1 & 0 & -\gamma_b & 0 \\ 0 & 1 & -\gamma_b(1-\rho_i)\psi_\pi & -\rho_i \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \; E = \begin{bmatrix} \gamma_f & 0 & 0 & 0 \\ \gamma_f(1-\rho_i)\psi_\pi & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{57}$$

and solve the corresponding generalized eigenvalue problem. As a result, I obtain the following ordered matrix of generalized eigenvalues Q and their associated matrix of eigenvectors V:

$$Q = \begin{pmatrix} \theta_1 & 0 & 0 & 0 \\ 0 & \rho_i & 0 & 0 \\ 0 & 0 & \theta_2 & 0 \\ 0 & 0 & 0 & \infty \end{pmatrix}, \; V = \begin{pmatrix} \frac{\theta_1-\rho_i}{(1-\rho_i)\psi_\pi} & 0 & \frac{\theta_2-\rho_i}{(1-\rho_i)\psi_\pi} & 0 \\ \theta_1 & \rho_i & \theta_2 & 1 \\ \frac{\theta_1-\rho_i}{(1-\rho_i)\psi_\pi\theta_1} & 0 & \frac{\theta_2-\rho_i}{(1-\rho_i)\psi_\pi\theta_2} & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \tag{58}$$

where $\theta_1 = \frac{1-\sqrt[2]{1-4\gamma_f\gamma_b}}{2\gamma_f}$ and $\theta_2 = \frac{1+\sqrt[2]{1-4\gamma_f\gamma_b}}{2\gamma_f}$ are defined exactly as in the univariate case. The matrices Q and V are already ordered so that the two stable eigenvalues come first.

From here it follows that $Q^1 = \begin{pmatrix} \theta_1 & 0 \\ 0 & \rho_i \end{pmatrix}$ and $V^{21} = \begin{pmatrix} \frac{\theta_1-\rho_i}{(1-\rho_i)\psi_\pi\theta_1} & 0 \\ 1 & 1 \end{pmatrix}$, so the companion quadratic matrix equation in (5) has the following solution:

$$\Theta = V^{21}Q^1\left(V^{21}\right)^{-1} = \begin{pmatrix} \theta_1 & 0 \\ (1-\rho_i)\psi_\pi\theta_1 & \rho_i \end{pmatrix}. \tag{59}$$

which is lower triangular. The solution Θ found in (59) permits splitting the backward- and forward-looking parts of the model as in Binder and Pesaran (1995, 1997). To do so, it requires two eigenvalues that are stable and lie within the unit circle to exist, i.e. it requires $|\theta_1| < 1$ and $|\rho_i| < 1$. Given that by construction I already assume that $0 < \rho_i < 1$, the solution Θ that I seek to characterize depends solely on whether the parameters γ_f and γ_b imply also that $|\theta_1| < 1$. If such a solution Θ exists, then the transformed endogenous variables $Z_t = W_t - \Theta W_{t-1} = (\pi'_t, i'_t)$ take the following form: $\pi'_t = \pi_t - \theta_1\pi_{t-1}$ where θ_1 is the same stable root as in the univariate case, and $i'_t = i_t - (1-\rho_i)\psi_\pi\theta_1\pi_{t-1} - \rho_i i_{t-1}$. Notice that the transformed short-term interest rate i'_t needs to be adjusted with its own lag and lagged inflation as well, while the adjustment for the inflation variable π'_t is exactly the same as in the univariate case.

Then, the forward-looking part of the augmented model of inflation can be expressed in the form of (6) as follows:

$$\Gamma_0 Z_t = \Gamma_1\mathbb{E}_t\left[Z_{t+1}\right] + \Gamma_2 X_t, \tag{60}$$

$$\Gamma_0 \equiv (I_k - \Phi_2\Theta) = \begin{pmatrix} 1-\gamma_f\theta_1 & 0 \\ -(1-\rho_i)\psi_\pi\gamma_f\theta_1 & 1 \end{pmatrix} = \begin{pmatrix} \gamma_f\theta_2 & 0 \\ -(1-\rho_i)\psi_\pi\gamma_f\theta_1 & 1 \end{pmatrix}, \tag{61}$$

$$\Gamma_1 \equiv \Phi_2 = \begin{pmatrix} \gamma_f & 0 \\ \gamma_f(1-\rho_i)\psi_\pi & 0 \end{pmatrix}, \; \Gamma_2 \equiv \Phi_3 = \begin{pmatrix} 1 & 0 \\ (1-\rho_i)\psi_\pi & 1 \end{pmatrix}. \tag{62}$$

Whenever Γ_0 is nonsingular, the system of structural relationships for the forward-looking part of the augmented inflation model implied by (6) can be expressed in the form of (7) as:[11]

$$Z_t = F\mathbb{E}_t\left[Z_{t+1}\right] + GX_t, \tag{63}$$

$$F \equiv \left(\Gamma_0\right)^{-1}\Gamma_1 = \begin{pmatrix} \frac{1}{\theta_2\gamma_f} & 0 \\ \frac{\theta_1}{\theta_2}\left(1-\rho_i\right)\psi_\pi & 1 \end{pmatrix}\begin{pmatrix} \gamma_f & 0 \\ \gamma_f\left(1-\rho_i\right)\psi_\pi & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\theta_2} & 0 \\ \left(1-\rho_i\right)\psi_\pi\left(\frac{1}{\theta_2}\right) & 0 \end{pmatrix}, \tag{64}$$

$$G \equiv \left(\Gamma_0\right)^{-1}\Gamma_2 = \begin{pmatrix} \frac{1}{\theta_2\gamma_f} & 0 \\ \frac{\theta_1}{\theta_2}\left(1-\rho_i\right)\psi_\pi & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \left(1-\rho_i\right)\psi_\pi & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\theta_2\gamma_f} & 0 \\ \left(1-\rho_i\right)\psi_\pi\left(\frac{1}{\theta_2\gamma_f}\right) & 1 \end{pmatrix}. \tag{65}$$

The solution of the augmented model in state-space form includes matrices A and B in (56) to represent the stochastic dynamics of the driving variables and the matrix equation in (9) to describe the mapping between the lagged driving variables and their innovations into the endogenous transformed variables in the solution of the forward-looking part of the LRE model. This, in turn, requires the conforming matrices C and D to satisfy the conditions given by (12) – (13).

Hence, matrices C and D must be tied to the matrices F, H, and G that arise from the canonical form of the forward-looking part of the model in (7) where F and H are given above in (64) – (65) and[12]

$$H \equiv -GA = -\begin{pmatrix} \frac{1}{\theta_2\gamma_f} & 0 \\ \left(1-\rho_i\right)\psi_\pi\left(\frac{1}{\theta_2\gamma_f}\right) & 1 \end{pmatrix}\begin{pmatrix} \eta & 0 \\ 0 & \rho_m \end{pmatrix} = \begin{pmatrix} -\frac{\eta}{\theta_2\gamma_f} & 0 \\ -\frac{\eta}{\theta_2}\frac{\psi_\pi}{\gamma_f}\left(1-\rho_i\right) & -\rho_m \end{pmatrix}. \tag{66}$$

From Proposition 2, I can check the existence and uniqueness of a solution C via the companion Sylvester equation in (17). I compute the eigenvalues of F (that is, $\lambda_1 = \frac{1}{\theta_2}$, $\lambda_2 = 0$) and the eigenvalues of A (that is, $\mu_1 = \eta$, $\mu_2 = \rho_m$). Then, given that $\lambda_i\mu_j \neq 1$, for all $i, j = 1, 2$ if and only if $\theta_2 \neq \eta$ and $\theta_2 \neq \rho_m$, I can conclude in that case that a solution C to the companion Sylvester equation exists and is unique.

A straightforward manipulation of the 2^2 equations implied by the Sylvester equation leads me to char-

[11] Notice that $\theta_1 + \theta_2 = \frac{1-\sqrt[2]{1-4\gamma_f\gamma_b}}{2\gamma_f} + \frac{1+\sqrt[2]{1-4\gamma_f\gamma_b}}{2\gamma_f} = \frac{1}{\gamma_f}$.

[12] Computing the eigenvalues of H (that is, $-\beta\frac{\eta}{\theta_2\gamma_f}$ and $-\rho$), I find them to be non-zero if and only if $\eta \neq 0$ and $\rho \neq 0$ since $\theta_2 > 1$ and by assumption $\gamma_f, \gamma_f, \beta > 0$. Hence, the matrix H is nonsingular and invertible.

acterize the conforming matrices C and D as follows:[13]

$$C=\begin{pmatrix} \frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\eta & 0 \\ \psi_\pi\left(1-\rho_i\right)\frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\eta & \rho_m \end{pmatrix},\; D\equiv CA^{-1}B=\begin{pmatrix} \frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\sigma_\delta & 0 \\ \psi_\pi\left(1-\rho_i\right)\frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\sigma_\delta & \sigma_\xi \end{pmatrix}. \tag{67}$$

Checking Condition 2 is straightforward to see that $rank\left(C\right)=2$ if and only if $\eta\neq 0$ and $\rho\neq 0$ since $\theta_2>1$ and by assumption $\gamma_f,\gamma_f>0$ and $\psi_\pi>1$ (the Taylor principle). All of this, in turn, implies that there exists a unique matrix C that solves the companion Sylvester equation in (17) and is also invertible. Hence, the inverse of C is given as:

$$C^{-1}=\begin{pmatrix} \gamma_f\left(\theta_2-\eta\right)\frac{1}{\eta} & 0 \\ -\psi_\pi\left(1-\rho_i\right)\frac{1}{\rho_m} & \frac{1}{\rho_m} \end{pmatrix}. \tag{68}$$

Therefore, the forward-looking part of the augmented inflation model has a VAR(1) representation in the form of (16) which can be expressed as:

$$Z_t=CAC^{-1}Z_{t-1}+CA^{-1}B\epsilon_t, \tag{69}$$

where

$$CAC^{-1}=\begin{pmatrix} \eta & 0 \\ \left(\eta-\rho_m\right)\psi_\pi\left(1-\rho_i\right) & \rho_m \end{pmatrix},\; CA^{-1}B=\begin{pmatrix} \frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\sigma_\delta & 0 \\ \psi_\pi\left(1-\rho_i\right)\frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\sigma_\delta & \sigma_\xi \end{pmatrix}. \tag{70}$$

Then, the finite-order VAR solution of the full-fledged LRE model in (19) becomes:

$$W_t=\Psi_1W_{t-1}+\Psi_2W_{t-2}+\Psi_3\epsilon_t, \tag{71}$$

[13] Using the properties of vectorization and the Kronecker product, I can write the companion Sylvester equation in simple linear form as $\mathcal{A}vec\left(C\right)=vec\left(H\right)$ where:

$$\mathcal{A}:=\left[\left(A^T\otimes F\right)-I_{k^2}\right]=\begin{pmatrix} \frac{\eta}{\theta_2}-1 & 0 & 0 & 0 \\ \left(1-\rho_i\right)\psi_\pi\left(\frac{\eta}{\theta_2}\right) & -1 & 0 & 0 \\ 0 & 0 & \frac{\rho_m}{\theta_2}-1 & 0 \\ 0 & 0 & \left(1-\rho_i\right)\psi_\pi\left(\frac{\rho_m}{\theta_2}\right) & -1 \end{pmatrix}.$$

From here it follows that:

$$vec\left(C\right)=\mathcal{A}^{-1}vec\left(H\right)=\begin{pmatrix} \frac{\theta_2}{\eta-\theta_2} & 0 & 0 & 0 \\ \frac{\eta}{\eta-\theta_2}\left(1-\rho_i\right)\psi_\pi & -1 & 0 & 0 \\ 0 & 0 & -\frac{\theta_2}{\theta_2-\rho_m} & 0 \\ 0 & 0 & -\frac{\rho_m}{\theta_2-\rho_m}\left(1-\rho_i\right)\psi_\pi & -1 \end{pmatrix}\begin{pmatrix} -\frac{\eta}{\theta_2\gamma_f} \\ -\frac{\eta}{\theta_2}\frac{\psi_\pi}{\gamma_f}\left(1-\rho_i\right) \\ 0 \\ -\rho_m \end{pmatrix}=\begin{pmatrix} \frac{\eta}{\gamma_f\left(\theta_2-\eta\right)} \\ \frac{\eta}{\gamma_f\left(\theta_2-\eta\right)}\psi_\pi\left(1-\rho_i\right) \\ 0 \\ \rho_m \end{pmatrix}$$

Undoing the vectorization, the solution C in (67) follows.

where

$$\Psi_1 \equiv \left(\Theta + CAC^{-1}\right) = \begin{pmatrix} \eta + \theta_1 & 0 \\ \psi_\pi \left(1-\rho_i\right)\left(\eta + \theta_1 - \rho_m\right) & \rho_i + \rho_m \end{pmatrix}, \tag{72}$$

$$\Psi_2 \equiv -CAC^{-1}\Theta = \begin{pmatrix} -\eta\theta_1 & 0 \\ -\psi_\pi \left(1-\rho_i\right)\eta\theta_1 & -\rho_i\rho_m \end{pmatrix}, \tag{73}$$

$$\Psi_3 \equiv CA^{-1}B = \begin{pmatrix} \frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\sigma_\delta & 0 \\ \psi_\pi \left(1-\rho_i\right)\frac{1}{\gamma_f}\left(\frac{1}{\theta_2-\eta}\right)\sigma_\delta & \sigma_\xi \end{pmatrix}. \tag{74}$$

From an economic point of view, the solution of the augmented model presented here indicates that there are no spillovers from lagged interest rates into current inflation. Spillovers are only from lagged inflation into the policy rate itself. The policy parameter ψ_π determines the magnitude of the spillovers from lagged inflation into the current policy rate—while the difference in persistence across shocks $(\eta + \theta_1 - \rho_m)$ influences the sign of the spillover from last period's inflation. Moreover, current monetary policy shocks do not contribute to inflation fluctuations. The policy parameter ψ_π plays a key role in explaining the contribution of the monetary policy shock innovation relative to that of the real marginal shock innovation in accounting for the policy rate volatility.

A Multivariate Model of Inflation. Monetary policy has no effect on inflation in the bivariate model given by (42), (43), (51) and (52). This is because the exogenous process for real marginal costs alone drives the dynamics of inflation via the hybrid Phillips curve. In fact, the solution of inflation is exactly the same as that of the univariate case and could have been derived separately since there are no linkages built into the model between the dynamics of inflation and the policy. A further extension of the inflation model that gives monetary policy a distinct role in the determination of real marginal costs and inflation is required. To do so, I follow the approach underlying the workhorse three-equation New Keynesian model which partly endogenizes the real marginal costs and connects them explicitly to monetary policy actions.

To be more precise, I retain the exogenous real marginal cost e_t in the hybrid Phillips curve equation but augment the specification in (42) with an endogenous real marginal cost component that is proportional to the output gap y_t, i.e.,

$$\pi_t = \gamma_f \mathbb{E}_t\left(\pi_{t+1}\right) + \gamma_b \pi_{t-1} + \kappa y_t + e_t, \tag{75}$$

where the parameter $\kappa > 0$ identifies the slope of the hybrid Phillips curve. With external habit formation, the household's utility is affected by aggregate consumption (Campbell and Cochrane (1999)). External habits, therefore, lead the output gap y_t to evolve according to the following hybrid dynamic IS equation:

$$y_t = (1-h)\,\mathbb{E}_t\left(y_{t+1}\right) + hy_{t-1} - \left(\frac{2h-1}{\sigma}\right)\left(i_t - \mathbb{E}_t\left(\pi_{t+1}\right) - r_t^n\right), \tag{76}$$

where $\sigma > 0$ determines the intertemporal elasticity of substitution and the coefficient $h \geq 0$ introduces external habit persistence in the specification. Needless to say, whenever $h = 0$ equation (76) collapses to the familiar time-separable dynamic IS equation. The natural rate r_t^n is assumed to follow an exogenously given first-order autoregressive process:

$$r_t^n = \vartheta r_{t-1}^n + \sigma_\zeta \zeta_t, \tag{77}$$

where ζ_t is assumed to be white noise with zero mean and variance of one, and uncorrelated at all leads and lags with δ_t and ξ_t. The persistence parameter $-1 < \vartheta < 1$ is expected to be less than one in absolute value to ensure the stationarity of the process, while the parameter $\sigma_\zeta > 0$ pins down the natural rate shock volatility. Finally, I also modify the monetary policy rule in (51) as follows:

$$i_t = \rho_i i_{t-1} + (1-\rho_i)\left[\psi_\pi \pi_t + \psi_y y_t\right] + m_t, \tag{78}$$

to allow policy to respond to the endogenous output gap with the corresponding policy parameters set as $\psi_\pi > 1$ and $\psi_y > 0$.[14]

Let me define the vector of endogenous variables as $W_t = (y_t, \pi_t, i_t)^T$, the driving variables as $X_t = (r_t^n, e_t, m_t)^T$, and the vector of innovations as $\epsilon_t = (\zeta_t, \delta_t, \xi_t)^T$. The forward-looking part of the three-equation New Keynesian model of inflation given by (75), (76), and (78) can be expressed as:

$$D_0 W_t = D_1 W_{t-1} + D_2 \mathbb{E}_t\left[W_{t+1}\right] + D_3 X_t, \tag{79}$$

$$D_0 = \begin{pmatrix} 1 & 0 & \left(\frac{2h-1}{\sigma}\right) \\ -\kappa & 1 & 0 \\ -(1-\rho_i)\psi_y & -(1-\rho_i)\psi_\pi & 1 \end{pmatrix},$$
$$D_1 = \begin{pmatrix} h & 0 & 0 \\ 0 & \gamma_b & 0 \\ 0 & 0 & \rho_i \end{pmatrix},\ D_2 = \begin{pmatrix} (1-h) & \left(\frac{2h-1}{\sigma}\right) & 0 \\ 0 & \gamma_f & 0 \\ 0 & 0 & 0 \end{pmatrix},\ D_3 = \begin{pmatrix} \left(\frac{2h-1}{\sigma}\right) & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{80}$$

and re-written, whenever D_0 is nonsingular, in the form of (1):

$$W_t = \Phi_1 W_{t-1} + \Phi_2 \mathbb{E}_t\left[W_{t+1}\right] + \Phi_3 X_t, \tag{81}$$
$$\Phi_1 \equiv (D_0)^{-1} D_1,\ \Phi_2 \equiv (D_0)^{-1} D_2,\ \Phi_3 \equiv (D_0)^{-1} D_3. \tag{82}$$

The shock processes in (43), (52), and (77) can be cast in the form indicated by the matrix equation (8) with conforming matrices A and B given by:

$$A = \begin{pmatrix} \vartheta & 0 & 0 \\ 0 & \eta & 0 \\ 0 & 0 & \rho_m \end{pmatrix},\ B = \begin{pmatrix} \sigma_\zeta & 0 & 0 \\ 0 & \sigma_\delta & 0 \\ 0 & 0 & \sigma_\xi \end{pmatrix}. \tag{83}$$

Then, the solution of the three-equation New Keynesian model can be derived following the steps of the procedure proposed in this paper by solving a companion quadratic matrix equation and a companion Sylvester equation.

In this case, I illustrate the solution of the full-fledged LRE model numerically taking advantage of the set of Matlab codes and functions that I have written to implement the procedures described in this paper.[15]

[14] While the output gap may not be observable, one can rewrite the New Keynesian model in terms of the observable output and a stochastic process for output potential. The output potential is driven by the same shock process as the natural rate, so the three-equation New Keynesian model can still be cast with three endogenous variables—all of them observable with output in place of the output gap—and three exogenous shock processes. A related illustration extending the workhorse New Keynesian model (defined on observable output, inflation and the policy rate) to a two-country setting can be found in Duncan and Martínez-García (2015).

[15] In terms of implementation, the solution of the full-fledged LRE model requires solving a quadratic matrix equation as in (5)

First, let me assume the parameters of the New Keynesian model specified for this application take the conventional parameterization presented in Table 1:

Table 1. Parameterization of the Three-Equation New Keynesian Model

Structural Parameter	Notation	Value
Forward-looking weight on Phillips curve	γ_f	0.7
Backward-looking weight on Phillips curve	γ_b	0.29
Slope of the Phillips curve	κ	0.5
Intertemporal elasticity of substitution	σ	1
External habit formation parameter	h	0.6
Monetary Policy		
Monetary policy inertia	ρ_i	0.85
Monetary policy response to inflation deviations	ψ_π	1.5
Monetary policy response to output gap deviations	ψ_y	0.5
Exogenous Shock Parameters		
Persistence of the natural interest rate shock	ϑ	0.95
Volatility of the natural interest rate shock	σ_ζ	1
Persistence of the cost-push shock	η	0.8
Volatility of the cost-push shock	σ_δ	2
Persistence of the monetary policy shock	ρ_m	0.3
Volatility of the monetary policy shock	σ_ξ	0.7

The parameter values are chosen to illustrate the qualitative features of the three-equation New Keynesian model and the performance of the algorithms for the solution of the companion quadratic matrix equation and the companion Sylvester equation. I use a laptop with Intel(R) Core(TM) i7 with 2.7GHz, 4 cores and 32GB of installed memory (RAM). The wall-clock time elapsed in computing and reporting the numerical given the defaults of the code is 0.150890 seconds (CPU time is: 0.1716 seconds). Using the iterative algorithm to compute the solution to the quadratic matrix equation and Matlab's own implementation of the Hessenberg-Schur algorithm for Lyapunov equations to speed up the computation, the elapsed time falls to 0.088264 seconds (CPU time is: 0.1248 seconds).

The code confirms that, given the parameterization of Table 1, the solution Θ to the companion quadratic matrix equation has its roots inside the unit circle. Furthermore, the code also reports that the solution C to the companion Sylvester equation exists, and is both unique and invertible. Therefore, a straightforward implementation of the algorithm implies that the the three-equation New Keynesian model has a finite-

and a companion Sylvester equation as in (17). In the univariate and bivariate illustrations presented before, the solution of the model can be achieved analytically with standard matrix algebra. However, the paper includes a straightforward Matlab code implementation to compute numerically the solution of the three-equation New Keynesian model via the companion quadratic matrix equation in (5) and the companion Sylvester equation in (17). I use those codes in the rest of this section and made them available in my website: https://sites.google.com/site/emg07uw/. The codes can be downloaded directly using this link: https://sites.google.com/site/emg07uw/econfiles/LRE_model_solution.zip?attredirects=0. Straightforward manipulations of those codes can be made to adapt them to other LRE models that can be cast in the canonical form investigated in this paper—users of the codes are invited to include a citation of this paper in their work.

The Matlab programs and functions appear free of errors, however I do appreciate all feedback, suggestions or corrections that you may have. While users are free to copy, modify and use the code for their work, I do not assume any responsibility for any remaining errors or for how the codes may be used or misused by users other than myself.

order VAR(2) representation given by (19) under the parameterization reported in Table 1 which takes the following form:

$$W_t = \Psi_1 W_{t-1} + \Psi_2 W_{t-2} + \Psi_3 \epsilon_t, \tag{84}$$

$$\Psi_1 = \begin{pmatrix} 1.7257 & -0.2344 & 0.0140 \\ 0.9303 & 0.6517 & 0.1666 \\ 0.3162 & 0.0616 & 1.1885 \end{pmatrix},$$
$$\Psi_2 = \begin{pmatrix} -0.7918 & 0.0567 & 0.0643 \\ -0.8523 & -0.1542 & 0.0417 \\ -0.2512 & -0.0305 & -0.2408 \end{pmatrix}, \tag{85}$$
$$\Psi_3 = \begin{pmatrix} 1.4816 & 0.0097 & -0.8349 \\ 3.3205 & 3.8860 & -1.7135 \\ 0.8582 & 0.8751 & 0.2518 \end{pmatrix},$$

where the corresponding coefficient matrices $\Psi_1 \equiv \left(\Theta + CAC^{-1}\right)$, $\Psi_2 \equiv -CAC^{-1}\Theta$, and $\Psi_3 \equiv CA^{-1}B$—for standard parameterizations—have non-zero entries everywhere. This is in contrast with the bivariate model presented where there were no spillovers from lagged interest rates to current inflation.

From an economic perspective, the solution of the three-equation New Keynesian model shows that there could be spillovers from lagged interest rates to inflation whenever the real marginal costs are partly endogenous and tied to short-term interest rate movements through a dynamic IS equation. Moreover, monetary policy shocks in the full-fledged New Keynesian model contribute to drive the inflation dynamics unlike in the bivariate version of the model presented earlier. Given that the finite-order VAR(2) solution of the workhorse New Keynesian model exists, one may be able to identify the structural monetary policy shock itself—as well as all other fundamental economic shocks driving the economy—directly from the observable data.

The impact of monetary policy on the volatility, cyclicality and persistence of the endogenous variables depends in nonlinear ways on the policy parameters, ψ_π and ψ_y, that describe the systematic part of monetary policy. Here, I exploit the finite-order VAR(2) representation to compute the theoretical moments of the workhorse New Keynesian model for the benchmark parameterization of ψ_π and ψ_y, but also for increasingly higher values of the anti-inflation bias ψ_π. I summarize the key findings describing the business cycle implications in Table 2 below:

Table 2. Key Business Cycle Moments at Different Degrees of Anti-Inflationary Bias

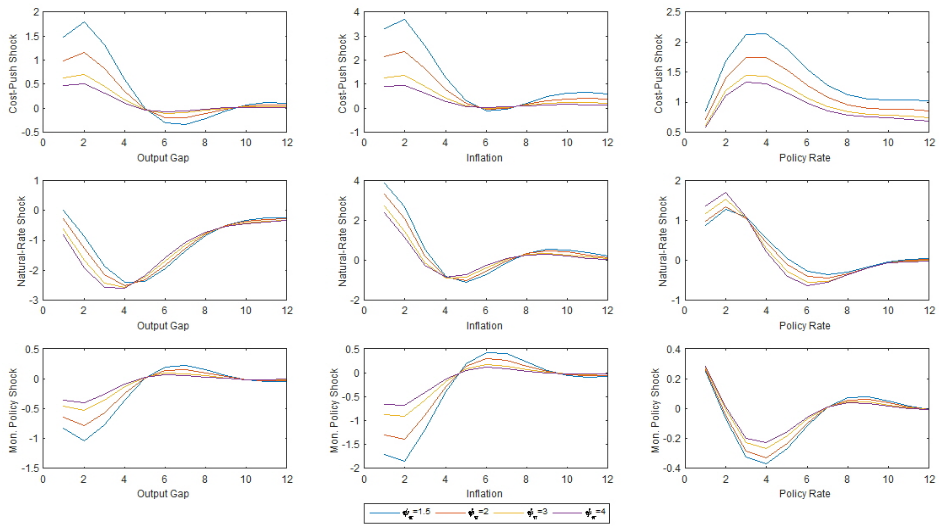
Mean	$\psi_\pi = 1.5$	$\psi_\pi = 2$	$\psi_\pi = 3$	$\psi_\pi = 4$
Output Gap	0	0	0	0
Inflation	0	0	0	0
Policy Rate	0	0	0	0
Std. Deviation				
Output Gap	5.76	5.39	5.31	5.36
Inflation	8.42	6.12	4.29	3.48
Policy Rate	6.32	5.42	4.86	4.73
Contemporaneous Comovement				
Output Gap	1	1	1	1
Inflation	0.46	0.30	0.17	0.11
Policy Rate	0.13	0.03	−0.08	−0.14
First-Order Autocorrelation				
Output Gap	0.88	0.91	0.93	0.93
Inflation	0.69	0.64	0.57	0.52
Policy Rate	0.95	0.94	0.92	0.90

Note: This table summarizes the key theoretical business cycle moments of the three-equation workhorse New Keynesian model keeping the parameterization invariant as in Table 1 except for the policy parameter ψ_π.

As can be seen in Table 2, the monetary policy parameter ψ_π can have a major impact over the business cycle even though the shock processes remain invariant. Generally, a higher value of ψ_π signals a stronger anti-inflation commitment on the part of the monetary authority and is associated with significant declines in inflation and policy rate volatility measured with the theoretical standard deviations—albeit not for the output gap. It is worth pointing out that higher values of the policy parameter lead to a weaker contemporaneous correlation between the output gap and inflation and to a reversal of the contemporaneous correlation between the output gap and the policy rate. I also find a significant weakening of the persistence of inflation as measured by its first-order autocorrelation.

All of this suggests that shifts in the patterns of endogenous volatility, cyclicality, and persistence do not necessarily reflect changes in the underlying shock process driving the economy but could also be the result of changes in monetary policy altering the monetary policy transmission mechanism itself. I illustrate those changes in the transmission mechanism by plotting in Figure 1 the corresponding one-standard deviation (theoretical) impulse response functions (IRFs) at different degrees of the anti-inflation bias, ψ_π.

Figure 1. Theoretical IRFs at Different Degrees of Anti-Inflationary Bias



Note: This figure displays the theoretical impulse response functions (IRFs) of the three-equation workhorse New Keynesian model keeping the parameterization invariant as in Table 1 except for the policy parameter ψ_π.

I note that a higher anti-inflation bias stance on monetary policy ($\uparrow \psi_\pi$) tends to dampen the impact on the endogenous variables of cost-push and monetary policy shock innovations. However, while the impact on inflation of natural-rate shock innovations also declines, a higher ψ_π tends to amplify the effect (at least in the initial quarters) of those same innovations in the output gap and the policy rate.

Hence, these findings confirm the view that identifying structural shocks—monetary policy shocks in particular—is only part of the task, and that the transmission mechanism and its spillovers ultimately depend in nonlinear ways on the features of the prevailing monetary policy regime—the policy parameters of the Taylor (1993) rule in this case—and other deep structural parameters related to preferences, technology, etc. Hence, the method proposed in this paper to solve canonical LRE models makes an important contribution to help us understand the transmission mechanism in a large class of macro models and its mapping into finite-order VAR specifications.

5 Concluding Remarks

I propose a novel approach to solve a large class of LRE models. I represent the LRE model in its canonical form as a system of expectational first-order difference equations. Then, I reduce the canonical form by splitting the backward-looking and forward-looking parts of the model via the solution of a companion quadratic matrix equation (Binder and Pesaran (1995, 1997)). The forward-looking part is solved through the method of undetermined coefficients and describes the unique solution—when one exists—by means of a linear state-space model. Finally, I check the existence and uniqueness of a solution to the purely forward-looking part of the LRE model and characterize its corresponding reduced-form solution in finite-order VAR form via a companion Sylvester equation.

An important contribution of the paper is the derivation of conditions under which the reduced-form, finite-order VAR solution of the forward-looking part of the LRE model via the well-known Sylvester matrix equation is well-defined. The approach proposed in the paper not only provides a way to decouple the canonical form of the LRE model and to characterize the solution in finite-order VAR form, but also checks its properties—existence, uniqueness and invertibility. Furthermore, the paper also makes a contribution to the existing computation literature with the development of an integrated algorithm to solve computationally for models for which a unique finite-order VAR representation exists.

Solving LRE models with a VAR representation by this method is straightforward to implement, efficient, and can be handled easily with standard matrix algebra and conventional computational methods. The paper provides a number of functions for the solution of the companion quadratic matrix equation and the companion Sylvester equation with an economically-relevant application to the study of the transmission mechanism of monetary policy. For this purpose, the paper provides an illustration of the procedure based on the workhorse New Keynesian model and discusses the fundamental features of the model that permit the propagation of monetary policy shocks to inflation and the output gap—with particular attention to the role of the systematic part of the monetary policy rule (which determines the response to inflation and output deviations from their targets). The paper also points out how under the conditions that ensure the existence of a finite-order VAR representation of the solution, the identification of fundamental shocks for empirical research—including the identification of monetary shocks—can be made tractable.

References

[1] Anderson, E. W., McGrattan, E. R., Hansen, L. P. and Sargent, T. J. (1996): "Mechanics of Forming and Estimating Dynamic Linear Economies." *Handbook of Computational Economics*, Vol. 1, pp. 171-252. Elsevier Science, Amsterdam.

[2] Bartels, R. H. and Stewart, G. W. (1972): "Algorithm 432: Solution of the matrix equation $AX+XB=C$." *Comm. ACM*, 15 (9), pp. 820-826.

[3] Binder, M. and Pesaran, M. H. (1995): "Multivariate Rational Expectations Models and Macroeconomic Modelling: A Review and Some New Results," in Pesaran, M. H. and Wickens, M. (eds.), *Handbook of Applied Econometrics: Macroeconomics*, pp. 139-187.

[4] Binder, M. and Pesaran, M. H. (1997): "Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and Their Fully Recursive Computation." *Econometric Theory*, Vol. 13 (6), pp. 877-888.

[5] Blanchard, O. J. and Kahn, C. M. (1980): "The Solution of Linear Difference Models Under Rational Expectations." *Econometrica*, Vol. 48 (5), pp. 1305-1311.

[6] Broze, L., Gouriéroux, C. and Szafarz, A. (1985): "Solutions of Dynamic Linear Rational Expectations Models." *Econometric Theory*, Vol. 1, pp. 341-368.

[7] Broze, L., Gouriéroux, C. and Szafarz, A. (1990): "Reduced Forms of Rational Expectations Models." Chur, Switzerland: Harwood Academic Publishers.

[8] Calvo, G. A. (1983): "Staggered Prices in a Utility-Maximizing Framework". *Journal of Monetary Economics*, Vol. 12 (3), pp. 383-398.

[9] Campbell, H. and Cochrane, J. (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, Vol. 107 (2), pp. 205-251.

[10] Christiano, L. J. (2002): "Solving Dynamic Equilibrium Models by a Method of Undetermined Coefficients." *Computational Economics*, Vol. 20 (1), pp. 21-55.

[11] Duncan, R. and Martínez-García, E. (2015): "Forecasting Local Inflation with Global Inflation: When Economic Theory Meets the Facts." Globalization and Monetary Policy Institute Working Paper No. 235, 2015.

[12] Fernández-Villaverde, J., Rubio-Ramírez, J., Sargent, T. and Watson, M. (2007): "ABCs (and Ds) of Understanding VARs." *American Economic Review*, 97 (3), pp. 1021-1026.

[13] Franchi, M. and Paruolo, P. (2015): "Minimality of State Space Solutions of DSGE Models and Existence Conditions for Their VAR Representation." *Computational Economics*, Vol. 46, pp. 613-626.

[14] Golub, G. H., Nash, S. and van Loan, C. F. (1979): "A Hessenberg-Schur Method for the Problem $AX + XB = C$." *IEEE Trans. Automat. Control*, AC-24(6), pp. 909-913.

[15] Golub, G. H. and van Loan, C. F. (1996): "The Hessenberg and Real Schur Forms." in *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins University Press, pp. 361-372.

[16] Hansen, L. and Sargent, T. (1980): "Formulating and Estimating Dynamic Linear Rational Expectations Models." *Journal of Economic Dynamics and Control*, Vol. 2, pp. 7-46.

[17] Horn, R. A. and Johnson, C. R. (1991): "Topics in Matrix Analysis." Cambridge University Press.

[18] Inoue, A. and Kilian, L. (2002): "Bootstrapping Smooth Functions of Slope Parameters and Innovation Variances in VAR(1) Models." *International Economic Review*, Vol. 43 (2), pp. 309-332.

[19] Jiang, T. and Wei, M. (2003): "On Solutions of the Matrix Equations $X - AXB = C$ and $X - A\overline{X}B = C$." *Linear Algebra and Its Applications*, Vol. 367, pp. 225-233.

[20] King, R. G. and Watson, M. W. (1998): "The Solution of Singular Linear Difference Systems Under Rational Expectations." *International Economic Review*, Vol. 39 (4), pp. 1015-1026.

[21] Klein, P. (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model." *Journal of Economic Dynamics and Control*, Vol. 24 (10), pp. 1405-1423.

[22] Lancaster, P. and Tismenetsky, M. (1985): "The Theory of Matrices with Applications. Second Ed." Academic Press, New York.

[23] Lin, M. and Wimmer, H. K. (2011): "The Generalized Sylvester Matrix Equation, Rank Minimization and Roth's Equivalence Theorem." *Bulletin of the Australian Mathematical Society*, Vol. 84, pp. 441-443.

[24] Lubik, T. A. and Schorfheide, F. (2003): "Computing Sunspot Equilibria in Linear Rational Expectations Models." *Journal of Economic Dynamics and Control*, Vol. 28 (2), pp. 273-285.

[25] Lucas, R. (1976): "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy*, Vol. 1, pp. 19-46.

[26] Muth, J. (1961): "Rational Expectations and the Theory of Price Movements." *Econometrica*, Vol. 29 (3), pp. 315-335.

[27] Ravenna, F. (2007): "Vector Autoregressions and Reduced Form Representations of DSGE Models." *Journal of Monetary Economics*, Vol. 54, pp. 2048-2064.

[28] Sargent, T. (1980): "Rational Expectations and the Reconstruction of Macroeconomics." *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 1980, pp. 15-19.

[29] Sima, V. and Benner, P. (2015): "Solving Linear Matrix Equations with SLICOT." National Institute for Research and Development in Informatics, Bucharest, Romania.

[30] Sims, C. A. (2002): "Solving Linear Rational Expectations Models." *Computational Economics*, Vol. 20 (1), pp. 1-20.

[31] Taylor, J. B. (1993): "Discretion versus Policy Rules in Practice". *Carnegie-Rochester Conference Series*, Vol. 39, pp. 195-214.

[32] Uhlig, H. (1999): "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily." in R. Marimon and A. Scott: *Computational Methods for the Study of Dynamic Economies*. Oxford and New York: Oxford University Press, 1999, pp. 30-61.

[33] Yun, T. (1996): "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles." *Journal of Monetary Economics*, Vol. 37 (2), pp. 345-370.

Appendix

A The LRE Model: From General Form to Canonical Form

The general form of a large class of multivariate LRE models can be written as

$$\Phi_{00} Y_t = \sum\nolimits_{i=1}^{m} \Phi_{i0} Y_{t-i} + \sum\nolimits_{i=0}^{m} \sum\nolimits_{j=1}^{n} \Phi_{ij} \mathbb{E}_{t-i} \left(Y_{t+j-i} \right) + \Omega e_t, \tag{86}$$

where Y_t and X_t are the $r \times 1$ vectors of the endogenous and driving variables of the model, Φ_{ij} for all $i = 0, ..., m$ and $j = 0, ..., n$ and Ω are $r \times r$ conforming matrices, and $\mathbb{E}_t (\cdot)$ represents the conditional expectations operator based on all current and lagged values of Y_t and e_t. I assume that Φ_{00} in (86) is nonsingular and, without loss of generality, set it equal to the identity matrix I_r.

As in Broze *et al.* (1985, 1990) and Binder and Pesaran (1995, 1997), the canonical form of (86) is given by

$$W_t = \Phi_1 W_{t-1} + \Phi_2 \mathbb{E}_t \left[W_{t+1} \right] + X_t, \tag{87}$$

where

$$\begin{aligned}
& W_t = \left(Y_t^T, Y_{t-1}^T, ..., Y_{t-m+1}^T \right)^T, \; U_t^T = \left(Y_t^T, \mathbb{E}_t \left(Y_{t+1}^T \right), ..., \mathbb{E}_t \left(Y_{t+n}^T \right) \right)^T, \\
& \Phi_1 = -D_0^{-1} D_1, \; \Phi_2 = -D_0^{-1} D_1, \\
& X_t = D_0^{-1} \overline{u}_t, \; \overline{u}_t = \left(u_t^T, 0_{q\times 1}^T, ..., 0_{q\times 1}^T \right)^T, \; u_t = \left(\left(\Omega e_t \right)^T, 0_{r\times 1}^T, ..., 0_{r\times 1}^T \right)^T,
\end{aligned}$$

where $\overline{u}_t$ is a column-vector of dimension $k = m(n+1)r$ and u_t is a column-vector of dimension $l = (n+1)r$. The $k \times k$ square matrices D_i, $i = -1, 0, 1$ are defined as

$$D_{-1} = \begin{pmatrix} \Psi_{-1} & 0_l & \ldots & 0_l \\ 0_l & 0_l & \ldots & 0_l \\ & & \ddots & \\ 0_l & 0_l & \ldots & 0_l \end{pmatrix},$$

$$D_0 = \begin{pmatrix} \Psi_0 & \Psi_1 & \ldots & \Psi_{m-1} \\ 0_l & I_l & \ldots & 0_l \\ & & \ddots & \\ 0_l & 0_l & \ldots & I_l \end{pmatrix},$$

$$D_1 = \begin{pmatrix} 0_l & 0_l & \ldots & 0_l & \Psi_m \\ -I_l & 0_l & \ldots & 0_l & 0_l \\ & & \ddots & & \\ 0_l & 0_l & \ldots & -I_l & 0_l \end{pmatrix},$$

where Ψ_i, $i = -1, 0, ..., m$ are square matrices of dimension l given by

$$\Psi_{-1} = \begin{pmatrix} 0_r & 0_r & ... & 0_r & 0_r \\ -I_r & 0_r & ... & 0_r & 0_r \\ & & \ddots & & \\ 0_r & 0_r & ... & -I_r & 0_r \end{pmatrix},$$

$$\Psi_0 = \begin{pmatrix} I_r & -A_{01} & ... & -A_{0n} \\ 0_r & I_r & ... & 0_r \\ & & \ddots & \\ 0_r & 0_r & ... & I_r \end{pmatrix},$$

$$\Psi_i = \begin{pmatrix} -A_{i0} & -A_{i1} & ... & -A_{in} \\ 0_r & 0_r & ... & 0_r \\ & & \ddots & \\ 0_r & 0_r & ... & 0_r \end{pmatrix} \text{ for } i = -1, 0, ..., m.$$

The quadratic determinantal equation method of Pesaran and Binder (1995, 1997) deals with the simultaneous dependence of W_t on its past and its future expected values by transforming the canonical system in (87) so that the resulting system in terms of the transformed variables contains only forward-looking and not backward-looking terms of the transformed variables.

B The Moving Average Representation of the LRE Model

The canonical state-space representation of a large class of linear (and linearized) rational expectations models can be given as in (8) – (9). With the assumption that all are square matrices, it follows from equation (8) that $[I_k - AL]\, X_t = B\epsilon_t$ where I_k is a conforming identity matrix of dimension k and L is the lag operator. If the eigenvalues of A are less than one in modulus, the solution has a moving average representation. In that case, X_t becomes a square summable polynomial given by $X_t = \sum_{j=0}^{\infty} [A]^j\, B\epsilon_{t-j}$; this can be shifted one period back and replaced in (9) to obtain $Z_t = C\sum_{j=0}^{\infty} [A]^j\, B\epsilon_{t-1-j} + D\epsilon_t$.

The moving average representation implied by the state-space solution in (8) – (9) can be inverted under a simple condition and, therefore, represented with a VAR(∞). Whenever the matrix D is nonsingular, replacing the vector ϵ_t in (8) using (9) gives

$$\left[I_k - \left(A - BD^{-1}C\right) L\right] X_t = BD^{-1}Z_t. \tag{88}$$

As noted in Fernández-Villaverde *et al.* (2007), if the eigenvalues of $A - BD^{-1}C$ are strictly less than one in modulus, the inverse of the operator on the left-hand side of (88) gives X_t as a square summable polynomial in L satisfying

$$X_t = \sum_{j=0}^{\infty} \left[A - BD^{-1}C\right]^j BD^{-1}Z_{t-j}. \tag{89}$$

Shifting this expression back one period and replacing it in (9), I obtain

$$Z_t = C\sum\nolimits_{j=0}^{\infty} \left[A - BD^{-1}C\right]^j BD^{-1} Z_{t-1-j} + D\epsilon_t. \tag{90}$$

In general, the VAR(∞) representation of the solution can be reasonably well-approximated by a finite-order structural VAR model as shown in Inoue and Kilian (2002).

In this paper, my methodological contribution is twofold: I propose a new approach to derive the state-space solution in (8) − (9) from the solution of a companion Sylvester equation and a testable condition under which a finite-order VAR representation will be exact rather than an approximation for the LRE model solution. The method proposed here is very efficient at computing the solution to a large class of LRE models. It also checks a straightforward condition—the invertibility of the solution to the companion Sylvester equation—and efficiently derives the corresponding finite-order VAR representation of the solution to the canonical LRE model when one exists.

C An Algorithm to Solve the Quadratic Matrix Equation

Binder and Pesaran (1995, 1997) solve the quadratic matrix equation in (5) for its (stable) real-valued solution Θ with a straightforward iterative algorithm.[16] Alternatively, I look for the stable solution of the quadratic matrix equation in (5) via a companion matrix form where the $k \times k$ matrices Φ_1 and Φ_2 that describe the quadratic matrix equation enter into the following $2k \times 2k$ matrices

$$D = \begin{bmatrix} I_k & -\Phi_1 \\ I_k & 0_k \end{bmatrix}, \; E = \begin{bmatrix} \Phi_2 & 0_k \\ 0_k & I_k \end{bmatrix}. \tag{91}$$

The solution to the generalized eigenvalue problem for the matrix pair (D, E) is a set of $2k$ eigenvalues q_k and their corresponding eigenvectors v_k such that $Dv_k = Ev_kq_k$. Assuming there are at least k stable eigenvalues (those who are inside the unit circle), I can then order the eigenvalues and corresponding eigenvectors so that the k stable eigenvalues come first. Partitioning the ordered matrix of eigenvalues Q and the matrix of eigenvectors V accordingly, i.e. $Q = \begin{bmatrix} Q^1 & 0_k \\ 0_k & Q^2 \end{bmatrix}$ and $V = \begin{bmatrix} V^{11} & V^{12} \\ V^{21} & V^{22} \end{bmatrix}$ where each block matrix is of dimension $k \times k$, the generalized eigenvalue problem can be stated in matrix form as

$$\begin{bmatrix} I_k & -\Phi_1 \\ I_k & 0_k \end{bmatrix} \begin{bmatrix} V^{11} & V^{12} \\ V^{21} & V^{22} \end{bmatrix} = \begin{bmatrix} \Phi_2 & 0_k \\ 0_k & I_k \end{bmatrix} \begin{bmatrix} V^{11} & V^{12} \\ V^{21} & V^{22} \end{bmatrix} \begin{bmatrix} Q^1 & 0_k \\ 0_k & Q^2 \end{bmatrix}, \tag{92}$$

and from here it follows that

$$\begin{bmatrix} V^{11} - \Phi_1 V^{21} & V^{12} - \Phi_1 V^{22} \\ V^{11} & V^{12} \end{bmatrix} = \begin{bmatrix} \Phi_2 V^{11} & \Phi_2 V^{12} \\ V^{21} & V^{22} \end{bmatrix} \begin{bmatrix} Q^1 & 0_k \\ 0_k & Q^2 \end{bmatrix} = \begin{bmatrix} \Phi_2 V^{11} Q^1 & \Phi_2 V^{12} Q^2 \\ V^{21} Q^1 & V^{22} Q^2 \end{bmatrix}. \tag{93}$$

[16] The accompanying codes provided with the paper include a Matlab function that implements the iterative solution of Binder and Pesaran (1995, 1995) together with the Matlab function that recovers the solution using alternatively the generalized eigenvalue problem algorithm described in this Appendix. For a discussion of the quadratic matrix equation and the iterative algorithm, the interested reader is also refered to Binder and Pesaran (1995, 1997) and the accompanying Matlab codes for reference: https://ideas.repec.org/c/dge/qmrbcd/73.html.

This block system implies that $V^{11} = V^{21}Q^1$ and $V^{11} - \Phi_1 V^{21} = \Phi_2 V^{11} Q^1$. Substituting the first expression into the second one gives that $V^{21}Q^1 - \Phi_1 V^{21} = \Phi_2 V^{21} Q^1 Q^1$; then, post-multiplying both sides by $\left(V^{21}\right)^{-1}$ and re-arranging, it follows that:

$$\Phi_2 \left(V^{21} \left(Q^1\right)^2 \left(V^{21}\right)^{-1} \right) - \left(V^{21} Q^1 \left(V^{21}\right)^{-1} \right) + \Phi_1 = 0_k. \tag{94}$$

Defining Θ to be $\Theta \equiv V^{21}Q^1 \left(V^{21}\right)^{-1}$ gives $\Theta^2 = V^{21}Q^1 \left(V^{21}\right)^{-1} V^{21}Q^1 \left(V^{21}\right)^{-1} = V^{21} \left(Q^1\right)^2 \left(V^{21}\right)^{-1}$. Therefore, under this definition of Θ, the expression above is observationally equivalent to the quadratic matrix equation in (5). This suffices to characterize a solution of the quadratic matrix equation based on its k stable eigenvalues.

If such a real-valued solution of the quadratic matrix equation exists with all its eigenvalues inside the unit circle, then it is straightforward to compute by constructing the matrices D and E and finding the ordered solution to the corresponding generalized eigenvalue problem for those matrices. The block matrix Q^1 contains the k stable eigenvalues and the block matrix V^{21} contains the corresponding eigenvectors from the generalized eigenvalue problem after partitioning, then I can construct the solution Θ for the quadratic matrix simply as $\Theta \equiv V^{21}Q^1 \left(V^{21}\right)^{-1}$.