FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

Optimal Policy in Rational Expectations Models: New Solution Algorithms

Richard Dennis
Federal Reserve Bank of San Francisco

September 2005

Working Paper 2001-09
http://www.frbsf.org/publications/economics/papers/2001/wp01-09bk.pdf

Optimal Policy in Rational Expectations Models: New Solution Algorithms*

Richard Dennis†
Federal Reserve Bank of San Francisco

September 2005

Abstract

This paper develops methods to solve for optimal discretionary policies and optimal commitment policies in rational expectations models. These algorithms, which allow the optimization constraints to be conveniently expressed in second-order structural form, are more general than existing methods and are simple to apply. We use several New Keynesian business cycle models to illustrate their application. Simulations show that the procedures developed in this paper can quickly solve small-scale models and that they can be usefully and effectively applied to medium- and large-scale models.

1 INTRODUCTION

Central banks and other policymakers are commonly modeled as agents whose objective is to minimize some loss function subject to constraints that contain forward-looking rational expectations (see Woodford (2002), Clarida, Galí, and Gertler, (1999), or the Taylor (1999) conference volume, for example). But of course Kydland and Prescott (1977) showed that in the absence of some commitment mechanism optimal policies were time-inconsistent. The modern policy literature addresses this time-consistency problem either by simply assuming that the policymaker can commit to not reoptimize or by explicitly modeling the strategic interactions that occur between the various economic agents in the model.[1] When the former approach is taken the policymaker commits to undertake a single optimization and implements

*I would like to thank the associate editor and two anonymous referees for comments. I would also like to thank Kirk Moore for research assistance. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

†Address for correspondence: Economic Research, Mail Stop 1130, Federal Reserve Bank of San Francisco, 101 Market St, CA 94105, USA. Email: richard.dennis@sf.frb.org.

[1] An alternative approach is to assume that the policymaker sets policy according to a "timeless perspective" (Woodford, 1999). According to the timelsss perspective the policymaker optimizing today behaves as it would have chosen to if it had optimized at a time far in the past. Useful discussions and analyses of timelessly optimal policies can be found in Giannoni and Woodford (2002) and Jensen and McCallum (2002).

the chosen policy in all subsequent periods. When the latter approach is pursued the technique is to solve for a subgame-perfect equilibrium. While other time-consistent equilibria could be studied,[2] it is most common to solve for Markov-perfect Stackelberg-Nash equilibria in which the policymaker is the Stackelberg leader and private-sector agents and future policymakers are Stackelberg followers.

Several methods for solving for optimal commitment policies and for Markov-perfect Stackelberg-Nash equilibria (optimal discretionary policies in what follows) have been developed. For commitment, available algorithms include those developed in Currie and Levine (1985, 1993) and Söderlind (1999), which are based on Hamiltonian and Lagrangian methods, respectively, and Backus and Driffill (1986), which transforms the solution to a dynamic programming problem to obtain the commitment policy. For discretion, two popular algorithms are those developed by Oudiz and Sachs (1985) and Backus and Driffill (1986). Both Oudiz and Sachs (1985) and Backus and Driffill (1986) solve for optimal discretionary rules using dynamic programming, but Oudiz and Sachs (1985) solve for the stationary feedback rule by taking the limit as $t \to -\infty$ of a finite horizon problem while Backus and Driffill (1986) solve the asymptotic problem directly.[3] By construction optimal discretionary policies are time-consistent.

A feature common to all of these solution methods is that they require the optimization constraints to be written in state-space form. An attractive feature of the state-space form is that it provides a compact and efficient representation of an economic model in a form that contains only first-order dynamics. However, reliance on the state-space form also has disadvantages. In particular, use of the state-space representation forces the distinction between predetermined and non-predetermined variables and often requires considerable effort to manipulate the model in the required form. Both of these requirements, but especially the latter, mean that these solution methods come with "overhead" costs that can be considerable.

In this paper we develop algorithms to solve for optimal commitment policies and optimal discretionary policies that differ from existing algorithms in that they allow the constraints to be written in structural form rather than in state-space form. As a broadbrush characterization, for small-scale models that have only a few non-predetermined variables, it is generally not too difficult to express the model in state-space form, in which case solving the model using a state-space solution method may be convenient. However, even for models

[2] See Holly and Hughes-Hallett (1989), Amman and Kendrick (1999), Chen and Zadrozney (2002), and Blake (2004), for other approaches.

[3] Söderlind (1999) provides a popular implementation of the Backus and Driffill (1986) algorithm. Krusell, Quadrini, and Rios-Rull (1997) provide an alternative, guess-and-verify, approach to solving for time-consistent policies.

that are only modestly complex it is often much easier to write the model in structural form, and the algorithms presented here have been developed with these models in mind.[4] While of interest in their own right, the algorithms developed here also have several other useful attributes. First, they do not require the predetermined variables to be separated from the non-predetermined variables; second, they can be easily applied to models whose optimization constraints contain the expectation of next period's policy instrument(s); and third, they supply the Euler equation for the optimal discretionary policy, which makes them particularly convenient when a "targeting rule" rather than an "instrument rule" (see Svensson, 2003) is sought.

Although the relative strength of these algorithms is that they can be easily applied to models that are difficult to manipulate into a state-space form, to illustrate their use we apply them to several, reasonably simple, New Keynesian models and compare their properties to existing solution methods. For this comparison, we take the models developed by Galí and Monacelli (2005) and Erceg, Henderson, and Levin (2000) and solve for optimal commitment policies and optimal discretionary policies for a range of policy regimes. Although the algorithms developed in this paper exploit less model structure than the state-space methods, they are still reasonably efficient. To demonstrate that they can usefully solve larger models, we also apply them to the model developed by Fuhrer and Moore (1995).[5]

The remainder of this paper is structured as follows. The following section introduces the components of the policymaker's optimization problem, describing the objective function and the equations that constrain the optimization process. Section 3 shows how to solve for optimal commitment rules. Section 4 turns to discretion. The discretionary problem is formulated in terms of an unconstrained optimization, which contrasts with the more usual dynamic programming or Lagrangian-based approaches. In section 5 we employ two New Keynesian models to calculate the computation times of a range of solution methods. Section 5 also discusses situations for which convergence problems can be encountered when solving for optimal policies. Section 6 concludes while appendices contain technical details.

[4] Dennis (2004a) shows how to solve for optimal simple rules when the optimization constraints are written in structural form.

[5] Dennis and Söderström (2005) apply the methods developed in this paper to the macro-policy model developed by Orphanides and Wieland (1998), which is somewhat more complicated than the Fuhrer and Moore (1995) model.

2 THE OPTIMIZATION PROBLEM

Let $\mathbf{y}_t$ be an $(n \times 1)$ vector of endogenous variables and $\mathbf{x}_t$ be a $(p \times 1)$ vector of policy instruments. For convenience, the variables in $\mathbf{y}_t$ and $\mathbf{x}_t$ represent deviations from nonstochastic steady-state values. The policymaker sets $\mathbf{x}_t$ to minimize the loss function[6]

$$Loss\,(t_0, \infty) = E_{t_0} \sum_{t=0}^{\infty} \beta^t \left[\mathbf{y}_t' \mathbf{W} \mathbf{y}_t + \mathbf{x}_t' \mathbf{Q} \mathbf{x}_t \right], \tag{1}$$

where $\beta \in (0, 1)$ is the discount factor and $\mathbf{W}$ $(n \times n)$ and $\mathbf{Q}$ $(p \times p)$ are symmetric, positive semi-definite, matrices containing policy preferences, or policy tastes. Here E_{t_0} represents the mathematical expectations operator conditional upon period t_0 information. Loss functions like (1) are widely employed in the literature because, with linear constraints, they lead to linear decision rules. Furthermore, as is now well documented, this loss function can represent a second-order approximation to a representative agent's utility function (see Diaz-Gimenez, 1999, for example) and in certain situations it collapses to an expression containing just output and inflation (Woodford, 2002).

The policymaker minimizes (1) subject to the following system of dynamic constraints[7]

$$\mathbf{A}_0 \mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{A}_2 E_t \mathbf{y}_{t+1} + \mathbf{A}_3 \mathbf{x}_t + \mathbf{A}_4 E_t \mathbf{x}_{t+1} + \mathbf{A}_5 \mathbf{v}_t, \tag{2}$$

where $\mathbf{v}_t \sim iid\,[\mathbf{0}, \mathbf{\Omega}]$ is an $(s \times 1, s \leq n)$ vector of innovations and the matrices $\mathbf{A}_0$, $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, $\mathbf{A}_4$, and $\mathbf{A}_5$ contain the model's structural parameters, which are conformable with $\mathbf{y}_t$, $\mathbf{x}_t$, and $\mathbf{v}_t$, as necessary. The dating convention in (2) is such that every element in $\mathbf{y}_{t-1}$ has a known value as of the beginning of period t. $\mathbf{A}_0$ is assumed to be non-singular.

It is not necessary to explicitly include the matrix $\mathbf{A}_5$ in (2) because it is always possible to include $\mathbf{v}_t$ within $\mathbf{y}_t$. However, its presence allows us to easily accommodate models where the innovations are scaled by structural parameters without having to expand the state vector or redefine the innovations. An important feature of (2) is that it contains not just the contemporaneous policy instruments, but also the expectation of next period's instrument vector, which is useful for models that contain an interest rate term structure, for example.[8]

[6] The omission of terms interacting $\mathbf{y}_t$ and $\mathbf{x}_t$ in equation (1) is without loss of generality. When the constraints are in structural form it is always possible in include $\mathbf{x}_t$ within $\mathbf{y}_t$ with $\mathbf{W}$ containing the relevant penalty terms.

[7] Binder and Pesaran (1995) show how a wide class of models, which can contain quite general lead-lag structures and expectations formed with varying information sets, can be written in terms of (2).

[8] Svensson (2000) also solves a model that contains the expected future value of the policy instrument. However, Svensson (2000) considers discretion, but not commitment.

3 Commitment

Under commitment the policymaker optimizes once and never reoptimizes. The assumption is that either the policymaker can adhere to its chosen policy or that the discount factor is sufficiently large to allow the chosen policy to be supported as a reputation equilibrium (see Currie and Levine, 1993, chapter 5). To solve for optimal commitment policies, we observe that from the definition of rational expectations we can write

$$\begin{aligned} \mathbf{y}_{t+1} &= E_t\mathbf{y}_{t+1} + \boldsymbol{\eta}_{t+1} && (3) \\ \mathbf{x}_{t+1} &= E_t\mathbf{x}_{t+1} + \boldsymbol{\mu}_{t+1}, && (4) \end{aligned}$$

where the expectation errors, $\boldsymbol{\eta}_{t+1}$ and $\boldsymbol{\mu}_{t+1}$, are martingale difference sequences. Substituting (3) and (4) into (2) gives

$$\begin{aligned} \mathbf{A}_0\mathbf{y}_t &= \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t+1} + \mathbf{A}_3\mathbf{x}_t + \mathbf{A}_4\mathbf{x}_{t+1} + \boldsymbol{\rho}_t && (5) \\ \boldsymbol{\rho}_t &= \mathbf{A}_5\mathbf{v}_t - \mathbf{A}_2\boldsymbol{\eta}_{t+1} - \mathbf{A}_4\boldsymbol{\mu}_{t+1}. && (6) \end{aligned}$$

To solve the optimization problem from the standpoint of period t_0 we form the Lagrangian

$$\begin{aligned} L &= E_{t_0}\sum_{t=t_0}^{\infty}\beta^{(t-t_0)}\left[\mathbf{y}_t'\mathbf{W}\mathbf{y}_t + \mathbf{x}_t'\mathbf{Q}\mathbf{x}_t\right. && (7) \\ &\quad \left. + 2\boldsymbol{\lambda}_t'\left(\mathbf{A}_0\mathbf{y}_t - \mathbf{A}_1\mathbf{y}_{t-1} - \mathbf{A}_2\mathbf{y}_{t+1} - \mathbf{A}_3\mathbf{x}_t - \mathbf{A}_4\mathbf{x}_{t+1} - \boldsymbol{\rho}_t\right)\right] \end{aligned}$$

and take derivatives with respect to $\mathbf{x}_t$, $\mathbf{y}_t$, and $\boldsymbol{\lambda}_t$, giving[9]

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{x}_t} &= \mathbf{Q}\mathbf{x}_t - \mathbf{A}_3'\boldsymbol{\lambda}_t - \beta^{-1}\mathbf{A}_4'\boldsymbol{\lambda}_{t-1} = \mathbf{0},\ t > t_0 && (8) \\ \frac{\partial L}{\partial \mathbf{y}_t} &= \mathbf{W}\mathbf{y}_t + \mathbf{A}_0'\boldsymbol{\lambda}_t - \beta^{-1}\mathbf{A}_2'\boldsymbol{\lambda}_{t-1} - \beta\mathbf{A}_1'E_t\boldsymbol{\lambda}_{t+1} = \mathbf{0},\ t > t_0 && (9) \\ \frac{\partial L}{\partial \boldsymbol{\lambda}_t} &= \mathbf{A}_0\mathbf{y}_t - \mathbf{A}_1\mathbf{y}_{t-1} - \mathbf{A}_2E_t\mathbf{y}_{t+1} - \mathbf{A}_3\mathbf{x}_t - \mathbf{A}_4E_t\mathbf{x}_{t+1} - \mathbf{A}_5\mathbf{v}_t = \mathbf{0},\ t \geq t_0 && (10) \\ \frac{\partial L}{\partial \mathbf{x}_t} &= \mathbf{Q}\mathbf{x}_t - \mathbf{A}_3'\boldsymbol{\lambda}_t = \mathbf{0},\ t = t_0 && (11) \\ \frac{\partial L}{\partial \mathbf{y}_t} &= \mathbf{W}\mathbf{y}_t + \mathbf{A}_0'\boldsymbol{\lambda}_t - \beta\mathbf{A}_1'E_t\boldsymbol{\lambda}_{t+1} = \mathbf{0},\ t = t_0, && (12) \end{aligned}$$

where $\mathbf{y}_{t_0-1} = \overline{\mathbf{y}}$. The fact that the policymaker behaves differently in the first period than it does in subsequent periods in reflected in the fact that (11) and (12) apply in the initial period ($t = t_0$) whereas (8) and (9) apply in all subsequent periods. However, the recursive nature of

[9] Because W and Q are symmetric positive semi-definite and the policy constraints are convex the Lagrangian is also convex and the first-order conditions are sufficient for locating the minimum.

the system can be restored by employing (8) - (10), for all $t \geq t_0$, but with the initial conditions $\mathbf{y}_{t_0-1} = \overline{\mathbf{y}}$ and $\boldsymbol{\lambda}_{t_0-1} = \mathbf{0}$.[10] The Lagrange multipliers equal zero at the start of the initial period because, as implied by (11) and (12), the policymaker exploits agent's expectations in the initial period, while promising never to do so in the future (Kydland and Prescott, 1980). It is worth noting that it is the very fact that the policymaker behaves differently in the initial period that provides the starting values for the Lagrange multipliers that are needed for the first-order conditions to have a unique, stable, rational expectations equilibrium. The presence of the expected future instrument vector in the optimization constraints presents no special difficulties because the policymaker optimizes only once and the dynamics arising from this term can easily be accounted for during that optimization.

Given $\mathbf{y}_{t_0-1} = \overline{\mathbf{y}}$ and $\boldsymbol{\lambda}_{t_0-1} = \mathbf{0}$, (8) – (10), can be solved in a number of ways. One possibility is to write them in second-order form

$$\begin{bmatrix} \mathbf{0} & \mathbf{A}_0 & -\mathbf{A}_3 \\ \mathbf{A}_0' & \mathbf{W} & \mathbf{0} \\ -\mathbf{A}_3' & \mathbf{0} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_t \\ \mathbf{y}_t \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{A}_1 & \mathbf{0} \\ \beta^{-1}\mathbf{A}_2' & \mathbf{0} & \mathbf{0} \\ \beta^{-1}\mathbf{A}_4' & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_{t-1} \\ \mathbf{y}_{t-1} \\ \mathbf{x}_{t-1} \end{bmatrix} \tag{13}$$

$$+ \begin{bmatrix} \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_4 \\ \beta\mathbf{A}_1' & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} E_t \begin{bmatrix} \boldsymbol{\lambda}_{t+1} \\ \mathbf{y}_{t+1} \\ \mathbf{x}_{t+1} \end{bmatrix} + \begin{bmatrix} \mathbf{A}_5 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} [\mathbf{v}_t] \tag{14}$$

and to apply an undetermined-coefficient technique, such as these developed by Binder and Pesaran (1995), McCallum (1999), or Uhlig (1999). An alternative approach is to write (8) – (10) in first-order form

$$\mathbf{M}\boldsymbol{\Gamma}_t = \mathbf{N}E_t\boldsymbol{\Gamma}_{t+1} + \boldsymbol{\Upsilon}_t, \tag{15}$$

where

$$\mathbf{M} = \begin{bmatrix} \mathbf{0} & -\mathbf{A}_1 & \mathbf{0} & \mathbf{A}_0 & -\mathbf{A}_3 \\ -\beta^{-1}\mathbf{A}_2' & \mathbf{0} & \mathbf{A}_0' & \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ -\beta^{-1}\mathbf{A}_4' & \mathbf{0} & -\mathbf{A}_3' & \mathbf{0} & \mathbf{Q} \end{bmatrix}, \boldsymbol{\Gamma}_t = \begin{bmatrix} \boldsymbol{\lambda}_{t-1} \\ \mathbf{y}_{t-1} \\ \boldsymbol{\lambda}_t \\ \mathbf{y}_t \\ \mathbf{x}_t \end{bmatrix},$$

$$\mathbf{N} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_4 \\ \mathbf{0} & \mathbf{0} & \beta\mathbf{A}_1' & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \boldsymbol{\Upsilon}_t = \begin{bmatrix} \mathbf{A}_5 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} [\mathbf{v}_t]$$

[10] More precisely, the constraints on $\boldsymbol{\lambda}_{t_0-1}$ are $\begin{bmatrix} \mathbf{A}_2' \\ \mathbf{A}_4' \end{bmatrix} \boldsymbol{\lambda}_{t_0-1} = \mathbf{0}$. Because $dim \left[\ker \left(\begin{bmatrix} \mathbf{A}_2' \\ \mathbf{A}_4' \end{bmatrix} \right) \right]$ is not necessarily zero, the restriction in the text, $\boldsymbol{\lambda}_{t_0-1} = \mathbf{0}$, is sufficient, but stronger than necessary.

and to solve the system using an eigenvalue method, such as Anderson and Moore (1985), Klein (2000), King and Watson (2002), Christiano (2002), or Sims (2002). Notice that the vectors in (15) have been allocated such that the predetermined variables, $\boldsymbol{\lambda}_{t-1}$ and $\mathbf{y}_{t-1}$, enter at the top of the system, the "stable" block, while the jump variables, $\boldsymbol{\lambda}_t$, $\mathbf{y}_t$, and $\mathbf{x}_t$, enter at the bottom, the "unstable" block. The order of the variables within these vectors is irrelevant since by construction every element in $\boldsymbol{\lambda}_{t-1}$ and $\mathbf{y}_{t-1}$ has a known value at the beginning of period t. While the order of the variables within the system is an issue for many solution methods, here one simply orders vectors, which makes these Euler equations particularly easy to solve.

Both undetermined-coefficient methods and eigenvalue methods return solutions in the form

$$\begin{bmatrix} \boldsymbol{\lambda}_t \\ \mathbf{y}_t \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}_{11} & \boldsymbol{\theta}_{12} & \mathbf{0} \\ \boldsymbol{\theta}_{21} & \boldsymbol{\theta}_{22} & \mathbf{0} \\ \boldsymbol{\varphi}_1 & \boldsymbol{\varphi}_2 & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_{t-1} \\ \mathbf{y}_{t-1} \\ \mathbf{x}_{t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\theta}_{13} \\ \boldsymbol{\theta}_{23} \\ \boldsymbol{\varphi}_3 \end{bmatrix} \begin{bmatrix} \mathbf{v}_t \end{bmatrix}. \tag{16}$$

The optimal commitment rule can be seen in (16) to depend on the vector of lagged Lagrange multipliers. These Lagrange multipliers enter the rule to ensure that today's policy validates how private sector expectations were formed in the past.

Remark 1: The Euler equations for optimal commitment policy, (8) – (10), do not depend on the covariance matrix of the shocks, $\boldsymbol{\Omega}$. Consequently, optimal commitment policies are certainty equivalent.

Remark 2: Let $\mathbf{z}_t \equiv \begin{bmatrix} \boldsymbol{\lambda}_t' & \mathbf{y}_t' & \mathbf{x}_t' \end{bmatrix}'$, $\widehat{\mathbf{K}} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q} \end{bmatrix}$, and, in obvious notation, (16) be written as $\mathbf{z}_t = \mathbf{H}\mathbf{z}_{t-1} + \mathbf{G}\mathbf{v}_t$. Then, the loss function can be written as $Loss\,(t,\infty) = \left[\mathbf{z}_{t-1}'\mathbf{H}'\widehat{\mathbf{P}}\mathbf{H}\mathbf{z}_{t-1} + \mathbf{v}_t'\mathbf{G}'\widehat{\mathbf{P}}\mathbf{G}\mathbf{v}_t + \frac{\beta}{1-\beta}tr\left(\mathbf{G}'\widehat{\mathbf{P}}\mathbf{G}\boldsymbol{\Omega}\right)\right]$, where $\widehat{\mathbf{P}} \equiv \widehat{\mathbf{K}} + \beta\mathbf{H}'\widehat{\mathbf{P}}\mathbf{H}$ (see Appendix A2).

Remark 3: A consequence of Remark 2 is that $\lim_{\beta\uparrow 1}(1-\beta)\,Loss\,(t,\infty) = tr\,(\mathbf{K}\boldsymbol{\Phi})$ where $\boldsymbol{\Phi}$ is the unconditional variance-covariance matrix of $\begin{bmatrix} \mathbf{y}_t' & \mathbf{x}_t' \end{bmatrix}'$ and $\mathbf{K} \equiv \begin{bmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix}$ (see Appendix A3).

4 Discretion

In this section we consider discretion and develop a numerical procedure that solves for Markov-perfect Stackelberg-Nash equilibria in which the policymaker optimizing today is the Stackelberg leader and private sector agents and future policymakers are Stackelberg followers. Broadly, the innovations in this section are that we allow expected future instruments to enter the optimization constraints and that we base the algorithm on a framework in which the optimization constraints are written in structural form, not in state-space form. The outcome is a solution procedure that is easy to apply and that eliminates the need to represent the optimization constraints in state-space form. The solution method also returns the Euler equations associated with the optimal discretionary policy, not just the decisions rules, which makes it particularly well suited for studying "targeting rules" (Svensson, 2003).

Because the system's state variables are $\mathbf{y}_{t-1}$ and $\mathbf{v}_t$, in the equilibrium that we seek the endogenous variables and the policy instruments will all be functions of these variables. Assume, then, that a stationary solution to the policymaker's optimization problem exists and is given by

$$\mathbf{y}_t = \mathbf{H}_1\mathbf{y}_{t-1} + \mathbf{H}_2\mathbf{v}_t \tag{17}$$

$$\mathbf{x}_t = \mathbf{F}_1\mathbf{y}_{t-1} + \mathbf{F}_2\mathbf{v}_t. \tag{18}$$

It is the time-invariant matrices $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{F}_1$, and $\mathbf{F}_2$ that we seek.

Substituting (17) and (18) into (2) gives

$$\mathbf{D}\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_3\mathbf{x}_t + \mathbf{A}_5\mathbf{v}_t \tag{19}$$

where

$$\mathbf{D} \equiv \mathbf{A}_0 - \mathbf{A}_2\mathbf{H}_1 - \mathbf{A}_4\mathbf{F}_1. \tag{20}$$

The matrix $\mathbf{D}$ embeds how future policymakers respond to movements in $\mathbf{y}_t$ and the policymaker setting $\mathbf{x}_t$ today allows for $\mathbf{D}$, rather than just $\mathbf{A}_0$, when calculating the impact of its policy decisions. Thus, future policymakers are followers with respect to current policymakers.

Using (17) and (18), we can express the loss function in terms of $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{y}_t$ and the instrument vector, $\mathbf{x}_t$ as (see Appendix B1)

$$Loss\,(t,\infty) = \mathbf{y}_t'\mathbf{P}\mathbf{y}_t + \mathbf{x}_t'\mathbf{Q}\mathbf{x}_t + \frac{\beta}{1-\beta}tr\left[\left(\mathbf{F}_2'\mathbf{Q}\mathbf{F}_2 + \mathbf{H}_2'\mathbf{P}\mathbf{H}_2\right)\mathbf{\Omega}\right] \tag{21}$$

where

$$\mathbf{P} \equiv \mathbf{W} + \beta\mathbf{F}_1'\mathbf{Q}\mathbf{F}_1 + \beta\mathbf{H}_1'\mathbf{P}\mathbf{H}_1. \tag{22}$$

While a Lagrangian could be used, it is just as simple to substitute (19) into (21) and transform what was a constrained optimization problem into an unconstrained optimization problem.

Following this substitution we have[11]

$$\begin{aligned} Loss\left(t,\infty\right) &= \left(\mathbf{A}_1\mathbf{y}_{t-1}+\mathbf{A}_3\mathbf{x}_t+\mathbf{A}_5\mathbf{v}_t\right)'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\left(\mathbf{A}_1\mathbf{y}_{t-1}+\mathbf{A}_3\mathbf{x}_t+\mathbf{A}_5\mathbf{v}_t\right) \\ &\quad +\mathbf{x}_t'\mathbf{Q}\mathbf{x}_t+\frac{\beta}{1-\beta}tr\left[\left(\mathbf{F}_2'\mathbf{Q}\mathbf{F}_2+\mathbf{H}_2'\mathbf{P}\mathbf{H}_2\right)\mathbf{\Omega}\right]. \end{aligned} \tag{23}$$

Differentiating (23) with respect to $\mathbf{x}_t$ and setting the resulting derivative equal to zero gives

$$\frac{\partial Loss\left(t,\infty\right)}{\partial\mathbf{x}_t}=\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\left(\mathbf{A}_1\mathbf{y}_{t-1}+\mathbf{A}_5\mathbf{v}_t\right)+\left(\mathbf{Q}+\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3\right)\mathbf{x}_t=\mathbf{0}. \tag{24}$$

Equation (24) can alternatively be expressed in terms of endogenous variables as

$$\mathbf{Q}\mathbf{x}_t+\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{y}_t=\mathbf{0}, \tag{25}$$

which is useful when a "targeting rule" rather than an "instrument rule" (Svensson, 2003) is of interest. Solving (24) for $\mathbf{x}_t$ produces

$$\begin{aligned} \mathbf{x}_t &= -\left(\mathbf{Q}+\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3\right)^{-1}\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\left(\mathbf{A}_1\mathbf{y}_{t-1}+\mathbf{A}_5\mathbf{v}_t\right) \\ &\equiv \mathbf{F}_1\mathbf{y}_{t-1}+\mathbf{F}_2\mathbf{v}_t, \end{aligned} \tag{26}$$

which when inserted back into (19) leads to

$$\begin{aligned} \mathbf{y}_t &= \mathbf{D}^{-1}\left(\mathbf{A}_1+\mathbf{A}_3\mathbf{F}_1\right)\mathbf{y}_{t-1}+\mathbf{D}^{-1}\left(\mathbf{A}_5+\mathbf{A}_3\mathbf{F}_2\right)\mathbf{v}_t \\ &\equiv \mathbf{H}_1\mathbf{y}_{t-1}+\mathbf{H}_2\mathbf{v}_t. \end{aligned} \tag{27}$$

Of course both $\mathbf{P}$ and $\mathbf{D}$ are implicit functions of $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{F}_1$, and $\mathbf{F}_2$. Thus (26) and (27) must be solved for a fixed point to obtain the desired solution. The numerical procedure is as follows

Step 1) Initialize $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{F}_1$, and $\mathbf{F}_2$.

Step 2) Solve for $\mathbf{D}$ according to (20) and for $\mathbf{P}$ from (22), iterating "backward through time" using a method such as the doubling algorithm.[12]

Step 3) Update $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{F}_1$, and $\mathbf{F}_2$ according to

$$\mathbf{F}_1 = -\left(\mathbf{Q}+\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3\right)^{-1}\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_1, \tag{28}$$

$$\mathbf{F}_2 = -\left(\mathbf{Q}+\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3\right)^{-1}\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_5, \tag{29}$$

$$\mathbf{H}_1 = \mathbf{D}^{-1}\left(\mathbf{A}_1+\mathbf{A}_3\mathbf{F}_1\right), \tag{30}$$

$$\mathbf{H}_2 = \mathbf{D}^{-1}\left(\mathbf{A}_5+\mathbf{A}_3\mathbf{F}_2\right). \tag{31}$$

[11] This transformation requires that $\mathbf{D}$ have full rank, which is invariably satisfied because $\mathbf{A}_0$ has full rank.

[12] A description of the doubling algorithm along with other methods for solving matrix Sylvester equations can be found in Anderson, Hansen, McGrattan, and Sargent (1996).

Step 4) Iterate over Steps 2 and 3 until convergence.

As should be clear from the four steps above, the procedure is easy to implement, requiring no more than standard matrix operations. It is worth noting that the procedure does not require the predetermined variables to be separated from the non-predetermined variables, thereby avoiding the matrix partitioning required by state-space methods.

Remark 4: The feedback parameter matrices, $\mathbf{F}_1$ and $\mathbf{F}_2$, are independent of $\mathbf{\Omega}$ and thus optimal discretionary policies are certainty equivalent. Moreover, $\mathbf{P}$ and $\mathbf{D}$, and hence $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{F}_1$, and $\mathbf{F}_2$, are independent of the economy's initial state.

Remark 5: Let $\mathbf{z}_t \equiv \begin{bmatrix} \mathbf{y}_t' & \mathbf{x}_t' \end{bmatrix}'$ and $\mathbf{K} \equiv \begin{bmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix}$. For given $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{H}_2$ matrices, and an $\mathbf{H}_1$ matrix whose spectral radius is less than one, the transition equation for $\mathbf{z}_t$ is $\mathbf{z}_t = \mathbf{H}\mathbf{z}_{t-1} + \mathbf{G}\mathbf{v}_t$, where $\mathbf{H}$ and $\mathbf{G}$ are formed from $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{H}_2$, and $\mathbf{H}_1$ in a straightforward way. Then (1) can be written as $Loss\,(t,\infty) = \mathbf{z}_{t-1}'\mathbf{H}_1'\widetilde{\mathbf{P}}\mathbf{H}_1\mathbf{z}_{t-1} + \mathbf{v}_t'\mathbf{G}'\widetilde{\mathbf{P}}\mathbf{G}\mathbf{v}_t + \frac{\beta}{1-\beta} tr\left(\mathbf{G}'\widetilde{\mathbf{P}}\mathbf{G}\mathbf{\Omega}\right)$, where $\widetilde{\mathbf{P}} \equiv \mathbf{K} + \beta\mathbf{H}'\widetilde{\mathbf{P}}\mathbf{H}$ (see Appendix B2).

Remark 6: As a consequence of Remark 5, $\lim_{\beta\uparrow 1}(1-\beta)\,Loss\,(t,\infty) = tr\,(\mathbf{K}\mathbf{\Phi})$, where $\mathbf{\Phi}$ is the unconditional covariance matrix of $\mathbf{z}_t$ (see Appendix B3).

5 COMPARING SOLUTION METHODS[13]

The previous two sections showed how to solve for optimal commitment policies and optimal discretionary policies when the optimization problem is formulated with the constraints in structural form. This contrasts with the approach taken in Oudiz and Sachs (1985), Currie and Levine (1985, 1993), Backus and Driffill (1986), and Söderlind (1999), which is to minimize

$$Loss\,(t,\infty) = E_t \sum_{j=0}^{\infty} \beta^j \left[\mathbf{p}_{t+j}'\mathbf{S}\mathbf{p}_{t+j} + 2\mathbf{p}_{t+j}'\mathbf{U}\mathbf{x}_{t+j} + \mathbf{x}_{t+j}'\mathbf{R}\mathbf{x}_{t+j}\right] \tag{32}$$

subject to

$$\begin{bmatrix} \mathbf{p}_{1t+1} \\ E_t\mathbf{p}_{2t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{1t} \\ \mathbf{p}_{2t} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} [\mathbf{x}_t] + \begin{bmatrix} \mathbf{s}_{t+1} \\ \mathbf{0} \end{bmatrix}, \tag{33}$$

where $\mathbf{p}_t \equiv \begin{bmatrix} \mathbf{p}_{1t}' & \mathbf{p}_{2t}' \end{bmatrix}'$ combines the predetermined variables, $\mathbf{p}_{1t}$, and the non-predetermined variables, $\mathbf{p}_{2t}$. The key difference between the formulation above and that employed in this paper lies largely in how the constraints are expressed, structural form, (2), or state-space form, (33).

[13] All simulations were conducted in Gauss 6.0.17 using a Pentium IV 2.6Ghz processor with 1.5GB RAM.

In this section we apply the algorithms developed in sections 3 and 4 to two optimization-based New Keynesian models and compare their computation times to state-space solution methods. Because optimal commitment policies and optimal discretionary policies are often studied in the monetary policy literature, the models that we analyze are drawn from that literature. We study the small open-economy model developed by Galí and Monacelli (2005) and closed-economy model developed by Erceg, Henderson, and Levin (2000) (EHL), both are representative of the models used to examine monetary policy issues. For discretion, we compare the algorithm developed in section 4 to those developed in Oudiz and Sachs (1985) and Backus and Driffill (1986). For commitment, we compare the algorithm presented in section 3 to those developed by Backus and Driffill (1986) and Söderlind (1999). Subsequently, we apply our procedures to the larger Fuhrer and Moore (1995) model and discuss situations where convergence problems can sometimes be encountered when solving for optimal policies.

Solving for optimal commitment policies using the approach of section 3 requires solving a rational expectations (RE) model and there are numerous techniques available to perform this task. We take the opportunity to compare two RE solution methods, the "brute force" undetermined-coefficient method presented in Binder and Pesaran (1995) and the QZ-decomposition method presented in Klein (2000). For the latter, we used both the *real* QZ decomposition and the *complex* QZ decomposition,[14] however only the computation times for the *real* QZ decomposition are reported since they were uniformly faster than those for the *complex* QZ decomposition.[15]

5.1 A small-scale open-economy model

In the open-economy model developed by Galí and Monacelli (2005), households consume an aggregate of domestically produced and imported goods. The law-of-one-price applies to imported goods. Domestic firms are monopolistically competitive and subject to Calvo-style price rigidities (Calvo, 1983). Financial assets are internationally tradable and nominal uncovered interest parity holds. The law-of-one-price together with the definition for consumer price inflation implies that the real exchange rate and the terms-of-trade are perfectly correlated, while nominal uncovered interest parity coupled with the law-of-one-price implies that real uncovered interest parity holds.

[14] Routines to perform the *real* QZ decomposition and the *complex* QZ decomposition in Gauss are available from Paul Söderlind's website. These excellent routines were employed in this study.

[15] Söderlind's (1999) method of solving for commitment equilibria also utilizes the *real* QZ decomposition.

For our purposes, the relevant model equations are[16]

$$\begin{aligned}
y_t &= E_t y_{t+1} - \frac{\omega_\alpha}{\sigma}\left(i_t - E_t \pi_{t+1}\right) + g_t \\
\pi_t &= \beta E_t \pi_{t+1} + \kappa_\alpha y_t + u_t \\
\pi_t^c &= \pi_t + \frac{\alpha}{1-\alpha}\left(q_t - q_{t-1}\right) \\
q_t &= E_t q_{t+1} - (1-\alpha)\left(i_t - E_t \pi_{t+1} - i_t^* + E_t \pi_{t+1}^*\right) + (1-\alpha)\,\epsilon_t,
\end{aligned}$$

where y_t denotes the output gap, π_t denotes domestic goods' inflation, π_t^c denotes consumer price inflation rate, q_t denotes the real exchange rate (an increase in q_t represents a real depreciation), i_t and i_t^* denote domestic and foreign nominal interest rates, respectively, π_t^* denotes foreign inflation, and g_t, u_t, and ϵ_t are white noise shocks.[17]

The central bank sets the nominal interest rate, i_t, to minimize the loss function

$$Loss(t,\infty) = E_t \sum_{j=0}^{\infty} \beta^j \left[\left(\pi_{t+j}^c\right)^2 + \lambda y_{t+j}^2 + \nu\left(i_{t+j} - i_{t+j-1}\right)^2\right], \tag{34}$$

where the policy preference parameters, λ and ν, are restricted to be non-negative. In the experiments that follow, we consider a range of policy preference parameters, allowing λ to take on the values 0, 1, and 3 and ν to take on the values 0, $\frac{1}{2}$ and 1.

5.1.1 Solution times

We first assume that policy is set with discretion. Table 1 reports the time it takes to solve the Galí-Monacelli model using the algorithm in section 4 and the algorithms in Oudiz and Sachs (1985), Backus and Driffill (1986), and Söderlind (1999).[18] The procedure presented in Söderlind (1999) is the same as Backus and Driffill (1986) and for this reason both algorithms are attributed the same computation time.

[16] The structural form and the state-space form used for the Galí-Monacelli model are contained in a technical appendix that is available upon request.

[17] We follow Galí and Monacelli (2004) and set $\eta = 1$, $\alpha = 0.4$, $\beta = 0.99$, $\theta = 0.75$, $\sigma = 1$, $\varphi = 3$, $\omega_\alpha = 1 + \alpha(2-\alpha)(\sigma\eta - 1)$, and $\kappa_\alpha = \frac{(1-\theta)(1-\beta\theta)}{\theta}\left(\varphi + \frac{\sigma}{\omega_\alpha}\right)$. We condition on the foreign variables and normalize their values to zero.

[18] The model was solved 100,000 times using each algorithm from which average solution times were computed. The convergence criteria were set so that the algorithms all returned solutions that were identical to thirteen significant figures.

Table 1: Solution Times under Discretion
(hundreds of seconds)

Regime	Section 4	Oudiz-Sachs	Backus-Driffill/ Söderlind
$(\lambda,\nu)=(0,0)$	0.09	0.16	0.16
$(\lambda,\nu)=\left(0,\frac{1}{2}\right)$	0.17	0.16	0.17
$(\lambda,\nu)=(0,1)$	0.19	0.17	0.18
$(\lambda,\nu)=(1,0)$	0.15	0.11	0.11
$(\lambda,\nu)=\left(1,\frac{1}{2}\right)$	0.14	0.12	0.12
$(\lambda,\nu)=(1,1)$	0.16	0.12	0.12
$(\lambda,\nu)=(3,0)$	0.08	0.08	0.08
$(\lambda,\nu)=\left(3,\frac{1}{2}\right)$	0.12	0.09	0.10
$(\lambda,\nu)=(3,1)$	0.12	0.10	0.10

The first point to note is that all of the algorithms are able to quickly obtain the optimal discretionary rule for each parameter combination considered. At the same time, the results indicate that the Oudiz-Sachs algorithm is the fastest for this model, but only marginally; the difference between the fastest algorithm (Oudiz-Sachs) and the slowest algorithm (section 4) is just 0.00012 second on average. It takes slightly longer to solve the model when the constraints are in structural form because the structural form does not exploit as much of the model's structure as the state-space form does. No convergence problems were encountered with any of the algorithms.

Table 2: Solution Times under Commitment
(hundreds of seconds)

	Section 3		Backus-Driffill	Söderlind
Regime	Binder-Pesaran	Real-QZ	Dyn. Prog.	Real-QZ
$(\lambda,\nu)=(0,0)$	–	–	–	–
$(\lambda,\nu)=\left(0,\frac{1}{2}\right)$	0.10	0.13	0.07	0.09
$(\lambda,\nu)=(0,1)$	0.12	0.12	0.09	0.09
$(\lambda,\nu)=(1,0)$	0.15	0.11	–	0.05
$(\lambda,\nu)=\left(1,\frac{1}{2}\right)$	0.21	0.12	0.14	0.08
$(\lambda,\nu)=(1,1)$	0.19	0.12	0.14	0.09
$(\lambda,\nu)=(3,0)$	0.27	0.11	–	0.05
$(\lambda,\nu)=\left(3,\frac{1}{2}\right)$	0.29	0.12	0.21	0.08
$(\lambda,\nu)=(3,1)$	0.30	0.12	0.21	0.08

Turning to the commitment solution methods, Table 2 shows that for the algorithms that use recursive methods – Binder-Pesaran and Backus-Driffill – the solution time is increasing in λ and ν. The algorithms that obtain the commitment policy using the *real* QZ decomposition are both quicker in general that the recursive methods and have solution times that are largely invariant to the loss function's parameterization. As with discretion, use of the structural form weighs on the computation time slightly. Table 2 also shows that for this model the Backus-Driffill algorithm cannot obtain the solution when $\nu = 0$, regardless of the value for λ.[19] In fact, none of the algorithms are able to obtain a solution when $\lambda = \nu = 0$. For the Binder-Pesaran method the matrix recursion used to solve the RE model does not converge, although the iterations do converge when ν is small, but non-zero. The procedures that rely on eigenvalue methods determine the correct number of unstable eigenvalues, but a singularity prevents the jump variables from being related to the predetermined variables in a way that eliminates the unstable dynamics.[20] The fact that the problem occurs regardless of whether the model is in structural or state-space form shows that it is not idiosyncratic to one particular formulation of the optimization problem.

5.2 A small-scale closed-economy model

The second model that we examine is that developed by Erceg, Henderson, and Levin (2000) (EHL). EHL assume that both prices and wages are subject to Calvo-contracts, but the model is otherwise a standard New Keynesian business cycle model with monopolistically competitive firms. When log-linearized about the model's Pareto optimal equilibrium, the relevant equations are[21]

$$
\begin{aligned}
g_t &= E_t g_{t+1} - \frac{1}{\sigma l_c}\left(i_t - E_t \pi_{t+1} - r_t^*\right) \\
\pi_t &= \beta E_t \pi_{t+1} + \kappa_\rho \left(\overline{w}_t - mpl_t\right) \\
w_t &= \beta E_t w_{t+1} + \kappa_\rho \left(mrs_t - \overline{w}_t\right) \\
mpl_t &= \overline{w}_t^* - \lambda_{mpl} g_t \\
mrs_t &= \overline{w}_t^* + \lambda_{mrs} g_t \\
\overline{w}_t &= \overline{w}_{t-1} + w_t - \pi_t \\
r_t^* &= \sigma l_c \left(E_t y_{t+1}^* - y_t^*\right) + \sigma l_q \left(E_t q_{t+1} - q_t\right) \\
\overline{w}_t^* &= w_x^* x_t + w_q^* q_t + w_z^* z_t \\
y_t^* &= y_x^* x_t + y_q^* q_t + y_z^* z_t \\
x_t &= \rho x_{t-1} + \varepsilon_{\pi t},
\end{aligned}
$$

[19] The source of this difficulty is a singularity in the transformation the algorithm uses to rotate the dynamic programming solution into state/co-state space.

[20] This instability differs from the standard form of instability encountered in rational expectations models, which is that the system contains more unstable eigenvalues than there are non-predetermined variables.

[21] The structural form and the state-space form used for the EHL model are contained in a technical appendix that is available upon request.

where g_t denotes the output gap, i_t denotes the short-term nominal interest rate, π_t denotes price inflation, $\overline{w}_t$ denotes the real wage, mpl_t denotes the marginal product of labor, w_t denotes wage inflation, mrs_t denotes the marginal rate of substitution between consumption and leisure, and r_t^*, $\overline{w}_t^*$, and y_t^* denote the Pareto optimal real interest rate, real wage, and level of output, respectively. Finally, x_t, z_t, and q_t, denote a technology shock, a leisure preference shock, and a consumption preference shock.[22]

For the purposes of this section, the central bank's loss function is assumed to be

$$Loss(t,\infty) = E_t \sum_{j=0}^{\infty} \beta^j \left[\pi_{t+j}^2 + \lambda g_{t+j}^2 + \nu \left(i_{t+j} - i_{t+j-1}\right)^2\right]$$

with the nominal interest rate, i_t, serving as the central bank's policy instrument.

5.2.1 Solution times

Table 3 reports the time taken to solve the model under discretion for different parameterizations of the policy objective function.[23]

Table 3: Solution Times under Discretion (hundreds of seconds)

Regime	Section 4	Oudiz-Sachs	Backus-Driffill/ Söderlind
$(\lambda,\nu) = (0,0)$	0.38	0.51	0.51
$(\lambda,\nu) = \left(0,\frac{1}{2}\right)$	1.67	4.02	2.07
$(\lambda,\nu) = (0,1)$	1.83	4.03	2.14
$(\lambda,\nu) = (1,0)$	4.54	2.40	2.52
$(\lambda,\nu) = \left(1,\frac{1}{2}\right)$	4.92	2.34	2.48
$(\lambda,\nu) = (1,1)$	5.03	2.33	2.49
$(\lambda,\nu) = (3,0)$	4.49	2.51	2.63
$(\lambda,\nu) = \left(3,\frac{1}{2}\right)$	4.99	2.48	2.60
$(\lambda,\nu) = (3,1)$	5.16	2.48	2.62

As with the Galí-Monacelli model, none of the algorithms had any difficulty solving the model for any of the parameterizations considered. Aside from two notable outliers, the Oudiz-Sachs algorithm is again the most efficient while the algorithm developed in section

[22] We parameterize the model according to Erceg, Henerson, and Levin (2000), setting $\sigma = \chi = 1.5$, $\alpha = 0.3$, $\theta_w = \theta_p = \frac{1}{3}$, $\zeta_w = \zeta_p = 0.75$, $\rho = 0.95$, $\beta = 0.99$, $\overline{C} = 3.163$, $\overline{Q} = 0.3163$, $\overline{N} = 0.27$, $\overline{Z} = 0.03$, $l_c = \frac{\overline{C}}{\overline{C}-\overline{Q}}$, $l_q = \frac{\overline{Q}}{\overline{C}-\overline{Q}}$, $l_n = \frac{\overline{N}}{1-\overline{N}-\overline{Z}}$, $l_z = \frac{\overline{Z}}{1-\overline{N}-\overline{Z}}$, $\Lambda = \alpha + \chi l_n + (1-\alpha)\sigma l_c$, $\lambda_{mpl} = \frac{\alpha}{1-\alpha}$, $\lambda_{mrs} = \sigma l_c + \frac{\chi l_n}{1-\alpha}$, $\kappa_p = \frac{\left(1-\zeta_p\right)\left(1-\beta\zeta_p\right)}{\zeta_p}$, $\kappa_w = \frac{(1-\zeta_w)(1-\beta\zeta_w)}{\left(1+\chi l_n\left(\frac{1+\theta_w}{\theta_w}\right)\right)\zeta_w}$, $w_x^* = \frac{\chi l_n + \alpha l_c}{\Lambda}$, $w_q^* = -\frac{\alpha\sigma l_q}{\Lambda}$, $w_z^* = \frac{\alpha\chi l_z}{\Lambda}$, $y_x^* = \frac{1+\chi l_n}{\Lambda}$, $y_q^* = \frac{(1-\alpha)\sigma l_q}{\Lambda}$, $y_z^* = \frac{-(1-\alpha)\chi l_z}{\Lambda}$.

[23] The model was solved 100,000 times using each algorithm from which average solution times were computed. The convergence criteria were set so that the algorithms all returned solutions that were identical to twelve significant figures.

4 is typically the slowest, taking on average 1.5 hundreds of a second longer the obtain the solution. While all of the algorithms are notably slower when solving the EHL model than when solving the Galí-Monacelli model, the relative inefficiency of the structural form method is more apparent here because the EHL model contains a larger number of non-predetermined variables, whose lagged values are all treated as candidate state variables by the algorithm.

Table 4: Solution Times under Commitment
(hundreds of seconds)

	Section 3		Backus-Driffill	Söderlind
Regime	Binder-Pesaran	Real-QZ	Dyn. Prog.	Real-QZ
$(\lambda, \nu) = (0, 0)$	2.81	0.26	–	0.10
$(\lambda, \nu) = \left(0, \frac{1}{2}\right)$	0.91	0.32	1.18	0.16
$(\lambda, \nu) = (0, 1)$	1.07	0.35	1.18	0.15
$(\lambda, \nu) = (1, 0)$	2.06	0.27	–	0.06
$(\lambda, \nu) = \left(1, \frac{1}{2}\right)$	1.82	0.29	1.23	0.12
$(\lambda, \nu) = (1, 1)$	1.77	0.30	2.08	0.13
$(\lambda, \nu) = (3, 0)$	2.93	0.25	–	0.07
$(\lambda, \nu) = \left(3, \frac{1}{2}\right)$	3.22	0.30	1.48	0.12
$(\lambda, \nu) = (3, 1)$	3.26	0.31	1.51	0.12

Table 4 reports computation times for the EHL model when solving for optimal commitment policies; several interesting results emerge. First, the computation time is longer when a recursive method is used, such as the Backus-Driffill algorithm or Binder and Pesaran's RE solution method. Second, optimal commitment policies can generally be obtained more quickly than optimal discretionary policies, even if a recursive method is used to solve for the optimal commitment policy. Third, as earlier, the Backus-Driffill algorithm is unable to solve for the optimal commitment policy when $\nu = 0$, but now none of the other algorithms have problems.

5.3 A larger model

The final model that we consider is the well-known model developed by Fuhrer and Moore (1995). The particular specification that we analyze comes from Fuhrer (1997, Table IV).

The model equations are given by

$$
\begin{aligned}
y_t &= 1.45y_{t-1} - 0.47y_{t-2} - 0.34\rho_{t-1} + \epsilon_t \\
\rho_t &= \frac{40}{41}E_t\rho_{t+1} + \frac{1}{41}\left(i_t - E_t\pi_{t+1}\right) \\
p_t &= 0.42w_t + 0.31w_{t-1} + 0.19w_{t-2} + 0.08w_{t-3} \\
\pi_t &= 4\left(p_t - p_{t-1}\right) \\
v_t &= 0.42\bar{w}_t + 0.31\bar{w}_{t-1} + 0.19\bar{w}_{t-2} + 0.08\bar{w}_{t-3} \\
\bar{w}_t &= 0.42v_t + 0.31E_tv_{t+1} + 0.19E_tv_{t+2} + 0.08E_tv_{t+3} \\
&\quad +0.002\left(0.42y_t + 0.31E_ty_{t+1} + 0.19E_ty_{t+2} + 0.08E_ty_{t+3}\right) + \varepsilon_t,
\end{aligned}
$$

where y_t represents the output gap, p_t represent the price level, w_t represents the nominal wage, π_t represents inflation, $\bar{w}_t$ represents the real wage, ρ_t represents the real yield on a ten-year bond, and i_t, the short-term nominal interest rate, serves as the central bank's policy instrument.

Unlike the optimization-based models analyzed above, which each contained only two endogenous state variables, the Fuhrer-Moore model contains seven endogenous state variables, and it also contains a relatively large number of non-predetermined variables. Despite its larger size, the Fuhrer-Moore model can easily be written in second-order structural form and solved using the algorithms developed in sections 3 and 4. As earlier, the central bank's loss function takes the form

$$Loss(t,\infty) = E_t\sum_{j=0}^{\infty}\beta^j\left[\pi_{t+j}^2 + \lambda y_{t+j}^2 + \nu\left(i_{t+j} - i_{t+j-1}\right)^2\right].$$

Table 5 reports the time taken (in hundreds of seconds) to solve the Fuhrer-Moore model using the algorithms presented in section 3 and 4 for different values of λ and ν. The discount factor, β, is set to 0.99.[24]

[24] The model was solved 10,000 times using each algorithm from which average solution times were computed.

Table 5: Solution Times (hundreds of seconds)			
	Discretion	Commitment	
Regime		Binder-Pesaran	Real-QZ
$(\lambda, \nu) = (0, 0)$	–	–	1.40
$(\lambda, \nu) = \left(0, \frac{1}{2}\right)$	4.36	8.37	1.63
$(\lambda, \nu) = (0, 1)$	5.17	9.20	1.63
$(\lambda, \nu) = (1, 0)$	11.70	8.94	1.47
$(\lambda, \nu) = \left(1, \frac{1}{2}\right)$	5.35	6.17	1.64
$(\lambda, \nu) = (1, 1)$	5.17	5.57	1.65
$(\lambda, \nu) = (3, 0)$	16.24	7.26	1.48
$(\lambda, \nu) = \left(3, \frac{1}{2}\right)$	9.26	7.53	1.66
$(\lambda, \nu) = (3, 1)$	8.73	7.21	1.72

Table 5 underscores the point that it is much more efficient to solve for optimal commitment policies the *real* QZ decomposition than it is to use the Binder-Pesaran solution method, which employs recursive iterations, and that optimal commitment policies can be obtained more quickly than optimal discretionary policies. Nevertheless, averaging across the parameterizations for which the solution could be obtained, it takes less than one tenth of a second to solve for the optimal discretionary policy. Perhaps the most notable feature of Table 5, however, is that when $\lambda = \nu = 0$ the optimal commitment policy is obtained when using the *real* QZ decomposition, but not (to sufficient accuracy) when using Binder-Pesaran. Moreover, for this parameterization of the loss function, the optimal discretionary policy could not be obtained.

5.4 Convergence issues

For the three models analyzed above, one or more of the solution algorithms could not obtain the optimal (commitment/discretionary) policy for some parameterizations of the loss function. Looking closely at the offending parameterizations, it is notable that they are all parameterizations for which no penalty is placed on the policy instruments, i.e., cases where the central bank (say) is not concerned about interest rate smoothing/stabilization. Similar instances are documented in Svensson (2000) and Dennis and Söderström (2005) and, together, these occurrence highlight the fact that the properties of the solution algorithms, particularly those for obtaining discretionary policies, have not been explored. In fact, when discussing their algorithm for obtaining optimal discretionary policies, Oudiz and Sachs (1985, pp311) comment "We do not know of any general result concerning the convergence of this process. However in our empirical applications we have not run into major problems." Similarly, Söderlind (1999, pp819), notes that "The general properties of this algorithm are unknown.

Practical experience suggests that it is often harder to find the discretionary equilibrium than the commitment equilibrium. It is unclear if this is due to the algorithm."

The approach taken in this paper provides insights as to why, and under what circumstances, convergence problems may be encountered when solving for optimal discretionary policies. For this purpose, the key equations in the algorithm are

$$\mathbf{P} = \mathbf{W} + \beta\mathbf{F}_1'\mathbf{Q}\mathbf{F}_1 + \beta\mathbf{H}_1'\mathbf{P}\mathbf{H}_1 \tag{35}$$

$$\mathbf{D} = \mathbf{A}_0 - \mathbf{A}_2\mathbf{H}_1 - \mathbf{A}_4\mathbf{F}_1 \tag{36}$$

$$\mathbf{H}_1 = \mathbf{D}^{-1}\left(\mathbf{I} - \mathbf{A}_3\left(\mathbf{Q} + \mathbf{A}_3'\mathbf{D}^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3\right)^{-1}\mathbf{A}_3'\mathbf{D}'^{-1}\mathbf{P}\mathbf{D}^{-1}\right)\mathbf{A}_1 \tag{37}$$

where (35) and (36) are taken directly from section 4 and (37) is obtained by substituting (28) into (30). By inspection, difficulties obtaining the optimal discretionary policy can arise when:

1. The discount factor is close to one and the $\mathbf{H}_1$ matrix associated with the optimal discretionary policy has a spectral radius close to one. In this situation it take a long time for recursive methods to obtain the fixed point $\mathbf{P}$. If this problem occurs, it may be useful to solve for $\mathbf{P}$ using a non-recursive method, such as the Hessenberg-Schur algorithm described in Anderson, et al. (1996).

2. The iterations produce a $\mathbf{D}$ matrix that is singular. Because $\mathbf{A}_2$ and $\mathbf{A}_4$ are typically sparse and $\mathbf{A}_0$ has full rank, a singularity in $\mathbf{D}$ is unlikely to occur.

3. The iterations produce a $\left(\mathbf{Q} + \mathbf{A}_3'\mathbf{D}^{-1}\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3\right)$ matrix that is singular. This will occur if both $\mathbf{Q}$ and $\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3$ equal null matrices, i.e. if the loss function does not penalize movements in the policy instrument(s) and the target variables are contemporaneously unaffected by movements in the policy instruments. When both $\mathbf{P}\mathbf{D}^{-1}\mathbf{A}_3$ and $\mathbf{Q}$ equal null matrices the optimal discretionary policy is not uniquely determined by setting the derivative of the loss function with respect to the policy instrument(s) to zero, see (24).

Among these three situations, the first will likely increase the computation time, the second is very unlikely to occur, and the third accounts for the problems encountered above, in Svensson (2000), and in Dennis and Söderström (2005), which all involve a loss function that does not penalize the policy instrument. In fact, the assumption that $\mathbf{Q}$ is positive definite – rather than just positive semi-definite –, which would preclude this problem, is standard in linear-quadratic control theory (Rustem and Zarrop, 1979; Anderson, et al. 1996, pp175).

We did not restrict $\mathbf{Q}$ to be positive definite in our analysis because policies where the central bank (say) does not smooth or stabilize interest rates are of broad interest and can be obtained for many models.

6 CONCLUSION

This paper has presented new algorithms to solve for optimal commitment rules and optimal discretionary rules in rational expectation models. These algorithms differ from those developed by Oudiz and Sachs (1985), Backus and Driffill (1986), and Söderlind (1999) in that they do not require the optimization constraints to be written in state-space form. This paper has also extended the class of models that can be analyzed, allowing expectations of future policy instruments to enter the optimization constraints, which can be useful when solving models that contain an interest rate term structure, and showed how to derive the Euler equation for the optimal discretionary policy, which is needed to study targeting rules.

After setting up the optimization problem in section 2, the approaches taken to solve for optimal commitment policies and optimal discretionary policies were described in sections 3 and 4, respectively. Section 5 took two optimization-based New Keynesian models and used a range of solution algorithms to solve them for optimal commitment policies and optimal discretionary policies for a variety of policy objective functions. For these models it was feasible to analytically derive state-space representations. While the relative advantage of the solution methods developed in this paper is that they save the user from having to manipulate a model into state-space form, computational experiments on these small-scale models showed that the state-space solution methods are generally more efficient, precisely because they exploit more of the model's structure. Nevertheless, the methods developed here can be employed to solve larger models and for small models they are efficient enough to be used in estimation exercises, which may involve a model being solved many times within a hill-climbing routine (see Dennis, 2004b, for an application). Section 5 also solved the Fuhrer-Moore model, a model that can be much more quickly expressed in structural form than in a state-space form.

The paper has also discussed some of the computational problems that can be encountered when solving for optimal policies and showed why, in practice, placing a small penalty on movements in the policy instrument(s), invariably overcomes these problems. These experiments also showed that it is generally more efficient to solve for optimal commitment policies by solving the underlying rational expectations model using the *real* QZ-decomposition rather than either the *complex* QZ-decomposition or the Binder-Pesaran algorithm.

A APPENDIX

A.1 Mathematical preliminary

Assume that $0 < \beta < 1$ and that the spectral radius of the matrix $\boldsymbol{\theta}$ is less than one, then the infinite series $\mathbf{S} = \sum_{j=0}^{\infty} \beta^j \boldsymbol{\theta}'^j \mathbf{W} \boldsymbol{\theta}^j$ is convergent. Consequently, $\beta \boldsymbol{\theta}' \mathbf{S} \boldsymbol{\theta} = \sum_{j=1}^{\infty} \beta^j \boldsymbol{\theta}'^j \mathbf{W} \boldsymbol{\theta}^j = \mathbf{S} - \mathbf{W}$, and $\mathbf{S}$ can be found as the fixed point of $\mathbf{S} = \mathbf{W} + \beta \boldsymbol{\theta}' \mathbf{S} \boldsymbol{\theta}$.

A.2 Establishing Remark 2

Let $\mathbf{z}_t \equiv \begin{bmatrix} \boldsymbol{\lambda}_t \\ \mathbf{y}_t \\ \mathbf{x}_t \end{bmatrix}$ and $\widehat{\mathbf{K}} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q} \end{bmatrix}$. From (27) the recursive equilibrium law of motion for $\mathbf{z}_t$ can be expressed as $\mathbf{z}_t = \mathbf{H}\mathbf{z}_{t-1} + \mathbf{G}\mathbf{v}_t$. Given this law of motion, we have

$$\begin{aligned} Loss\,(t,\infty) &= E_t \sum_{j=0}^{\infty} \beta^j \mathbf{z}'_{t+j} \mathbf{K} \mathbf{z}_{t+j} \\ &= \left[\mathbf{z}'_t \left(\sum_{j=0}^{\infty} \beta^j \mathbf{H}'^j \mathbf{K} \mathbf{H}^j \right) \mathbf{z}_t + \frac{\beta}{1-\beta} \left(\sum_{j=0}^{\infty} \beta^j tr\left(\mathbf{G}' \mathbf{H}_1'^j \mathbf{K} \mathbf{H}_1^j \mathbf{G} \boldsymbol{\Omega} \right) \right) \right], \end{aligned}$$

which using the result regarding convergent infinite series in Appendix A1 can be simplified to

$$Loss\,(t,\infty) = \left[\mathbf{z}'_t \widehat{\mathbf{P}} \mathbf{z}_t + \frac{\beta}{1-\beta} tr\left(\mathbf{G}' \widehat{\mathbf{P}} \mathbf{G} \boldsymbol{\Omega} \right) \right],$$

where $\widehat{\mathbf{P}} \equiv \widehat{\mathbf{K}} + \beta \mathbf{H}' \widehat{\mathbf{P}} \mathbf{H}$. Finally, again using the law of motion for $\mathbf{z}_t$ we obtain

$$Loss\,(t,\infty) = \left[\mathbf{z}'_{t-1} \mathbf{H}' \widehat{\mathbf{P}} \mathbf{H} \mathbf{z}_{t-1} + \mathbf{v}'_t \mathbf{G}' \widehat{\mathbf{P}} \mathbf{G} \mathbf{v}_t + \frac{\beta}{1-\beta} tr\left(\mathbf{G}' \widehat{\mathbf{P}} \mathbf{G} \boldsymbol{\Omega} \right) \right],$$

as required.

A.3 Establishing Remark 3

From Appendix A2 the loss function under commitment can be written as

$$Loss\,(t,\infty) = \left[\mathbf{z}'_t \widehat{\mathbf{P}} \mathbf{z}_t + \frac{\beta}{1-\beta} tr\left(\mathbf{G}' \widehat{\mathbf{P}} \mathbf{G} \boldsymbol{\Omega} \right) \right], \tag{A.1}$$

where $\widehat{\mathbf{P}} \equiv \widehat{\mathbf{K}} + \beta \mathbf{H}' \widehat{\mathbf{P}} \mathbf{H}$, and the economy's evolution is given by

$$\mathbf{z}_t = \mathbf{H}\mathbf{z}_{t-1} + \mathbf{G}\mathbf{v}_t. \tag{A.2}$$

From A.2 the unconditional variance-covariance matrix for $\mathbf{z}_t$ is the fixed point of

$$\mathbf{\Sigma} = \mathbf{H\Sigma H}' + \mathbf{G\Omega G}', \tag{A.3}$$

from which the unconditional variance-covariance matrix for $\begin{bmatrix} \mathbf{y}_t' & \mathbf{x}_t' \end{bmatrix}'$, $\mathbf{\Phi}$, can be obtained. Multiplying A.1 through by $(1-\beta)$ gives

$$\begin{aligned}
(1-\beta)\, Loss\,(t,\infty) &= (1-\beta)\,\mathbf{z}_t'\widehat{\mathbf{P}}\mathbf{z}_t + \beta tr\left(\mathbf{G}'\widehat{\mathbf{P}}\mathbf{G\Omega}\right) \\
&= (1-\beta)\,\mathbf{z}_t'\widehat{\mathbf{P}}\mathbf{z}_t + \beta tr\left(\widehat{\mathbf{P}}\mathbf{G\Omega G}'\right).
\end{aligned}$$

Now employing A.3 we have

$$\begin{aligned}
(1-\beta)\, Loss\,(t,\infty) &= (1-\beta)\,\mathbf{z}_t'\widehat{\mathbf{P}}\mathbf{z}_t + \beta tr\left[\widehat{\mathbf{P}}\left(\mathbf{\Sigma} - \mathbf{H\Sigma H}'\right)\right] \\
&= (1-\beta)\,\mathbf{z}_t'\widehat{\mathbf{P}}\mathbf{z}_t + \beta tr\left(\widehat{\mathbf{P}}\mathbf{\Sigma}\right) - \beta tr\left(\mathbf{H}'\widehat{\mathbf{P}}\mathbf{H\Sigma}\right) \\
&= (1-\beta)\,\mathbf{z}_t'\widehat{\mathbf{P}}\mathbf{z}_t + \beta tr\left(\widehat{\mathbf{P}}\mathbf{\Sigma}\right) - tr\left[\left(\widehat{\mathbf{P}} - \widehat{\mathbf{K}}\right)\mathbf{\Sigma}\right] \\
&= (1-\beta)\,\mathbf{z}_t'\widehat{\mathbf{P}}\mathbf{z}_t - (1-\beta)\, tr\left(\widehat{\mathbf{P}}\mathbf{\Sigma}\right) + tr\left(\widehat{\mathbf{K}}\mathbf{\Sigma}\right).
\end{aligned}$$

Provided the spectral radius of $\mathbf{H}$ is less than one, $\widehat{\mathbf{P}}$ will remain bounded even in the limit as $\beta \uparrow 1$. Therefore, $\lim\limits_{\beta\uparrow 1} Loss\,(t,\infty) = tr\left(\widehat{\mathbf{K}}\mathbf{\Sigma}\right)$. However, defining $\mathbf{K} \equiv \begin{bmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix}$ so that $\widehat{\mathbf{K}} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K} \end{bmatrix}$ only the elements in $\mathbf{\Sigma}$ associated with $\mathbf{y}_t$ and $\mathbf{x}_t$ (i.e., $\mathbf{\Phi}$) are relevant. Consequently, we have $\lim\limits_{\beta\uparrow 1} Loss\,(t,\infty) = tr\,(\mathbf{K\Phi})$, as required.

B APPENDIX

B.1 Deriving equation (21)

We show that

$$\begin{aligned}
Loss\,(t,\infty) &= E_t \sum_{j=0}^{\infty} \beta^j \left(\mathbf{y}_{t+j}'\mathbf{W}\mathbf{y}_{t+j} + \mathbf{x}_{t+j}'\mathbf{Q}\mathbf{x}_{t+j}\right) \\
&= \mathbf{y}_t'\mathbf{P}\mathbf{y}_t + \mathbf{x}_t'\mathbf{Q}\mathbf{x}_t + \frac{\beta}{1-\beta} tr\left[\left(\mathbf{F}_2'\mathbf{Q}\mathbf{F}_2 + \mathbf{H}_2'\mathbf{P}\mathbf{H}_2\right)\mathbf{\Omega}\right],
\end{aligned}$$

where $\mathbf{P} \equiv \mathbf{S} + \beta\mathbf{R} = \mathbf{W} + \beta\mathbf{F}_1'\mathbf{Q}\mathbf{F}_1 + \beta\mathbf{H}_1'\mathbf{P}\mathbf{H}_1$.

To establish this equality we exploit the properties of convergent geometric series, and the requirement that, in equilibrium, $\mathbf{y}_{t+j} = \mathbf{H}_1\mathbf{y}_{t+j-1} + \mathbf{H}_2\mathbf{v}_{t+j}$, and $\mathbf{x}_{t+j} = \mathbf{F}_1\mathbf{y}_{t+j-1} + \mathbf{F}_2\mathbf{v}_{t+j}$,

$\forall\ j > 0$. First write

$$E_t \sum_{j=0}^{\infty} \beta^j \left(\mathbf{y}'_{t+j} \mathbf{W} \mathbf{y}_{t+j} + \mathbf{x}'_{t+j} \mathbf{Q} \mathbf{x}_{t+j} \right) = \mathbf{x}'_t \mathbf{Q} \mathbf{x}_t + E_t \sum_{j=0}^{\infty} \beta^j \left(\mathbf{y}'_{t+j} \mathbf{W} \mathbf{y}_{t+j} \right) + E_t \sum_{j=1}^{\infty} \beta^j \left(\mathbf{x}'_{t+j} \mathbf{Q} \mathbf{x}_{t+j} \right). \tag{B.1}$$

The first term on the RHS of B1 is in the appropriate form. We will treat the second and third terms on the RHS separately, beginning with the second term. We can write

$$E_t \sum_{j=0}^{\infty} \beta^j \left(\mathbf{y}'_{t+j} \mathbf{W} \mathbf{y}_{t+j} \right) = \mathbf{y}'_t \left(\sum_{j=0}^{\infty} \beta^j \mathbf{H}_1'^j \mathbf{W} \mathbf{H}_1^j \right) \mathbf{y}_t + \beta \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} \beta^{(l+j)} tr \left(\mathbf{H}'_2 \mathbf{H}_1'^l \mathbf{W} \mathbf{H}_1^l \mathbf{H}_2 \mathbf{\Omega} \right).$$

Assuming that the spectral radius of $\mathbf{H}_1$ is less than one, the result in Appendix A1 leads to

$$E_t \sum_{j=0}^{\infty} \beta^j \left(\mathbf{y}'_{t+j} \mathbf{W} \mathbf{y}_{t+j} \right) = \mathbf{y}'_t \mathbf{S} \mathbf{y}_t + \frac{\beta}{1-\beta} tr \left(\mathbf{H}'_2 \mathbf{S} \mathbf{H}_2 \mathbf{\Omega} \right), \tag{B.2}$$

where $\mathbf{S} \equiv \mathbf{W} + \beta \mathbf{H}'_1 \mathbf{S} \mathbf{H}_1$.

Turning to the third term on the RHS of B.1

$$E_t \sum_{j=1}^{\infty} \beta^j \left(\mathbf{x}'_{t+j} \mathbf{Q} \mathbf{x}_{t+j} \right) = \beta \mathbf{y}'_t \left(\sum_{j=0}^{\infty} \beta^j \mathbf{H}_1'^j \mathbf{F}'_1 \mathbf{Q} \mathbf{F}_1 \mathbf{H}_1^j \right) \mathbf{y}_t + \beta \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} \beta^{(l+j)} tr \left(\mathbf{H}_2'^l \mathbf{F}'_1 \mathbf{Q} \mathbf{F}_1 \mathbf{H}_2^l \mathbf{\Omega} \right).$$

Again, provided the spectral radius of $\mathbf{H}_1$ is less than one, the result in Appendix A1 gives $\mathbf{R} \equiv \mathbf{F}'_1 \mathbf{Q} \mathbf{F}_1 + \beta \mathbf{H}'_1 \mathbf{R} \mathbf{H}_1$. Therefore,

$$E_t \sum_{j=1}^{\infty} \beta^j \left(\mathbf{x}'_{t+j} \mathbf{Q} \mathbf{x}_{t+j} \right) = \beta \mathbf{y}'_t \mathbf{R} \mathbf{y}_t + \frac{\beta}{1-\beta} tr \left(\mathbf{F}'_2 \mathbf{Q} \mathbf{F}_2 \mathbf{\Omega} \right) + \frac{\beta^2}{1-\beta} tr \left(\mathbf{H}'_2 \mathbf{R} \mathbf{H}_2 \mathbf{\Omega} \right). \tag{B.3}$$

Substituting B.2 and B.3 back into B.1 gives

$$Loss\,(t,\infty) = \mathbf{y}'_t \left(\mathbf{S} + \beta \mathbf{R} \right) \mathbf{y}_t + \mathbf{x}'_t \mathbf{Q} \mathbf{x}_t + \frac{\beta}{1-\beta} tr \left(\mathbf{F}'_2 \mathbf{Q} \mathbf{F}_2 \mathbf{\Omega} \right) + \frac{\beta}{1-\beta} tr \left[\mathbf{H}'_2 \left(\mathbf{S} + \beta \mathbf{R} \right) \mathbf{H}_2 \mathbf{\Omega} \right].$$

Finally, let $\mathbf{P} \equiv \mathbf{S} + \beta \mathbf{R} = \mathbf{W} + \beta \mathbf{F}'_1 \mathbf{Q} \mathbf{F}_1 + \beta \mathbf{H}'_1 \mathbf{P} \mathbf{H}_1$ giving

$$Loss\,(t,\infty) = \mathbf{y}'_t \mathbf{P} \mathbf{y}_t + \mathbf{x}'_t \mathbf{Q} \mathbf{x}_t + \frac{\beta}{1-\beta} tr \left(\mathbf{F}'_2 \mathbf{Q} \mathbf{F}_2 \mathbf{\Omega} \right) + \frac{\beta}{1-\beta} tr \left(\mathbf{H}'_2 \mathbf{P} \mathbf{H}_2 \mathbf{\Omega} \right),$$

or

$$Loss\,(t,\infty) = \mathbf{y}_t'\mathbf{P}\mathbf{y}_t + \mathbf{x}_t'\mathbf{Q}\mathbf{x}_t + \frac{\beta}{1-\beta} tr\left[\left(\mathbf{F}_2'\mathbf{Q}\mathbf{F}_2 + \mathbf{H}_2'\mathbf{P}\mathbf{H}_2\right)\mathbf{\Omega}\right],$$

as required.

B.2 Establishing Remark 5

Let $\mathbf{z}_t \equiv \left[\begin{array}{cc} \mathbf{y}_t' & \mathbf{x}_t' \end{array}\right]'$, then the policy objective function, B.1, can be written as

$$Loss\,(t,\infty) = E_t \sum_{j=0}^{\infty} \beta^j \mathbf{z}_{t+j}' \mathbf{K} \mathbf{z}_{t+j}, \tag{B.4}$$

where $\mathbf{K} \equiv \left[\begin{array}{cc} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{array}\right]$, and

$$\mathbf{z}_{t+j} = \left[\begin{array}{cc} \mathbf{H}_1 & \mathbf{0} \\ \mathbf{F}_1 & \mathbf{0} \end{array}\right] \mathbf{z}_{t+j-1} + \left[\begin{array}{c} \mathbf{H}_2 \\ \mathbf{F}_2 \end{array}\right] \mathbf{v}_{t+j} \equiv \mathbf{H}\mathbf{z}_{t+j-1} + \mathbf{G}\mathbf{v}_{t+j}, \qquad \forall j > 0. \tag{B.5}$$

Employing B.5 in B.4 gives

$$\begin{aligned} Loss\,(t,\infty) &= E_t \sum_{j=0}^{\infty} \beta^j \mathbf{z}_{t+j}' \mathbf{K} \mathbf{z}_{t+j} \\ &= \mathbf{z}_t' \left(\sum_{j=0}^{\infty} \beta^j \mathbf{H}'^j \mathbf{K} \mathbf{H}^j \right) \mathbf{z}_t + \frac{\beta}{1-\beta} \sum_{j=0}^{\infty} \beta^j tr\left(\mathbf{G}'\mathbf{H}'^j \mathbf{K} \mathbf{H}^j \mathbf{G} \mathbf{\Omega}\right). \end{aligned}$$

Now, exploiting the properties of convergent geometric series from Appendix A1, we have

$$Loss\,(t,\infty) = \mathbf{z}_t' \widetilde{\mathbf{P}} \mathbf{z}_t + \frac{\beta}{1-\beta} tr\left(\mathbf{G}'\widetilde{\mathbf{P}}\mathbf{G}\mathbf{\Omega}\right),$$

or

$$Loss\,(t,\infty) = \mathbf{z}_{t-1}' \mathbf{H}' \widetilde{\mathbf{P}} \mathbf{H} \mathbf{z}_{t-1} + \mathbf{v}_t' \mathbf{G}' \widetilde{\mathbf{P}} \mathbf{G} \mathbf{v}_t + \frac{\beta}{1-\beta} tr\left(\mathbf{G}'\widetilde{\mathbf{P}}\mathbf{G}\mathbf{\Omega}\right),$$

where $\widetilde{\mathbf{P}} \equiv \mathbf{K} + \beta \mathbf{H}' \widetilde{\mathbf{P}} \mathbf{H}$, as necessary.

B.3 Establishing Remark 6

In this appendix we establish that $\lim\limits_{\beta \uparrow 1} (1-\beta)\, Loss\,(t,\infty) = tr\,(\mathbf{K}\mathbf{\Phi})$. Recall that the policy objective function can be written as

$$Loss\,(t,\infty) = \mathbf{y}_t'\mathbf{P}\mathbf{y}_t + \mathbf{x}_t'\mathbf{Q}\mathbf{x}_t + \frac{\beta}{1-\beta} tr\left[\left(\mathbf{F}_2'\mathbf{Q}\mathbf{F}_2 + \mathbf{H}_2'\mathbf{P}\mathbf{H}_2\right)\mathbf{\Omega}\right],$$

where $\mathbf{P} \equiv \mathbf{W}+\beta\mathbf{F}_1'\mathbf{Q}\mathbf{F}_1+\beta\mathbf{H}_1'\mathbf{P}\mathbf{H}_1$. The components of the unconditional variance-covariance matrices for $\mathbf{y}_t$ and $\mathbf{x}_t$ are given by

$$\begin{aligned}
\mathbf{\Phi}_y &= \mathbf{H}_1\mathbf{\Phi}_y\mathbf{H}_1'+\mathbf{H}_2\mathbf{\Omega}\mathbf{H}_2' && \text{(B.6)}\\
\mathbf{\Phi}_x &= \mathbf{F}_1\mathbf{\Phi}_y\mathbf{F}_1'+\mathbf{F}_2\mathbf{\Omega}\mathbf{F}_2' && \text{(B.7)}\\
\mathbf{\Phi}_{yx} &= \mathbf{H}_1\mathbf{\Phi}_y\mathbf{F}_1'+\mathbf{H}_2\mathbf{\Omega}\mathbf{F}_2'\\
\mathbf{\Phi}_{xy} &= \mathbf{F}_1\mathbf{\Phi}_y\mathbf{H}_1'+\mathbf{F}_2\mathbf{\Omega}\mathbf{H}_2'.
\end{aligned}$$

Scaling the policy objective function by $(1-\beta)$ gives

$$(1-\beta)\,Loss\,(t,\infty) = (1-\beta)\left(\mathbf{y}_t'\mathbf{P}\mathbf{y}_t+\mathbf{x}_t'\mathbf{Q}\mathbf{x}_t\right)+\beta tr\left[\left(\mathbf{F}_2'\mathbf{Q}\mathbf{F}_2+\mathbf{H}_2'\mathbf{P}\mathbf{H}_2\right)\mathbf{\Omega}\right].$$

Exploiting the properties of the trace operator results in

$$(1-\beta)\,Loss\,(t,\infty) = (1-\beta)\left(\mathbf{y}_t'\mathbf{P}\mathbf{y}_t+\mathbf{x}_t'\mathbf{Q}\mathbf{x}_t\right)+\beta tr\left(\mathbf{Q}\mathbf{F}_2\mathbf{\Omega}\mathbf{F}_2'+\mathbf{P}\mathbf{H}_2\mathbf{\Omega}\mathbf{H}_2'\right).$$

Employing B.6 and B.7 gives

$$\begin{aligned}
(1-\beta)\,Loss\,(t,\infty) &= (1-\beta)\left(\mathbf{y}_t'\mathbf{P}\mathbf{y}_t+\mathbf{x}_t'\mathbf{Q}\mathbf{x}_t\right)\\
&\quad +\beta tr\left[\mathbf{Q}\left(\mathbf{\Phi}_x-\mathbf{F}_1\mathbf{\Phi}_y\mathbf{F}_1'\right)+\mathbf{P}\left(\mathbf{\Phi}_y-\mathbf{H}_1\mathbf{\Phi}_y\mathbf{H}_1'\right)\right]\\
&= (1-\beta)\left(\mathbf{y}_t'\mathbf{P}\mathbf{y}_t+\mathbf{x}_t'\mathbf{Q}\mathbf{x}_t\right)\\
&\quad +\beta tr\left[\mathbf{Q}\mathbf{\Phi}_x+\mathbf{W}\mathbf{\Phi}_y-(1-\beta)\left(\mathbf{F}_1'\mathbf{Q}\mathbf{F}_1+\mathbf{H}_1'\mathbf{P}\mathbf{H}_1\right)\right].
\end{aligned}$$

Because the spectral radius of $\mathbf{H}_1$ is less than one, $\mathbf{P}$ remains bounded even as $\beta \uparrow 1$. Thus, $\lim_{\beta\uparrow 1}(1-\beta)\left(\mathbf{y}_t'\mathbf{P}\mathbf{y}_t+\mathbf{x}_t'\mathbf{Q}\mathbf{x}_t\right)=\mathbf{0}$ and $\lim_{\beta\uparrow 1}(1-\beta)\left(\mathbf{F}_1'\mathbf{Q}\mathbf{F}_1+\mathbf{H}_1'\mathbf{P}\mathbf{H}_1\right)=\mathbf{0}$, which gives us the result that $\lim_{\beta\uparrow 1}(1-\beta)\,Loss\,(t,\infty)=tr\left(\mathbf{Q}\mathbf{\Phi}_x+\mathbf{W}\mathbf{\Phi}_y\right)=tr\left(\mathbf{K}\mathbf{\Phi}\right)$, where $\mathbf{\Phi} \equiv \begin{bmatrix} \mathbf{\Phi}_y & \mathbf{\Phi}_{yx} \\ \mathbf{\Phi}_{xy} & \mathbf{\Phi}_x \end{bmatrix}$.

References

[1] Amman, H. and D. Kendrick (1999) Linear quadratic optimization for models with rational expectations. *Macroeconomic Dynamics* 3, 534-543.

[2] Anderson, E. Hansen, L. McGrattan, E. and T. Sargent (1996) Mechanics of forming and estimating dynamic linear economies. in Amman, H. Kendrick, D. and J. Rust (eds) *Handbook of Computational Economics*, Volume 1, Chapter 4, North Holland, New York.

[3] Anderson, G. and G. Moore (1985) A linear algebraic procedure for solving linear perfect foresight models. *Economics Letters* 17, 247-252.

[4] Backus, D. and J. Driffill (1986) The consistency of optimal policy in stochastic rational expectations models. Centre for Economic Policy Research Discussion Paper #124.

[5] Binder, M. and H. Pesaran (1995) Multivariate rational expectations models and macroeconomic modeling: a review and some new results. in Pesaran, H. and M. Wickens (eds) *Handbook of Applied Econometrics*, Blackwell Publishers.

[6] Blake, A. (2004) Open loop time consistency for linear rational expectations models. *Economics Letters* 82, 21-27.

[7] Calvo, G. (1983) Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12, 383-398.

[8] Chen, B. and P. Zadrozny (2002) An alternative feedback solution for the infinite-horizon, linear-quadratic, dynamic, Stackelberg game. *Journal of Economic Dynamics and Control* 26, 1397-1416.

[9] Christiano, L. (2002) Solving dynamic equilibrium models by a method of undetermined coefficients. *Computational Economics* 20, 21-55.

[10] Clarida, R. Galí, J. and M. Gertler (1999) The science of monetary policy: a new Keynesian perspective. *Journal of Economic Literature* 37, 1661-1707.

[11] Currie, D. and P. Levine (1985) Macroeconomic policy design in an interdependent world. in Buiter, W. and R. Marston (eds) *International Economic Policy Coordination*, Cambridge University Press, Cambridge.

[12] Currie, D. and P. Levine (1993) *Rules, Reputation and Macroeconomic Policy Coordination*, Cambridge University Press.

[13] Dennis, R. (2004a) Solving for optimal simple rules in rational expectations models. *Journal of Economic Dynamics and Control* 28, 1635-1660.

[14] Dennis, R. (2004b) Inferring policy objectives from economic outcomes. *Oxford Bulletin of Economics and Statistics* 66, 735-764.

[15] Dennis, R. and U. Söderström (2005) How important is precommitment for monetary policy. *Journal of Money, Credit, and Banking*, forthcoming.

[16] Diaz-Gimenez, J. (1999) Linear quadratic approximations: an introduction. in Marimon, R. and A. Scott (eds) *Computational Methods for the Study of Dynamic Economies*, Oxford University press, Oxford, United Kingdom.

[17] Erceg, C. Henderson, D. and A. Levin (2000) Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics* 46, 281-313.

[18] Fuhrer, J. (1997) Inflation/output variance trade-offs and optimal monetary policy. *Journal of Money, Credit, and Banking* 29, 214-234.

[19] Fuhrer, J. and G. Moore (1995) Inflation persistence. *Quarterly Journal of Economics* 110, 127-159.

[20] Galí, J. and T. Monacelli (2005) Monetary policy and exchange rate volatility in a small open economy. *Review of Economic Studies*, 72, pp707-734.

[21] Giannoni, M. and M. Woodford (2002) Optimal interest-rate rules: I. General Theory. Princeton University mimeo.

[22] Holly, S. and A. Hughes-Hallett (1989) *Optimal Control, Expectations and Uncertainty*, Cambridge University press, Cambridge, United Kingdom.

[23] Jensen, C. and B. McCallum (2002) The non-optimality of proposed monetary policy rules under timeless perspective commitment. *Economics Letters* 77, 163-168.

[24] King, R. and M. Watson (2002) System reduction and solution algorithms for singular linear difference systems under rational expectations. *Computational Economics* 20, 57-86.

[25] Klein, P. (2000) Using the generalized Schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control* 24, 1405-1423.

[26] Krusell, P. Quadrini, V. and J-V. Rios-Rull (1997) Politico-economic equilibrium and economic growth. *Journal of Economic Dynamics and Control* 21, 243-272.

[27] Kydland, F. and E. Prescott (1977) Rules rather than discretion: the inconsistency of optimal plans. *Journal of Political Economy* 87, 473-492.

[28] Kydland, F. and E. Prescott (1980) Dynamic optimal taxation, rational expectations and optimal control. *Journal of Economic Dynamics and Control* 2, 79-91.

[29] McCallum, B. (1999) Solutions to linear rational expectations models: a compact exposition. *Economics Letters* 61, 143-147.

[30] Orphanides, A. and V. Wieland (1998) Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero. Board of Governors of the Federal Reserve, Finance and Economics Discussion Paper#1998-35.

[31] Oudiz, G. and J. Sachs (1985) International policy coordination in dynamic macroeconomic models. in Buiter, W. and R. Marston (eds) *International Economic Policy Coordination*, Cambridge University Press, Cambridge, 275-319.

[32] Rustem, B. and M. Zarrop (1979) A Newton-type method for the optimization and control of non-linear econometric models. *Journal of Economic Dynamics and Control* 1, 283-300.

[33] Sims, C. (2002) Solving linear rational expectations models. *Computational Economics* 20, 1-20.

[34] Söderlind, P. (1999) Solution and estimation of RE macromodels with optimal policy. *European Economic Review* 43, 813-823.

[35] Svensson, L. (2000) Open-economy inflation targeting. *Journal of International Economics* 50, 155-183.

[36] Svensson, L. (2003) What is wrong with Taylor rules? Using judgement in monetary policy through targeting rules. *Journal of Economic Literature* June, 426-477.

[37] Taylor, J. (1999) *Monetary Policy Rules*. University of Chicago Press, Chicago.

[38] Uhlig, H. (1999) A toolkit for analysing nonlinear dynamic stochastic models easily. in Marimon, R. and A. Scott (eds) *Computational Methods for the Study of Dynamic Economies*, Oxford University Press, New York.

[39] Woodford, M. (1999) Commentary: How should monetary policy be conducted in an era of price stability? Presented at the Symposium on *New Challenges for Monetary Policy*, organized by the Federal Reserve Bank of Kansas City at Jackson Hole, Wyoming, August 26-28.

[40] Woodford, M. (2002) *Interest & Prices*. Princeton University Press, Princeton, New Jersey.