Wohltmann, Hans-Werner; Winkler, Roland C.

Working Paper

Rational expectations models with anticipated shocks and optimal policy: a general solution method and a new Keynesian example

Economics working paper / Christian-Albrechts-Universität Kiel, Department of Economics, No. 2009,01

Provided in cooperation with:
Christian-Albrechts-Universität Kiel (CAU)

Christian-Albrechts-Universität zu Kiel

Department of Economics

Economics Working Paper
No 2009-01

rational expectations models with anticipated shocks and optimal policy: a general solution method and a new keynesian example

by Hans-Werner Wohltmann and Roland Winkler

Rational Expectations Models with Anticipated Shocks and Optimal Policy: A General Solution Method and a New Keynesian Example*

Hans-Werner Wohltmann‡ and Roland Winkler§

Department of Economics, Christian-Albrechts-University of Kiel, Olshausenstr. 40, D-24098 Kiel, Germany

February 4, 2009

Abstract

The purpose of this paper is to show how to solve linear dynamic rational expectations models with anticipated shocks by using the generalized Schur decomposition method. Furthermore, we determine the optimal unrestricted and restricted policy responses to anticipated shocks. We demonstrate our solution method by means of a micro-founded hybrid New Keynesian model and show that anticipated cost-push shocks entail higher welfare losses than unanticipated shocks of equal size.

JEL classification: C61, C63, E52
Keywords: Anticipated Shocks, Optimal Monetary Policy, Rational Expectations, Generalized Schur Decomposition, Welfare Effects

*This paper is a revised and significantly extended version of our paper "Solution of RE Models with Anticipated Shocks and Optimal Policy", Economics Working Paper 2007-32, Department of Economics, Christian-Albrechts-Universität Kiel.

‡Corresponding author: Phone: ++49-431-880-1449, Fax: ++49-431-880-2228, E-mail: wohltmann@economics.uni-kiel.de

§E-mail: winkler@economics.uni-kiel.de

1 Introduction

Recently, a number of macroeconometric studies emphasize the role of anticipated shocks as sources of macroeconomic fluctuations. Beaudry and Portier (2006) find that more than one half of business cycle fluctuations are caused by news concerning future technological opportunities. Davis (2007) and Fujiwara, Hirose, and Shintani (2008) analyze the importance of anticipated shocks in large scale DSGE models and report that these disturbances are important components of aggregate fluctuations. Schmitt-Grohé and Uribe (2008) conduct a Bayesian estimation of a real-business cycle model and find that anticipated shocks are the most important source of aggregate fluctuations. In particular, they report that anticipated shocks explain two thirds of the volatility in consumption, output, investment, and employment.

In light of these findings, Wohltmann and Winkler (2008) investigate, whether the anticipation of future cost-push shocks has a stabilizing effect on the economy and thus reduces the welfare loss compared to unanticipated shocks. In order to provide analytical results which do not rely on calibrations, they consider the baseline New Keynesian model with purely forward-looking IS and Phillips curves. This enables them to derive an analytical solution of welfare as a function of the time span between the anticipation and the realization of the shock. They find that – for empirically plausible degrees of nominal rigidity – the anticipation of a future cost-push shock leads to a higher welfare loss than an analogous unanticipated shock.

In order to conduct an analysis of the (welfare) effects of anticipated shocks in more elaborate models, this paper presents a general solution method for linear dynamic rational expectations models with anticipated shocks and optimal policy. Our method extends the work of Söderlind (1999), who uses the generalized Schur decomposition method, advocated by Klein (2000), to solve linear rational expectations models with optimal policy. However, Söderlind (1999) only considers stochastic models with white noise shocks which are, by definition, unpredictable. In the case of anticipated shocks, the occurrence of all future shocks is known exactly at the time when the solution of the model is computed. Our method also contains unanticipated shocks as a limiting case.

As an economic example, we lay out a calibrated New Keynesian model for a closed and cashless economy with internal habit formation in consumption preferences, a variant of Calvo price staggering with partial indexation to past inflation and a time-varying wage mark-up which represents a typical cost-push shock. We compare the effects of mark-up shocks under optimal monetary policy for different lengths of the anticipation period. Our results confirm the finding of Wohltmann and Winkler (2008) who show that anticipated shocks entail higher welfare losses than unexpected cost shocks.

The paper is organized as follows. Section 2 discusses optimal policies in RE models with *anticipated* temporary shocks. We first determine the optimal unrestricted policy under precommitment and calculate the minimum value of the intertemporal loss function. We then consider (optimal) simple rules and demonstrate how the Schur decomposition can be used to solve the model under these conditions. Section 3 derives the hybrid New Keynesian model, presents the welfare-theoretic loss function and discusses the effects of anticipated and unanticipated cost-push shocks. Finally, section 4 provides concluding remarks. In the Appendix, we present a short discussion of the well known stochastic case with white noise shocks.

2 The Model

In this paper we discuss the following linear expectational difference equations

$$A \begin{pmatrix} w_{t+1} \\ \mathrm{E}_t\, v_{t+1} \end{pmatrix} = B \begin{pmatrix} w_t \\ v_t \end{pmatrix} + Cu_t + D\nu_{t+1} \tag{1}$$

where w_t is an $n_1 \times 1$ vector of predetermined variables, assuming w_0 given, v_t an $n_2 \times 1$ vector of non-predetermined variables, u_t an $m \times 1$ vector of policy instruments, and ν_{t+1} an $r \times 1$ vector of exogenous shocks. The matrices A and B are $n \times n$ (where $n = n_1 + n_2$), while the matrices C and D are $n \times m$ and $n \times r$ respectively. We allow matrix A to be singular which is the case if static (intratemporal) equations are included among the dynamic relationships. The vector w, composed of backward-looking variables can include exogenous variables following autoregressive processes. $\mathrm{E}_t\, v_{t+1}$ denotes model consistent (rational) expectations of v_{t+1} formed at time t. We assume that the shocks are anticipated by the public in advance and take the following form

$$\nu_t = \begin{cases} \overline{\nu} & \text{for} \quad t = \tau > 0 \\ 0 & \text{for} \quad t \neq \tau \end{cases} \tag{2}$$

where $\overline{\nu} = (\overline{\nu}_1, \dots, \overline{\nu_r})'$ is a constant non-zero $r \times 1$ vector. It is assumed that at time $t = 0$ the public anticipates a shock of the form outlined in (2) to take place at some future date $\tau > 0$. Note that τ also defines the lengths of the anticipation period. Since the shocks are anticipated by the public we have $\mathrm{E}_t\, \nu_{t+1} = \nu_{t+1}$. For notational convenience, we define the $n \times 1$ vector $k_t = (w_t', v_t')'$ and the $n_3 \times 1$ target vector $s_t = \tilde{A}k_t + \tilde{B}u_t$, where the matrices $\tilde{A}$ and $\tilde{B}$ are $n_3 \times n$ and $n_3 \times m$ respectively. Assume that the policy maker´s welfare loss at time t is given by

$$J_t = \frac{1}{2}\,\mathrm{E}_t \sum_{i=0}^{\infty} \lambda^i \{ s_{t+i}' W_1 s_{t+i} + u_{t+i}' W_2 u_{t+i} \} \tag{3}$$

where W_1 and W_2 are symmetric and non-negative definite matrices and λ is a discount factor with $0 < \lambda \leq 1$. We can rewrite J_t as

$$J_t = \frac{1}{2}\,\mathrm{E}_t \sum_{i=0}^{\infty} \lambda^i \{ k_{t+i}' \tilde{W} k_{t+i} + 2k_{t+i}' P u_{t+i} + u_{t+i}' R u_{t+i} \} \tag{4}$$

where $\tilde{W} = \tilde{A}' W_1 \tilde{A}$ and $R = W_2 + \tilde{B}' W_1 \tilde{B}$ are symmetric and non-negative definite and $P = \tilde{A}' W_1 \tilde{B}$.

2.1 Optimal Policy with Precommitment

In the following, the policy maker´s optimal policy rule at time $t = 0$ is developed. It is assumed that the policy maker is able to commit to such a rule. From the Lagrangian

$$\mathcal{L}_0 = \frac{1}{2} E_0 \sum_{t=0}^{\infty} \lambda^t \{ k_t' \tilde{W} k_t + 2 k_t' P u_t + u_t' R u_t \\ + 2\rho_{t+1}' [B k_t + C u_t + D \nu_{t+1} - A k_{t+1}] \} \tag{5}$$

with the $n \times 1$ multiplier ρ_{t+1}, we get the first-order conditions with respect to ρ_{t+1}, k_t, and u_t:

$$\begin{pmatrix} A & 0_{n\times m} & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times m} & \lambda B' \\ 0_{m\times n} & 0_{m\times m} & -C' \end{pmatrix} \begin{pmatrix} k_{t+1} \\ u_{t+1} \\ \rho_{t+1} \end{pmatrix} \\ = \begin{pmatrix} B & C & 0_{n\times n} \\ -\lambda \tilde{W} & -\lambda P & A' \\ P' & R & 0_{m\times n} \end{pmatrix} \begin{pmatrix} k_t \\ u_t \\ \rho_t \end{pmatrix} + \begin{pmatrix} D \\ 0_{n\times r} \\ 0_{m\times r} \end{pmatrix} \nu_{t+1} \tag{6}$$

To solve the system of equations (6), expand the state and costate vector k_t and ρ_t as $(w_t', v_t')'$ and $(p_{wt}', p_{vt}')'$ respectively and re-order the rows of the $(2n+m)\times 1$ vector $(k_t', u_t', \rho_t')'$ by placing the predetermined vector p_{vt} after w_t. Since v_t is forward-looking with arbitrarily chosen initial value v_0, the corresponding Lagrange multiplier p_{vt} is predetermined with initial value $p_{v0} = 0$. Re-order the columns of the $(2n + m) \times (2n + m)$ matrices in (6) according to the re-ordering of $(k_t', u_t', \rho_t)'$ and write the result as

$$F \begin{pmatrix} \tilde{w}_{t+1} \\ \tilde{v}_{t+1} \end{pmatrix} = G \begin{pmatrix} \tilde{w}_t \\ \tilde{v}_t \end{pmatrix} + \begin{pmatrix} D \\ 0_{n\times r} \\ 0_{m\times r} \end{pmatrix} \nu_{t+1} \tag{7}$$

where $\tilde{w}_t = (w_t', p_{vt}')'$ and $\tilde{v}_t = (v_t', u_t', p_{wt}')'$. The $n \times 1$ vector $\tilde{w}_t$ contains the 'backward-looking' variables of (6) while the $(n + m) \times 1$ vector $\tilde{v}_t$ contains the 'forward-looking' variables.

Equation (6) implies that the $(2n + m) \times (2n + m)$ matrix F is singular. To solve equation (7) we apply the generalized Schur decomposition method (Söderlind, 1999; Klein, 2000). The decomposition of the square matrices F and G is given by

$$F = \overline{Q}' S \overline{Z}', \quad G = \overline{Q}' T \overline{Z}' \tag{8}$$

or equivalently

$$QFZ = S, \quad QGZ = T \tag{9}$$

where Q, Z, S, and T are square matrices of complex numbers, S and T are upper triangular and Q and Z are unitary, i.e.

$$Q \cdot \overline{Q}' = \overline{Q}' \cdot Q = I_{(2n+m)\times(2n+m)} = Z \cdot \overline{Z}' = \overline{Z}' \cdot Z \tag{10}$$

where the non-singular matrix $\overline{Q}'$ is the transpose of $\overline{Q}$, which denotes the complex conjugate of Q. $\overline{Z}'$ is the transpose of the complex conjugate of Z. The matrices S and T can be arranged in such a way that the block with the stable generalized eigenvalues (the ith diagonal element of T divided by the ith diagonal element of S) comes first. Premultiply both sides of equation (7) with Q and define auxiliary variables $\tilde{z}_t$ and $\tilde{x}_t$ so that

$$\begin{pmatrix} \tilde{z}_t \\ \tilde{x}_t \end{pmatrix} = \overline{Z}' \begin{pmatrix} \tilde{w}_t \\ \tilde{v}_t \end{pmatrix} \tag{11}$$

Partitioning the triangular matrices S and T in order to conform with $\tilde{z}$ and $\tilde{x}$ and set

$$Q \begin{pmatrix} D \\ 0_{n\times r} \\ 0_{m\times r} \end{pmatrix} = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \tag{12}$$

where Q_1 is $n \times r$ and Q_2 is $(n+m)\times r$. We then obtain the equivalent system

$$\begin{pmatrix} S_{11} & S_{12} \\ 0_{(n+m)\times n} & S_{22} \end{pmatrix} \begin{pmatrix} \tilde{z}_{t+1} \\ \tilde{x}_{t+1} \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ 0_{(n+m)\times n} & T_{22} \end{pmatrix} \begin{pmatrix} \tilde{z}_t \\ \tilde{x}_t \end{pmatrix} + \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \nu_{t+1} \tag{13}$$

where the $n\times n$ matrix S_{11} and the $(n+m)\times(n+m)$ matrix T_{22} are invertible while S_{22} is singular. The square matrix T_{11} may also be singular. The lower block of equation (13) contains the unstable generalized eigenvalues and must be solved forward. Since

$$\tilde{x}_{t+s} = M_2 \tilde{x}_{t+s+1} - T_{22}^{-1} Q_2 \nu_{t+s+1} \quad (s = 0, 1, 2, \ldots) \tag{14}$$

where $M_2 = T_{22}^{-1} S_{22}$, the unique stable solution for $\tilde{x}_t$ is given by

$$\begin{aligned} \tilde{x}_t &= -\sum_{s=0}^{\infty} M_2^s T_{22}^{-1} Q_2 \,\mathrm{E}_t\, \nu_{t+s+1} \\ &= \begin{cases} -M_2^{\tau-1-t} T_{22}^{-1} Q_2 \overline{\nu} & \text{for} \quad 0 \le t < \tau \\ 0 & \text{for} \quad t \ge \tau \end{cases} \end{aligned} \tag{15}$$

The upper block of (13) contains the stable generalized eigenvalues and can be solved backward. Since

$$\tilde{z}_{t+1} = M_1 \tilde{z}_t + S_{11}^{-1}(T_{12}\tilde{x}_t - S_{12}\tilde{x}_{t+1}) + S_{11}^{-1} Q_1 \nu_{t+1} \tag{16}$$

where $M_1 = S_{11}^{-1} T_{11}$ (which in general is not invertible), the general solution is given by

$$\begin{aligned} \tilde{z}_t &= M_1^t K + \sum_{s=0}^{t-1} M_1^{t-s-1} S_{11}^{-1} (T_{12}\tilde{x}_s - S_{12}\tilde{x}_{s+1} + Q_1 \nu_{s+1}) \\ &= \begin{cases} M_1^t K + \sum_{s=0}^{t-1} M_1^{t-s-1} S_{11}^{-1}(T_{12}\tilde{x}_s - S_{12}\tilde{x}_{s+1}) & \text{for} \quad 0 \le t < \tau \\ \\ M_1^t K + \sum_{s=0}^{\tau-1} M_1^{t-s-1} S_{11}^{-1}(T_{12}\tilde{x}_s - S_{12}\tilde{x}_{s+1}) & \\ \quad + M_1^{t-\tau} S_{11}^{-1} Q_1 \overline{\nu} & \text{for} \quad t \ge \tau \end{cases} \end{aligned} \tag{17}$$

where $\tilde{x}_s$ is defined in (15).

The solution for $t \geq \tau$ can be rewritten as

$$\tilde{z}_t = M_1^{t-\tau} \tilde{K} \quad \text{for } \ t \geq \tau \tag{18}$$

where

$$\tilde{K} = M_1^{\tau} K + S_{11}^{-1} Q_1 \overline{\nu} + \sum_{s=0}^{\tau-1} M_1^{\tau-s-1} S_{11}^{-1} (T_{12} \tilde{x}_s - S_{12} \tilde{x}_{s+1}) \tag{19}$$

Since

$$\tilde{x}_s = \begin{cases} -M_2^{\tau-1-s} T_{22}^{-1} Q_2 \overline{\nu} & \text{for } 0 \leq s < \tau \\ 0 & \text{for } s \geq \tau \end{cases} \tag{20}$$

we can write $\tilde{K}$ as

$$\tilde{K} = M_1^{\tau} K + S_{11}^{-1} Q_1 \overline{\nu} + [-\tilde{W}_1 + M_1 \tilde{W}_2] T_{22}^{-1} Q_2 \overline{\nu} \tag{21}$$

where

$$\tilde{W}_1 = \sum_{s=0}^{\tau-1} M_1^{\tau-s-1} S_{11}^{-1} T_{12} M_2^{\tau-s-1} = \sum_{k=0}^{\tau-1} M_1^k S_{11}^{-1} T_{12} M_2^k \tag{22}$$

and

$$\tilde{W}_2 = \sum_{s=0}^{\tau-2} M_1^{\tau-s-2} S_{11}^{-1} S_{12} M_2^{\tau-s-2} = \sum_{k=0}^{\tau-2} M_1^k S_{11}^{-1} S_{12} M_2^k \tag{23}$$

$\tilde{W}_1$ as well as $\tilde{W}_2$ is a finite geometric sum of matrices and can be written as

$$\tilde{W}_1 = S_{11}^{-1} T_{12} - M_1^{\tau} S_{11}^{-1} T_{12} M_2^{\tau} + M_1 \tilde{W}_1 M_2 \tag{24}$$

and

$$\tilde{W}_2 = S_{11}^{-1} S_{12} - M_1^{\tau-1} S_{11}^{-1} S_{12} M_2^{\tau-1} + M_1 \tilde{W}_2 M_2 \tag{25}$$

To solve for $\tilde{W}_1$ and $\tilde{W}_2$ respectively, we use the matrix identities (Rudebusch and Svensson 1999; Klein, 2000) $\text{vec}\,(A + B) = \text{vec}\,(A) + \text{vec}\,(B)$ and $\text{vec}\,(ABC) = [C' \otimes A]\,\text{vec}\,(B)$ where $\text{vec}\,(A)$ denotes the vector of stacked column vectors of the matrix A and $\otimes$ denotes the Kronecker product of matrices.

We then obtain from (24) and (25)

$$\text{vec}\, \tilde{W}_1 - [M_2' \otimes M_1]\, \text{vec}\, \tilde{W}_1 = \text{vec}\, [S_{11}^{-1} T_{12} - M_1^{\tau} S_{11}^{-1} T_{12} M_2^{\tau}] \tag{26}$$

and

$$\text{vec}\, \tilde{W}_2 - [M_2' \otimes M_1]\, \text{vec}\, \tilde{W}_2 = \text{vec}\, [S_{11}^{-1} S_{12} - M_1^{\tau} S_{11}^{-1} S_{12} M_2^{\tau}] \tag{27}$$

with the solution

$$\text{vec}\,\tilde{W}_1 = [I - M_2' \otimes M_1]^{-1} \cdot \text{vec}\,[S_{11}^{-1} T_{12} - M_1^{\tau} S_{11}^{-1} T_{12} M_2^{\tau}] \tag{28}$$

$$\text{vec}\,\tilde{W}_2 = [I - M_2' \otimes M_1]^{-1} \cdot \text{vec}\,[S_{11}^{-1} S_{12} - M_1^{\tau-1} S_{11}^{-1} S_{12} M_2^{\tau-1}] \tag{29}$$

According to (17) and (20), the solution of $\tilde{z}_t$ for the anticipation period $0 < t < \tau$ can be rewritten as

$$\tilde{z}_t = M_1^t K + [-W_{1t}^* + W_{2t}^*]\, T_{22}^{-1} Q_2 \overline{\nu} \qquad \text{for } 0 \le t < \tau \tag{30}$$

with

$$W_{1t}^* = \sum_{s=0}^{t-1} M_1^{t-s-1} S_{11}^{-1} T_{12} M_2^{\tau-s-1} = \sum_{k=0}^{t-1} M_1^k S_{11}^{-1} T_{12} M_2^{\tau-t+k} \tag{31}$$

and

$$W_{2t}^* = \sum_{s=1}^{t} M_1^{t-s} S_{11}^{-1} S_{12} M_2^{\tau-s-1} = \sum_{k=0}^{t-1} M_1^k S_{11}^{-1} S_{12} M_2^{\tau-1-t+k} \tag{32}$$

W_{1t}^* satisfies the matrix equation[2]

$$W_{1t}^* = S_{11}^{-1} T_{12} M_2^{\tau-t} - M_1^t S_{11}^{-1} T_{12} M_2^{\tau} + M_1 W_{1t}^* M_2 \quad (0 \le t < \tau) \tag{33}$$

with the solution

$$\text{vec}\,W_{1t}^* = [I - M_2' \otimes M_1]^{-1} \cdot \text{vec}\,(S_{11}^{-1} T_{12} M_2^{\tau-t} - M_1^t S_{11}^{-1} T_{12} M_2^{\tau}) \tag{34}$$

The matrix W_{2t}^* satisfies the equation[3]

$$W_{2t}^* = S_{11}^{-1} S_{12} M_2^{\tau-1-t} - M_1^t S_{11}^{-1} S_{12} M_2^{\tau-1} + M_1 W_{2t}^* M_2 \quad (0 \le t < \tau) \tag{35}$$

with the solution

$$\text{vec}\,W_{2t}^* = [I - M_2' \otimes M_1]^{-1} \cdot \text{vec}\,(S_{11}^{-1} S_{12} M_2^{\tau-1-t} - M_1^t S_{11}^{-1} S_{12} M_2^{\tau-1}) \tag{36}$$

The constant K can be determined using the initial value of the predetermined vector $\tilde{w}$. By premultiplying equation (11) with Z and by partitioning the matrix Z to conform with the dimension of $\tilde{z}$ and $\tilde{x}$, we obtain

$$\begin{pmatrix} \tilde{w}_t \\ \tilde{v}_t \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \begin{pmatrix} \tilde{z}_t \\ \tilde{x}_t \end{pmatrix} \tag{37}$$

[2]Note that equation (33) is also well-defined for $t = \tau$. In this case it is equivalent to (24) implying $W_{1\tau}^* = \tilde{W}_1$.

[3]For $t = \tau - 1$ equation (35) is equivalent to (25) so that $W_{2\tau-1}^* = \tilde{W}_2$. Then, according to (21), $\tilde{K} = \tilde{z}_\tau = M_1^\tau K + S_{11}^{-1} Q_1 \overline{\nu} + [-W_{1\tau}^* + M_1 W_{2\tau-1}^*] T_{22}^{-1} Q_2 \overline{\nu}$.

The definition of W_{1t}^* implies that W_{1t}^* also satisfies the dynamic equation

$$W_{1t+1}^* = S_{11}^{-1} T_{12} M_2^{\tau-(t+1)} + M_1 W_{1t}^* \quad (0 \le t \le \tau - 1)$$

with the initial value $W_{1\,0}^* = 0$. Analogical, W_{2t}^* satisfies the matrix difference equation

$$W_{2t+1}^* = S_{11}^{-1} S_{12} M_2^{\tau-t-2} + M_1 W_{2t}^* \quad (W_{2\,0}^* = 0)$$

which only holds for $0 \le t < \tau - 1$ since $M_2^{\tau-(\tau-1)-2} = M_2^{-1}$ generally does not exist.

and therefore

$$\tilde{w}_0 = Z_{11}\tilde{z}_0 + Z_{12}\tilde{x}_0 \tag{38}$$

with $\tilde{w}_0 = (w_0', 0_{n_2\times 1}')'$, $\tilde{z}_0 = K$, and

$$\tilde{x}_0 = -M_2^{\tau-1}T_{22}^{-1}Q_2\overline{\nu} \tag{39}$$

where it is assumed that $\tau > 0$.[4] Equation (38) implies

$$K = Z_{11}^{-1}\tilde{w}_0 - Z_{11}^{-1}Z_{12}\tilde{x}_0 \tag{40}$$

provided the inverse Z_{11}^{-1} exists. A necessary condition is that the dynamic system (7) has the saddle path property, i.e., that the number of backward-looking variables ($n_1 + n_2 = n$) coincides with the number of stable generalized eigenvalues (Söderlind, 1999; Klein, 2000].

In the case $\tau > 0$ we can assume $w_0 = 0$ so that according to (39) the constant K can be written as

$$K = Z_{11}^{-1}Z_{12}M_2^{\tau-1}T_{22}^{-1}Q_2\overline{\nu} \tag{41}$$

The solution to the state vector $(\tilde{z}_t, \tilde{x}_t)'$ for $0 \le t < \tau$ now reads as follows

$$\begin{pmatrix} \tilde{z}_t \\ \tilde{x}_t \end{pmatrix} = \Xi_t T_{22}^{-1}Q_2\overline{\nu} \qquad \text{for } \ 0 \le t < \tau \tag{42}$$

where

$$\Xi_t = \begin{pmatrix} \phi_t^* \\ -M_2^{\tau-1-t} \end{pmatrix} \qquad (0 \le t < \tau) \tag{43}$$

and[5]

$$\phi_t^* = M_1^t Z_{11}^{-1}Z_{12}M_2^{\tau-1} - W_{1t}^* + W_{2t}^* \tag{44}$$

If Z_{11} is invertible, equation (37) implies

$$\tilde{v}_t = Z_{21}\tilde{z}_t + Z_{22}\tilde{x}_t = Z_{21}(Z_{11}^{-1}\tilde{w}_t - Z_{11}^{-1}Z_{12}\tilde{x}_t) + Z_{22}\tilde{x}_t = N\tilde{w}_t + \hat{Z}\tilde{x}_t \tag{45}$$

where $N = Z_{21}Z_{11}^{-1}$ and $\hat{Z} = Z_{22} - Z_{21}Z_{11}^{-1}Z_{12}$. Write equation (45) as

$$\begin{pmatrix} v_t \\ u_t \\ p_{w\,t} \end{pmatrix} = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \\ N_{31} & N_{32} \end{pmatrix} \begin{pmatrix} w_t \\ p_{v\,t} \end{pmatrix} + \begin{pmatrix} \hat{Z}_1 \\ \hat{Z}_2 \\ \hat{Z}_3 \end{pmatrix} \tilde{x}_t \tag{46}$$

[4] In the special case $\tau = 0$ (unanticipated shocks) we have $\tilde{x}_0 = 0$ and $\tilde{z}_t = (S_{11}^{-1}T_{11})^t K + (S_{11}^{-1}T_{11})^t S_{11}^{-1}Q_1\overline{\nu}$ implying $\tilde{z}_0 = K + S_{11}^{-1}Q_1\overline{\nu}$ and $K = Z_{11}^{-1}\tilde{w}_0 - S_{11}^{-1}Q_1\overline{\nu}$ with $w_0 \neq 0$. By contrast, the initial value w_0 can be normalized to zero if $\tau > 0$.

[5] ϕ_t^* satisfies the dynamic equation

$$\phi_{t+1}^* = M_1\phi_t^* + S_{11}^{-1}[-T_{12}M_2 + S_{12}]M_2^{\tau-t-2}$$

where the time index t must be restricted to $0 \le t < \tau - 1$.

and assume the $n_2 \times n_2$ matrix N_{12} is invertible. The optimal policy rule under commitment can then be written as

$$
\begin{aligned}
u_t &= N_{21} w_t + N_{22} p_{v\,t} + \hat{Z}_2 \tilde{x}_t \\
&= N_{21} w_t + N_{22} N_{12}^{-1} (v_t - N_{11} w_t - \hat{Z}_1 \tilde{x}_t) + \hat{Z}_2 \tilde{x}_t \\
&= N_{22} N_{12}^{-1} v_t + (N_{21} - N_{22} N_{12}^{-1} N_{11}) w_t + (\hat{Z}_2 - N_{22} N_{12}^{-1} \hat{Z}_1) \tilde{x}_t
\end{aligned} \tag{47}
$$

where $\tilde{x}_t$ is given by (15). For $t < \tau$, u_t depends on the auxiliary variable $\tilde{x}_t$, while for $t \geq \tau$, u_t is only a linear function of the predetermined state variables w_t and p_{vt}, where p_{vt} can be substituted with the original state variables v_t and w_t.

Minimum Value of the Loss Function

To determine the minimum value of the loss function J_t at time $t = 0$, we express J_t as function of $\tilde{w}$ and $\tilde{v}$. The loss function (4) can be written as

$$
J_t = \frac{1}{2} \sum_{i=0}^{\infty} \lambda^i (k'_{t+i}, u'_{t+i}) H \begin{pmatrix} k_{t+i} \\ u_{t+i} \end{pmatrix} = \frac{1}{2} \sum_{i=0}^{\infty} \lambda^i (w'_{t+i}, v'_{t+i}, u'_{t+i}) H \begin{pmatrix} w_{t+i} \\ v_{t+i} \\ u_{t+i} \end{pmatrix} \tag{48}
$$

where the $(n+m) \times (n+m)$ matrix H is given by

$$
H = \begin{pmatrix} \tilde{W} & P \\ P' & R \end{pmatrix} \tag{49}
$$

with $H = H'$. Define the $n_1 \times n$ matrix $\tilde{D}_1$ and the $(n_2+m) \times (n+m)$ matrix $\tilde{D}_2$ by $\tilde{D}_1 = (I_{n_1 \times n_1}, 0_{n_1 \times n_2})$ and $\tilde{D}_2 = (I_{(n_2+m)\times(n_2+m)}, 0_{(n_2+m)\times n_1})$, respectively. Then $w = \tilde{D}_1 (w', p'_v)' = \tilde{D}_1 \tilde{w}'$, $(v', u')' = \tilde{D}_2 (v', u', p'_w)' = \tilde{D}_2 \tilde{v}'$, $(w', v', u')' = \tilde{D}(\tilde{w}', \tilde{v}')'$ with

$$
\begin{aligned}
\tilde{D} &= \begin{pmatrix} \tilde{D}_1 & 0_{n_1 \times (n+m)} \\ 0_{(n_2+m)\times n} & \tilde{D}_2 \end{pmatrix} \\
&= \begin{pmatrix} I_{n_1 \times n_1} & 0_{n_1 \times n_2} & 0_{n_1 \times (n_2+m)} & 0_{n_1 \times n_1} \\ 0_{(n_2+m)\times n_1} & 0_{(n+m)\times n_2} & I_{(n_2+m)\times(n_2+m)} & 0_{(n_2+m)\times n_1} \end{pmatrix}
\end{aligned} \tag{50}
$$

which is a $(n+m) \times (2n+m)$ matrix. The loss function J_t can now be rewritten as

$$
J_t = \frac{1}{2} \sum_{i=0}^{\infty} \lambda^i (\tilde{w}'_{t+i}, \tilde{v}'_{t+i}) \tilde{D}' H \tilde{D} \begin{pmatrix} \tilde{w}_{t+i} \\ \tilde{v}_{t+i} \end{pmatrix} = J_t^{(1)} + J_t^{(2)} \tag{51}
$$

where

$$
J_t^{(1)} = \frac{1}{2} \sum_{i=0}^{\tau-1} \lambda^i (\tilde{w}'_{t+i}, \tilde{v}'_{t+i}) \tilde{D}' H \tilde{D} \begin{pmatrix} \tilde{w}_{t+i} \\ \tilde{v}_{t+i} \end{pmatrix} \tag{52}
$$

and

$$
J_t^{(2)} = \frac{1}{2} \sum_{i=\tau}^{\infty} \lambda^i (\tilde{w}'_{t+i}, \tilde{v}'_{t+i}) \tilde{D}' H \tilde{D} \begin{pmatrix} \tilde{w}_{t+i} \\ \tilde{v}_{t+i} \end{pmatrix} \tag{53}
$$

First, we calculate $J_t^{(2)}$. For $t \geq \tau$, we have $\tilde{v}_t = N\tilde{w}_t$ and $\tilde{w}_t = Z_{11}\tilde{z}_t$, where $N = Z_{21}Z_{11}^{-1}$. We then obtain $(\tilde{w}_t', \tilde{v}_t')' = \tilde{N}\tilde{w}_t = \tilde{N}Z_{11}\tilde{z}_t$, where $\tilde{N} = (I_{n\times n}, N')'$ is a $(2n+m)\times n$ matrix. $J_t^{(2)}$ can then be rewritten as

$$J_t^{(2)} = \frac{1}{2}\sum_{i=\tau}^{\infty}\lambda^i Z_{11}'\tilde{z}_{t+i}'\tilde{N}'\tilde{D}'H\tilde{D}\tilde{N}Z_{11}\tilde{z}_{t+i} = \frac{1}{2}\sum_{i=\tau}^{\infty}\lambda^i Z_{11}'\tilde{z}_{t+i}'H^*Z_{11}\tilde{z}_{t+i} \tag{54}$$

with $H^* = \tilde{N}'\tilde{D}'H\tilde{D}\tilde{N}$ is a symmetric $n\times n$ matrix. Inserting (18) in (54) we obtain

$$\begin{aligned} J_t^{(2)} &= \frac{1}{2}(M_1^t\tilde{K})'\lambda^\tau\left(\sum_{i=\tau}^{\infty}\lambda^{i-\tau}(Z_{11}M_1^{i-\tau})'H^*(Z_{11}M_1^{i-\tau})\right)M_1^t\tilde{K} \\ &= \frac{1}{2}\lambda^\tau\varphi_t'V^*\varphi_t = \frac{1}{2}\lambda^\tau\,\mathrm{trace}(V^*\varphi_t\varphi_t') \end{aligned} \tag{55}$$

where $\varphi_t = M_1^t\tilde{K}$ and V^* is the convergent geometric sum of matrices

$$V^* = \sum_{i=\tau}^{\infty}\lambda^{i-\tau}(Z_{11}M_1^{i-\tau})'H^*(Z_{11}M_1^{i-\tau}) \tag{56}$$

which is of dimension $n\times n$ and satisfies the matrix equation

$$V^* = Z_{11}'H^*Z_{11} + \lambda M_1'V^*M_1 \tag{57}$$

with the solution

$$\mathrm{vec}\,(V^*) = [I - \lambda M_1'\otimes M_1]^{-1}\,\mathrm{vec}\,(Z_{11}'H^*Z_{11}) \tag{58}$$

For $t = 0$ we obtain from (55)

$$J_0^{(2)} = \frac{1}{2}\lambda^\tau\,\mathrm{trace}(V^*\varphi_0\varphi_0') = \frac{1}{2}\lambda^\tau\,\mathrm{trace}(V^*\tilde{K}\tilde{K}') \tag{59}$$

with $\tilde{K}$ given by (21).

The next step is the calculation of the finite sum $J_t^{(1)}$ as defined in (52). Because $(\tilde{w}_t', \tilde{v}_t')' = Z(\tilde{z}_t', \tilde{x}_t')'$, we can write $J_0^{(1)}$ as

$$J_0^{(1)} = \frac{1}{2}\sum_{i=0}^{\tau-1}\lambda^i(\tilde{z}_i', \tilde{x}_i')Z'\tilde{D}'H\tilde{D}Z\begin{pmatrix}\tilde{z}_i\\ \tilde{x}_i\end{pmatrix} = \frac{1}{2}\sum_{t=0}^{\tau-1}\lambda^t(\tilde{z}_t', \tilde{x}_t')\tilde{H}\begin{pmatrix}\tilde{z}_t\\ \tilde{x}_t\end{pmatrix} \tag{60}$$

where $\tilde{H} = Z'\tilde{D}'H\tilde{D}Z$.

Inserting the solution formulas for $\tilde{z}_t$ and $\tilde{x}_t$ in (60), we obtain the expression

$$\begin{aligned} J_0^{(1)} &= \frac{1}{2}(T_{22}^{-1}Q_2\overline{\nu})'\left[\sum_{t=0}^{\tau-1}\lambda^t\Xi_t'\tilde{H}\Xi_t\right](T_{22}^{-1}Q_2\overline{\nu}) \\ &= \frac{1}{2}\mu'W^*\mu = \frac{1}{2}\,\mathrm{trace}\left(W^*\mu\mu'\right) \end{aligned} \tag{61}$$

where $\mu = T_{22}^{-1} Q_2 \overline{\nu}$ and $W^* = \sum_{t=0}^{\tau-1} \lambda^t \Xi_t' \tilde{H} \Xi_t$.

Ξ_t satisfies the matrix difference equation

$$\begin{aligned}
\Xi_{t+1} &= \begin{pmatrix} \phi_{t+1}^* \\ -M_2^{\tau-1-(t+1)} \end{pmatrix} \\
&= \begin{pmatrix} M_1 \phi_t^* \\ 0 \end{pmatrix} + \begin{pmatrix} S_{11}^{-1}[-T_{12}M_2 + S_{12}] \\ -I \end{pmatrix} M_2^{\tau-1-(t+1)} \\
&= \begin{pmatrix} M_1 \phi_t^* \\ -M_2^{\tau-1-t} \end{pmatrix} + \begin{pmatrix} S_{11}^{-1}[-T_{12}M_2 + S_{12}] \\ M_2 - I \end{pmatrix} M_2^{\tau-1-(t+1)} \\
&= \tilde{M} \Xi_t + \Omega M_2^{\tau-t-2} \qquad (0 \leq t < \tau - 1)
\end{aligned} \tag{62}$$

with

$$\tilde{M} = \begin{pmatrix} M_1 & 0 \\ 0 & I \end{pmatrix}, \qquad \Omega = \begin{pmatrix} S_{11}^{-1}[-T_{12}M_2 + S_{12}] \\ M_2 - I \end{pmatrix} \tag{63}$$

and the initial value

$$\Xi_0 = \begin{pmatrix} \phi_0^* \\ -M_2^{\tau-1} \end{pmatrix} = \begin{pmatrix} Z_{11}^{-1} Z_{12} \\ -I \end{pmatrix} M_2^{\tau-1} \tag{64}$$

Note that the dynamic equation (62) is not defined for $t = \tau - 1$, since $M_2 = T_{22}^{-1} S_{22}$ is generally not invertible. The solution time path for Ξ_t $(0 \leq t < \tau - 1)$ can be obtained by either solving equation (62) backward or – if possible – by solving equation (62) forward.

Solving (62) backward in time yields

$$\Xi_t = \tilde{M}^t \Xi_0 + \sum_{s=0}^{t-1} \tilde{M}^{t-s-1} \Omega M_2^{\tau-s-2} \tag{65}$$

To obtain the forward solution assume that $M_1 = S_{11}^{-1} T_{11}$ is invertible. Then $\tilde{M}^{-1}$ exists and equation (62) can be written as

$$\Xi_t = \tilde{M}^{-1} \Xi_{t+1} - \tilde{M}^{-1} \Omega M_2^{\tau-t-2} \tag{66}$$

Given

$$\Xi_{\tau-1} = \begin{pmatrix} \phi_{\tau-1}^* \\ -I \end{pmatrix} \tag{67}$$

we obtain recursively for $t = \tau - n$:

$$\begin{aligned}
\Xi_{\tau-n} &= \left(\tilde{M}^{-1}\right)^{n-1} \Xi_{\tau-1} - \left(\tilde{M}^{-1}\right)^{n-1} \Omega - \left(\tilde{M}^{-1}\right)^{n-2} \Omega M_2 \\
&\quad - \left(\tilde{M}^{-1}\right)^{n-3} \Omega M_2^2 - \cdots - \left(\tilde{M}^{-1}\right) \Omega M_2^{n-2} \\
&= \left(\tilde{M}^{-1}\right)^{n-1} \Xi_{\tau-1} - \sum_{k=1}^{n-1} \left(\tilde{M}^{-1}\right)^{n-k} \Omega M_2^{k-1}
\end{aligned}$$

With $t = \tau - n$ we then get the forward solution

$$\Xi_t = \left(\tilde{M}^{-1}\right)^{\tau-t-1} \Xi_{\tau-1} - \sum_{k=1}^{\tau-t-1} \left(\tilde{M}^{-1}\right)^{\tau-t-k} \Omega M_2^{k-1} \quad (0 \le t < \tau - 2) \tag{68}$$

The total loss under the optimal unrestricted policy under commitment is now given by

$$J_0 = J_0^{(1)} + J_0^{(2)} = \frac{1}{2}\,\mathrm{trace}(W^* \mu\mu') + \frac{1}{2}\lambda^\tau\,\mathrm{trace}(V^* \tilde{K}\tilde{K}') \tag{69}$$

Obviously, the value of J_0 depends on the size of the lead time τ. In New Keynesian models we often have a hump-shaped pattern for the function $J_0 = J_0(\tau)$ where J_0 is increasing in τ for small values of τ (see Section 3).

In the limiting case of unanticipated shocks ($\tau = 0$), the total loss is given by

$$J_0 = J_0^{(2)} = \frac{1}{2}\tilde{K}' V^* \tilde{K} \tag{70}$$

where

$$\tilde{K} = K\Big|_{\tau=0} + S_{11}^{-1} Q_1 \overline{\nu} = Z_{11}^{-1}\tilde{w}_0 - S_{11}^{-1} Q_1 \overline{\nu} + S_{11}^{-1} Q_1 \overline{\nu} = Z_{11}^{-1}\tilde{w}_0 \tag{71}$$

Then

$$J_0 = \frac{1}{2}\tilde{w}_0' Z_{11}^{-1'} V^* Z_{11}^{-1}\tilde{w}_0 = \frac{1}{2}\tilde{w}_0' V \tilde{w}_0' = \frac{1}{2}\,\mathrm{trace}(V\tilde{w}_0\tilde{w}_0') \tag{72}$$

where

$$\tilde{w}_0\tilde{w}_0' = \begin{pmatrix} w_0 \\ p_{v\,0} \end{pmatrix} (w_0', p_{v\,0}') = \begin{pmatrix} w_0 w_0' & 0_{n_1\times n_2} \\ 0_{n_2\times n_2} & 0_{n_2\times n_2} \end{pmatrix} \tag{73}$$

and $V = Z_{11}^{-1'} V^* Z_{11}^{-1}$ satisfies the matrix equation

$$\begin{aligned} V &= Z_{11}^{-1'} V^* Z_{11}^{-1} = H^* + \lambda Z_{11}^{-1'} M_1' V^* M_1 Z_{11}^{-1} \\ &= H^* + \lambda Z_{11}^{-1'} M_1' Z_{11}' Z_{11}'^{-1} V^* Z_{11}^{-1} Z_{11} M_1 Z_{11}^{-1} = H^* + \lambda \Gamma' V \Gamma \end{aligned} \tag{74}$$

with $\Gamma = Z_{11} M_1 Z_{11}^{-1}$.

2.2 (Optimal) Simple Rules

The policy maker could alternatively commit to a suboptimal simple rule of the form

$$u_t = \Lambda k_t + \Psi\,\mathrm{E}_t\,k_{t+1} \tag{75}$$

where the constant matrices Λ and Ψ are $m \times n$. Assuming rational expectations and exogenous shocks of the form (2) which are anticipated in $t = 0$, we obtain the dynamic system

$$\begin{pmatrix} A & 0_{n\times m} \\ \Psi & 0_{m\times m} \end{pmatrix} \begin{pmatrix} k_{t+1} \\ u_{t+1} \end{pmatrix} = \begin{pmatrix} B & C \\ -\Lambda & I_{m\times m} \end{pmatrix} \begin{pmatrix} k_t \\ u_t \end{pmatrix} + \begin{pmatrix} D \\ 0_{m\times r} \end{pmatrix} \nu_{t+1} \tag{76}$$

The generalized Schur decomposition yields the system of equations

$$F \begin{pmatrix} \tilde{w}_{t+1} \\ \tilde{v}_{t+1} \end{pmatrix} = G \begin{pmatrix} \tilde{w}_t \\ \tilde{v}_t \end{pmatrix} + \begin{pmatrix} D \\ 0_{m\times r} \end{pmatrix} \nu_{t+1} \tag{77}$$

where $\tilde{w} = w$ is an $n_1 \times 1$ vector, $\tilde{v} = (v', u')'$ is an $(n_2+m)\times 1$ vector and where the square matrices F and G are $(n+m)\times(n+m)$ with the decomposition $QFZ = S$ and $QGZ = T$, where Q, Z, S, and T are $(n+m)\times(n+m)$ matrices. Since

$$\begin{pmatrix} \tilde{w} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \begin{pmatrix} \tilde{z} \\ \tilde{x} \end{pmatrix} \tag{78}$$

the matrices Z_{11}, Z_{12}, Z_{21}, and Z_{22} are now $n_1\times n_1$, $n_1\times(n_2+m)$, $(n_2+m)\times n_1$, and $(n_2+m)\times(n_2+m)$ respectively. The auxiliary variables $\tilde{z}$ and $\tilde{x}$ satisfy the system of equations

$$\begin{pmatrix} S_{11} & S_{12} \\ 0_{(n_2+m)\times n_1} & S_{22} \end{pmatrix} \begin{pmatrix} \tilde{z}_{t+1} \\ \tilde{x}_{t+1} \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ 0_{(n+m)\times n_1} & T_{22} \end{pmatrix} \begin{pmatrix} \tilde{z}_t \\ \tilde{x}_t \end{pmatrix} + \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \nu_{t+1} \tag{79}$$

where S_{11} and T_{11} are $n_1\times n_1$ matrices, S_{22} and T_{22} are $(n_2+m)\times(n_2+m)$ and S_{12} and T_{12} are $n_1\times(n_2+m)$. The matrices Q_1 and Q_2 are $n_1\times r$ and $(n_2+m)\times r$ respectively with

$$\begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} = Q \begin{pmatrix} D \\ 0_{m\times r} \end{pmatrix} \tag{80}$$

The solution of (79) is given by (15) and (17). For $t \geq \tau$, we obtain $\tilde{v}_t = N\tilde{w}_t = Nw_t$, where $N = Z_{21}Z_{11}^{-1}$ is now an $(n_2+m)\times n_1$ matrix.

The loss function (51) simplifies to

$$J_t = \frac{1}{2}\sum_{i=0}^{\infty} \lambda^i (w'_{t+i}, \tilde{v}'_{t+i}) H \begin{pmatrix} w_{t+i} \\ \tilde{v}_{t+i} \end{pmatrix} \tag{81}$$

since $\tilde{D}_1 = I_{n_1\times n_1}$, $\tilde{D}_2 = I_{(n_2+m)\times(n_2+m)}$ and therefore $\tilde{D} = I_{(n+m)\times(n+m)}$ (cf. (50)). J_t can be partitioned using (51). $J_t^{(2)}$ can be written as (54) with $H^* = \tilde{N}'H\tilde{N}$ and $\tilde{N} = (I_{n_1\times n_1}, N')'$. The value of the loss function J_0 for given matrices Λ and Ψ is given by $J_0 = J_0^{(1)} + J_0^{(2)}$, where $J_0^{(1)}$ and $J_0^{(2)}$ are defined in (59) and (61) respectively.

The minimization of J_0 with respect to the coefficients of the matrices Λ and Ψ yields an optimal simple rule of the form (75).

3 Example: A Hybrid New Keynesian Model

The model is a standard New Keynesian model for a closed and cashless economy with the additional features of internal habit formation in consumption preferences and a variant of the Calvo (1983) mechanism with partial indexation of non-optimized prices to past inflation.[6] The economy consists of final

[6]Similar models are applied by Smets and Wouters (2003), Giannoni and Woodford (2004), or Casares (2006).

goods producers, labor bundlers, households, and intermediate goods producers.

Final goods producers use a continuum of intermediate goods $Y_t(i)$ to produce the homogenous final good Y_t in a perfectly competitive market. A final goods producer maximizes his profits $P_tY_t - \int_0^1 P_t(i)Y_t(i)di$, subject to the following CES production function

$$Y_t = \left(\int_0^1 Y_t(i)^{\frac{1}{1+\lambda_p}} di \right)^{1+\lambda_p} \tag{82}$$

where P_t is the price of the final good, $P_t(i)$ is the price of the intermediate good i, and $(1 + \lambda_p)$ is the mark-up in the intermediate goods market.

The first-order condition for profit maximization yields the demand function for intermediate good i

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\frac{(1+\lambda_p)}{\lambda_p}} Y_t \tag{83}$$

and the equation for marginal costs

$$P_t = \left(\int_0^1 P_t(i)^{-\frac{1}{\lambda_p}} di \right)^{-\lambda_p} \tag{84}$$

Analogously to final goods producers, labor bundlers buy differentiated labor types $N_t(j)$, aggregate them to N_t and sell it to the intermediate goods producers under perfectly competitive conditions. A bundler maximizes his profits $W_tN_t - \int_0^1 W_t(j)N_t(j)dj$, subject to the following CES aggregation function

$$N_t = \left(\int_0^1 N_t(j)^{\frac{1}{1+\lambda_{w,t}}} dj \right)^{1+\lambda_{w,t}} \tag{85}$$

W_t is the price of the labor bundle N_t, $W_t(j)$ denotes the price of labor type j and $(1 + \lambda_{w,t})$ is the time-varying wage mark-up.

The first-order condition for profit maximization yields the demand function for labor type j

$$N_t(j) = \left(\frac{W_t(j)}{W_t} \right)^{-\frac{(1+\lambda_{w,t})}{\lambda_{w,t}}} N_t \tag{86}$$

and the wage index equation

$$W_t = \left(\int_0^1 W_t(j)^{-\frac{1}{\lambda_{w,t}}} dj \right)^{-\lambda_{w,t}} \tag{87}$$

The economy is made up by a continuum of households, indexed by $j \in [0, 1]$. Each household j is a monopolistic supplier of labor type $N_t(j)$. The household determines the amount of the final good $C_t(j)$ for consumption, its one-period nominal bond holdings $B_t(j)$, and chooses the wage for its labor type $W_t(j)$ in order to maximize its lifetime utility

$$\mathrm{E}_t \sum_{k=0}^{\infty} \beta^k \left(\frac{1}{1-\sigma} \left(C_t(j) - hC_{t-1}(j)\right)^{1-\sigma} - \frac{1}{1+\eta} N_t(j)^{1+\eta} \right) \tag{88}$$

where β is the discount factor, $\sigma \geq 1$ is the inverse of the intertemporal elasticity of substitution in consumption, and η is the inverse of the labor supply elasticity. $C_{t-1}(j)$ is the consumption of the jth household in period $t-1$ and $N_t(j)$ are the total hours worked. We assume $h \geq 0$ to allow for internal habit formation in consumption. Maximization of (88) is subjected to the labor demand function (86) and the households' period-by-period budget constraint is given by

$$C_t(j) + \frac{B_t(j)}{P_t} = \frac{W_t(j)}{P_t} N_t(j) + \frac{R_{t-1} B_{t-1}(j)}{P_t} + D_t^r(j) \tag{89}$$

where R_t is the one-period gross nominal interest rate on households jth nominal bond holdings $B_t(j)$, and $D_t^r(j)$ are dividends, expressed in real terms.

The first-order conditions for this maximization problem are given by

$$\beta R_t \, \mathrm{E}_t \, \pi_{t+1}^{-1} = \mathrm{E}_t \left[\frac{(C_t - hC_{t-1})^{-\sigma} - h\beta(C_{t+1} - hC_t)^{-\sigma}}{(C_{t+1} - hC_t)^{-\sigma} - h\beta(C_{t+2} - hC_{t+1})^{-\sigma}} \right] \tag{90}$$

$$\frac{W_t}{P_t} = (1 + \lambda_{w,t}) \; \mathrm{E}_t \left[\frac{N_t^{\eta}}{(C_t - hC_{t-1})^{-\sigma} - h\beta(C_{t+1} - hC_t)^{-\sigma}} \right] \tag{91}$$

where $\pi_t = P_t/P_{t-1}$ is the gross rate of price inflation. We make use of the fact that all households are faced with the same optimization problem and hence, choose the same amount of consumption $C_t(j) = C_t$, the same nominal wage $W_t(j) = W_t$, and supply the same amount of labor $N_t(j) = N_t$.

Each intermediate goods producer is a monopolistic supplier of the intermediate good $i \in [0,1]$. Firm i uses the amount $N_t(i)$ of homogenous labor and the constant returns to scale technology $Y_t(i) = N_t(i)$, to produce its intermediate good $Y_t(i)$. Real marginal costs are the same for all firms and is given by $MC_t(i) = W_t/P_t$.

The price-setting decision for profit-maximization is constrained by a standard Calvo mechanism. In each period, the intermediate goods producer faces the constant probability $1-\theta$ of being allowed to re-optimize his price $P_t(i)$. We follow Smets and Wouters (2003) by assuming that a firm which cannot re-optimize his price, resets the price according to $P_t(i) = P_{t-1}(i)\pi_{t-1}^{\gamma}$, where γ is the degree of price indexation. The firm chooses $P_t(i)$ in order to maximize

$$\mathrm{E}_t \sum_{k=0}^{\infty} \theta^k \Delta_{t,t+k} \left(\frac{P_t(i)\Pi_{t,t+k-1}}{P_{t+k}} Y_{t+k}(i) - MC_{t+k} Y_{t+k}(i) \right) \tag{92}$$

subject to the sequence of demand functions

$$Y_{t+k}(i) = \left(\frac{P_t(i)\Pi_{t,t+k-1}}{P_{t+k}} \right)^{-\frac{(1+\lambda_p)}{\lambda_p}} Y_{t+k} \qquad \text{for } k = 0, 1, 2, \ldots \tag{93}$$

where $\Delta_{t,t+k}$ denotes the stochastic discount factor for real payoffs and $\Pi_{t,t+k-1} = \pi_t^\gamma \pi_{t+1}^\gamma \dots \pi_{t+k-1}^\gamma = (P_{t+k-1}/P_{t-1})^\gamma$.

The first-order condition for the price-setting problem yields

$$P_t^*(i) = (1+\lambda_p)\frac{\mathrm{E}_t \sum_{k=0}^{\infty} \theta^k \Delta_{t,t+k} MC_{t+k} (P_{t+k}/\Pi_{t,t+k-1})^{(1+\lambda_p)/\lambda_p} Y_{t+k}}{\mathrm{E}_t \sum_{k=0}^{\infty} \theta^k \Delta_{t,t+k} (P_{t+k}/\Pi_{t,t+k-1})^{-1/\lambda_p} Y_{t+k}} \tag{94}$$

Dividing equation (94) by P_t yields

$$\frac{P_t^*(i)}{P_t} = \mu_p \frac{\mathrm{E}_t \sum_{k=0}^{\infty} \theta^k \Delta_{t,t+k} MC_{t+k} \left(\frac{P_{t+k}}{P_t}\right)^{\frac{1+\lambda_p}{\lambda_p}} \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{-\frac{\gamma(1+\lambda_p)}{\lambda_p}} Y_{t+k}}{\mathrm{E}_t \sum_{k=0}^{\infty} \theta^k \Delta_{t,t+k} \left(\frac{P_{t+k}}{P_t}\right)^{\frac{1}{\lambda_p}} \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{-\frac{\gamma}{\lambda_p}} Y_{t+k}} \tag{95}$$

where $\mu_p = 1 + \lambda_p$.

Since all firms which are allowed to re-optimize their price will choose the same price $P_t^*(i) = P_t^*$, the price index (84) can be rewritten as

$$1 = \theta \left(\frac{\pi_{t-1}^\gamma}{\pi_t}\right)^{-\lambda_p} + (1-\theta)\left(\frac{P_t^*}{P_t}\right)^{-\lambda_p} \tag{96}$$

Log-linearizing equation (96) yields

$$\hat{P}_t^* - \hat{P}_t = \frac{\theta}{1-\theta}(\hat{\pi}_t - \gamma \hat{\pi}_{t-1}) \tag{97}$$

Note that we use the convention that a hat above a variable denotes the percentage deviation from its steady-state value.

By combining the latter equation with the log-linearized price-setting condition (95), we finally obtain

$$\hat{\pi}_t = \frac{\gamma}{1+\beta\gamma}\hat{\pi}_{t-1} + \frac{\beta}{1+\beta\gamma}\mathrm{E}_t \hat{\pi}_{t+1} + \Theta \hat{MC}_t \tag{98}$$

where $\Theta = \frac{(1-\beta\theta)(1-\theta)}{\theta(1+\beta\gamma)}$. By log-linearizing the optimality condition (91), using the log-linearized overall resource constraint $\hat{Y}_t = \hat{C}_t$ and using the fact that $\widehat{W_t/P_t} = \widehat{MC}_t$ and $\hat{Y}_t = \hat{N}_t$, we obtain

$$\widehat{MC}_t = \hat{\lambda}_{w,t} + (\eta + \delta_1)\hat{Y}_t - \delta_2 \hat{Y}_{t-1} - \beta\delta_2 \mathrm{E}_t \hat{Y}_{t+1} \tag{99}$$

where $\delta_1 = \frac{\sigma(1+\beta h^2)}{(1-h)(1-\beta h)}$, $\delta_2 = \frac{h\sigma}{(1-h)(1-\beta h)}$. The log-linearized mark-up $\hat{\lambda}_{w,t}$ is described by the AR(1) process

$$\hat{\lambda}_{w,t} = \rho \hat{\lambda}_{w,t-1} + e_t \tag{100}$$

By inserting the latter equation into equation (98), we obtain a hybrid Phillips curve that follows

$$\hat{\pi}_t = \omega_1 \mathrm{E}_t \hat{\pi}_{t+1} + \omega_2 \hat{\pi}_{t-1} + \omega_3 \hat{Y}_t - \omega_4 \hat{Y}_{t-1} - \beta\omega_4 \mathrm{E}_t \hat{Y}_{t+1} + \Theta \hat{\lambda}_{w,t} \tag{101}$$

where $\omega_1 = \frac{\beta}{1+\beta\gamma}$, $\omega_2 = \frac{\gamma}{1+\beta\gamma}$ $\omega_3 = \Theta(\eta + \delta_1)$, and $\omega_4 = \Theta\delta_2$.

Note that in our model the level of output in the absence of nominal rigidities (the natural level) Y_t^n is constant. Thus, the linearized output $\hat{Y}_t$ coincides with the linearized output gap $\hat{Y}_t^g = \hat{Y}_t - \hat{Y}_t^n$, where $\hat{Y}_t^n = 0$. Further note that for $\gamma = 0$, equation (101) collapses into the purely forward-looking New Keynesian Phillips curve.

By log-linearizing the optimality condition (90) and using $\hat{Y}_t = \hat{C}_t$, we obtain

$$\hat{Y}_t = \kappa_1 \rho \hat{Y}_{t-1} + \kappa_2 \operatorname{E}_t \hat{Y}_{t+1} - \kappa_3 \operatorname{E}_t \hat{Y}_{t+2} - \kappa_4(\hat{R}_t - \operatorname{E}_t \hat{\pi}_{t+1}) \tag{102}$$

where $\kappa_1 = \frac{h}{1+h+\beta h^2}$, $\kappa_2 = \frac{1+\beta h+\beta h^2}{1+h+\beta h^2}$, $\kappa_3 = \frac{\beta h}{1+h+\beta h^2}$, and $\kappa_4 = \frac{(1-h)(1-\beta h)}{\sigma(1+h+\beta h^2)}$. Note that for $h = 0$, we obtain the purely forward-looking New Keynesian IS curve.

Following Woodford (2003, Ch. 6) and Giannoni and Woodford (2004), a second-order approximation to the households' utility yields a loss function of the form

$$J_0 = \operatorname{E}_0 \sum_{t=0}^{\infty} \beta^t \left((\hat{\pi}_t - \gamma\hat{\pi}_{t-1})^2 + \alpha_Y (\hat{Y}_t - \delta\hat{Y}_{t-1})^2 \right) \tag{103}$$

where $\alpha_y = \frac{\Theta h \sigma \lambda_p}{(1+\lambda_p)\delta(1-\beta h)(1-h)}$ and δ is the smaller root of the quadratic equation

$$\frac{h\sigma}{(1-\beta h)(1-h)}(1+\beta\delta^2) = \left(\eta + \frac{\sigma}{(1-\beta h)(1-h)}(1+\beta h^2)\right)\delta \tag{104}$$

We follow Giannoni and Woodford (2004) and Casares (2006) by assuming that the monetary authority is concerned about the volatility of the nominal interest rate. Therefore, we augment the welfare-theoretic loss function by the additional term $\alpha_R \hat{R}_t^2$, where α_R measures the weight on interest rate stabilization.

The monetary authority then seeks to minimize the loss function

$$J_0 = \operatorname{E}_0 \sum_{t=0}^{\infty} \beta^t \left((\hat{\pi}_t - \gamma\hat{\pi}_{t-1})^2 + \alpha_Y (\hat{Y}_t - \delta\hat{Y}_{t-1})^2 + \alpha_R \hat{R}_t^2 \right) \tag{105}$$

subject to the model equations (100), (101), and (102). Note that in our model, the discount factor for the policy-maker, λ, is equal to the household's discount factor β.

In order to solve the model by using the methods outlined in Section 2, we define the policy objective parameters $\hat{Y}_t^o = \hat{Y}_t - \delta\hat{Y}_{t-1}$ and $\hat{\pi}_t^o = \hat{\pi}_t - \gamma\hat{\pi}_{t-1}$. Furthermore, we define the auxiliary variables $\widetilde{\pi}_t = \hat{\pi}_{t-1}$, $\widetilde{Y}_t = \hat{Y}_{t-1}$, and $s_t = \operatorname{E}_t \hat{\pi}_{t+1}$. If we add the definition of the real interest rate $\hat{r}_t = \hat{R}_t - \operatorname{E}_t \hat{\pi}_{t+1}$, we finally obtain a 3×1 vector w_t of predetermined variables given by $w_t = (\hat{\lambda}_{w,t}, \widetilde{\pi}_t, \widetilde{Y}_t)'$, a 6×1 vector v_t of non-predetermined variables given by $v_t = (\hat{\pi}_t, \hat{Y}_t, s_t, \hat{r}_t, \hat{\pi}_t^o, \hat{y}_t^o)'$, the vector of policy instruments u_t which is simply the scalar $u_t = \hat{R}_t$, and the 1×1 shock vector $\nu_t = e_t$. The 9×9 matrices A and

B are given by

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\beta}{1+\beta\gamma} & -\beta\omega_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \kappa_4 & \kappa_2 & -\kappa_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -\Theta & -\frac{\gamma}{1+\beta\gamma} & \omega_4 & 1 & -\omega_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\kappa_1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \gamma & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \delta & 0 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

while the 9×1 matrices C and D are

$$C = \begin{pmatrix} 0 & 0 & 0 & 0 & \kappa_4 & 0 & 1 & 0 & 0 \end{pmatrix}'$$
$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}'$$

Finally, the matrices $\tilde{W}$, P, and R are given by $P = \mathbf{0}_{9\times 9}$, $R = \alpha_R$, and

$$\tilde{W} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha_Y \end{pmatrix}$$

We complete the description of the model by presenting the calibration. The time unit is one quarter. The discount rate is equal to $\beta = 0.99$, implying a quarterly steady-state real interest rate of approximately one percent. The intertemporal elasticity of substitution, σ, is assumed to $\sigma = 2$. We follow Casares (2006) and set the habit formation parameter to $h = 0.85$ implying that the weight on lagged output in the IS equation is 1/3. The calibrated $\eta = 3$ implies a labor supply elasticity with respect to the real wage of 1/3. λ_p is set to 8/7 which implies a steady-state mark-up in the goods market of approximately

14 percent. We assume the linearized wage mark-up $\hat{\lambda}_{w,t}$ to be persistent and choose ρ equal to 0.8. The Calvo parameter θ is set to 0.75 implying an average duration of price contracts of one year. The price indexation parameter γ is set to 0.45 which is roughly equal to the value reported by Smets and Wouters (2003). This implies that the weight on lagged inflation in the Phillips curve equation is 0.31.

The parameter values chosen for our model imply a weight on output in the policy-makers' objective function of approximately $\alpha_Y = 0.69$. Following Casares (2006), we set $\alpha_R = 0.0088$ implying a small preference for interest rate smoothing.

For the analysis concerning anticipated and unanticipated shocks, we assume that the economy is in a deterministic steady-state until period $t = 0$. In the case of an unanticipated shock, the mark-up $\hat{\lambda}_{w,t}$ jumps by one percent in period $t = 0$ and begins to fall thereafter. In the case of an anticipated shock, the agents anticipate in period $t = 0$ that a one percent increase in the mark-up will take place at some future date $\tau > 0$. They also know that the mark-up will subsequently decline according to the autoregressive process (100), where now $e_t = 1$ for $t = \tau$ and $e_t = 0$ for $t \neq \tau$. Note that τ also defines the lengths of the anticipation period or the time interval between $t = 0$ and $t = \tau$. In order to obtain impulse response functions and welfare results, we simulate dynamic adjustment paths and the welfare loss function by using the methods outlined in Section 2.[7]

Figure 1 depicts the impulse response functions of inflation, output, nominal, and real interest rates under the unrestricted optimal monetary policy. The solid lines with circles represent the responses to an unforeseen cost-push shock that emerged in period $t = 0$. The solid lines with squares, triangles, and stars represent responses to a cost-push shock whose realization in period $\tau = 1$, $\tau = 2$, or $\tau = 3$ is anticipated in period $t = 0$.

An *unanticipated* rise in the wage mark-up puts upward pressure on the prices of intermediate goods and hence on inflation. Despite the instantaneous jump in inflation, the real interest rate rises due to the sharp increase in the nominal interest rate. The increase in the real interest rate induces households to postpone consumption which implies an abrupt drop in output. Subsequently, the nominal interest rate continues to rise. This leads – in conjunction with the decline in inflation – to hump-shaped response functions of the real interest rate and output.

In the case of *anticipated* shocks, the optimal policy calls for a decline in nominal and real interest rates in response to the anticipation of a future rise in marginal costs. At the latest with the occurrence of the anticipated shock in period τ, the nominal and real interest rates start to rise and display a hump-shaped development. Inflation declines in response to the anticipation of the future rise in marginal costs. After this initial decline, inflation starts to rise and peaks in the period when the anticipated shock materializes. Output displays a hump-shaped downturn, starting at the point of anticipation, $t = 0$. The drop

[7]Matlab codes can be downloaded from the author's webpage at http://www.wiso.uni-kiel.de/vwlinstitute/Wohltmann/REAS_solution.zip.

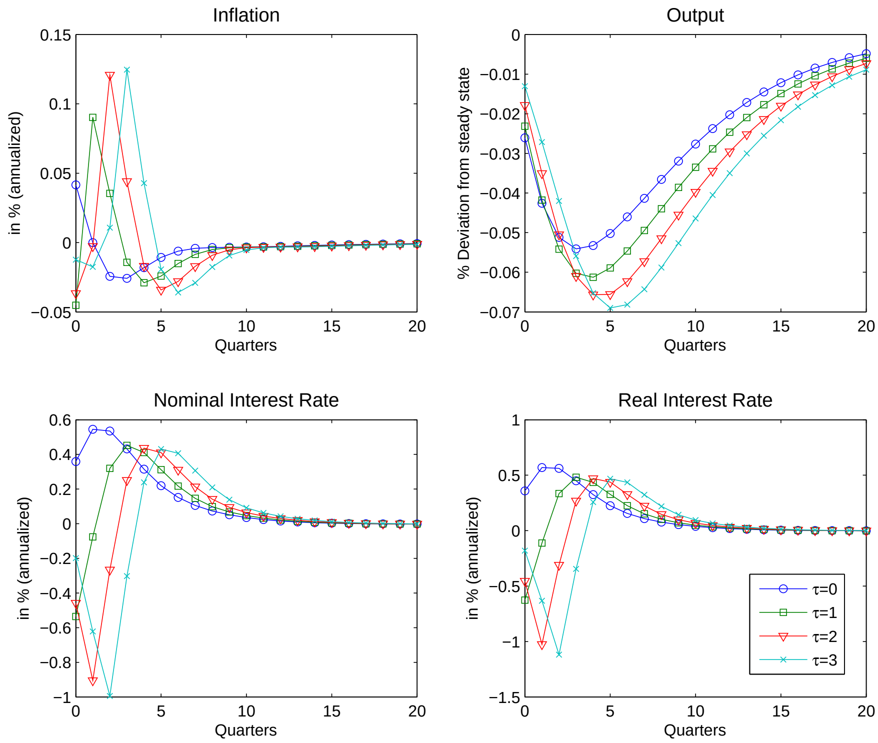

Figure 1: Impulse response functions under unrestricted optimal monetary policy.

Notes: Solid lines with circles denote responses to an unanticipated cost-push shock, solid lines with squares, triangles, and stars denote responses to an anticipated cost-push shock taking place in period $\tau = 1$, $\tau = 2$, and $\tau = 3$.

in output is thereby amplified by the lengths of the anticipation period, τ.

Notably, the anticipation of future shocks leads to an increase in the persistence (or volatility) of inflation, output as well as nominal and real interest rates which increases in lead time τ. Thereby, persistence is measured as the total variation of a variable over time, i.e. by its intertemporal deviation from its initial steady-state. The impact or anticipation effect, however, is inversely related to the time span between anticipation and realization of the cost-push shock. It measures the initial jump of a variable taking place at the time of anticipation.

The opposing effects of anticipations are shown in Figure 2 which displays the welfare loss as a function of the time span between the anticipation and the occurrence of the cost-push shock. The welfare function exhibits a hump-shaped pattern implying that for a realistic time span between the anticipation and the realization of cost-push shocks, anticipated shocks entail higher welfare losses than unanticipated shocks of equal size. The rationale is that the anticipation effect is dominated by the persistence effect. A welfare gain from anticipating can only be achieved for very large values of τ. Besides the anticipation effect, this can also be explained by discounting the realization impacts from period τ to period $t = 0$.

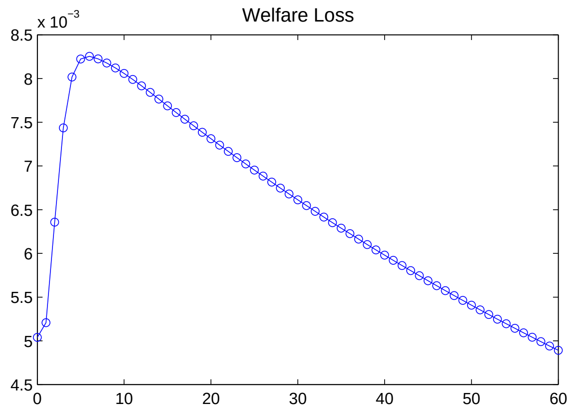

Figure 2: Welfare loss for different lengths of the anticipation period under unrestricted optimal monetary policy

The results we obtained from our simulations show that the welfare loss of anticipated cost-shocks exceeds the welfare loss of an unanticipated cost-shock of equal magnitude for plausible lengths of the anticipation period. Hence, our results strongly support the findings of Wohltmann and Winkler (2008) who report a similar result within the purely forward-looking canonical New Keynesian model.

4 Conclusion

In this paper, we presented a method to solve linear dynamic rational expectations models with anticipated shocks and optimal policy by using the generalized Schur decomposition method. Furthermore, we determine the optimal unrestricted and restricted policy responses to anticipated shocks. Our approach also allows for the evaluation of the widely discussed case of unpredictable shocks and can therefore be seen as a generalization of the methods summarized by Söderlind (1999). We demonstrated our method by means of a calibrated New Keynesian model with internal habit formation in consumption preferences, a variant of Calvo price staggering with partial indexation to past inflation, a time-varying wage mark-up which represents a typical cost-push shock, and a utility-based loss function. We simulated the model economy's responses to unanticipated and anticipated cost-push shocks under the unrestricted optimal monetary policy. We then showed that anticipated shocks amplify both, the stagflationary effects of cost-push shocks and the overall welfare loss. Hence, our results strongly support the previous work by Wohltmann and Winkler (2008) who find welfare-reducing effects of anticipations within the purely forward-looking canonical New Keynesian model.

Appendix

The Stochastic Case

We now assume that ν_{t+1} is an $r \times 1$ vector of independent and identically distributed white noise disturbances with variance-covariance matrix $\Sigma_{\nu\nu} = E(\nu_t \nu_t')$. The i.i.d shocks are, by definition, unpredictable ($\tau = 0$) and occur at time $t = 0$. Since $\mathrm{E}_t(\nu_{t+1}) = 0_{r\times 1}$, equation (7) implies

$$F \cdot \mathrm{E}_t \begin{pmatrix} \tilde{w}_{t+1} \\ \tilde{v}_{t+1} \end{pmatrix} = G \begin{pmatrix} \tilde{w}_t \\ \tilde{v}_t \end{pmatrix} \tag{A1}$$

The Schur decomposition yields the following system of equations

$$\begin{pmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{pmatrix} \mathrm{E}_t \begin{pmatrix} \tilde{z}_{t+1} \\ \tilde{x}_{t+1} \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix} \begin{pmatrix} \tilde{z}_t \\ \tilde{x}_t \end{pmatrix} \tag{A2}$$

where

$$\begin{pmatrix} \tilde{w}_t \\ \tilde{v}_t \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \begin{pmatrix} \tilde{z}_t \\ \tilde{x}_t \end{pmatrix} \tag{A3}$$

and $\tilde{x}_t = 0$ for all $t \geq T = 0$. Partitioning the matrices A and B in equation (1) to conform with the dimension of w_t and v_t, i.e.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{A4}$$

Equation (1) then implies

$$A_{11} w_{t+1} + A_{12}\, \mathrm{E}_t\, v_{t+1} = B_{11} w_t + B_{12} v_t + C_1 u_t + D_1 \nu_{t+1} \tag{A5}$$

and

$$A_{11}\, \mathrm{E}_t\, w_{t+1} + A_{12}\, \mathrm{E}_t\, v_{t+1} = B_{11} w_t + B_{12} v_t + C_1 u_t \tag{A6}$$

where

$$C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}, \qquad D = \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \tag{A7}$$

From (A5) and (A6) we get

$$A_{11}(w_{t+1} - \mathrm{E}_t\, w_{t+1}) = D_1 \nu_{t+1} \tag{A8}$$

so that

$$w_{t+1} - \mathrm{E}_t\, w_{t+1} = A_{11}^{-1} D_1 \nu_{t+1} \tag{A9}$$

holds (provided A_{11}^{-1} exists). The corresponding equation for the costate vector p_v is given by (Backus and Driffill, 1986)

$$p_{v,t+1} - \mathrm{E}_t\, p_{v,t+1} = 0_{n_2 \times 1} \tag{A10}$$

Defining $\tilde{w}_t = (w_t', p_{vt}')'$ and using equations (A2) and (A3) then imply

$$\tilde{w}_{t+1} - \mathrm{E}_t\, \tilde{w}_{t+1} = Z_{11}(\tilde{z}_{t+1} - \mathrm{E}_t\, \tilde{z}_{t+1}) = Z_{11}(\tilde{z}_{t+1} - S_{11}^{-1}T_{11}\tilde{z}_t) = \begin{pmatrix} A_{11}^{-1}D_1\nu_{t+1} \\ 0_{n_2\times 1} \end{pmatrix} \tag{A11}$$

and therefore

$$\tilde{z}_{t+1} = (S_{11}^{-1}T_{11})\tilde{z}_t + Z_{11}^{-1}\begin{pmatrix} A_{11}^{-1}D_1\nu_{t+1} \\ 0_{n_2\times 1} \end{pmatrix} = (S_{11}^{-1}T_{11})\tilde{z}_t + Z_{11}^{-1}\begin{pmatrix} A_{11}^{-1}D_1 \\ 0_{n_2\times r} \end{pmatrix}\nu_{t+1} \tag{A12}$$

The solution of the VAR(1) process (A12) has the general form

$$\tilde{z}_t = (S_{11}^{-1}T_{11})^t K + \sum_{s=0}^{t-1}(S_{11}^{-1}T_{11})^{t-s-1}Z_{11}^{-1}\begin{pmatrix} A_{11}^{-1}D_1 \\ 0_{n_2\times r} \end{pmatrix}\nu_{s+1} \tag{A13}$$

where

$$K = \tilde{z}_0 = Z_{11}^{-1}\tilde{w}_0 = Z_{11}^{-1}\begin{pmatrix} w_0 \\ 0_{n_2\times 1} \end{pmatrix} \tag{A14}$$

Since $\mathrm{E}_0\,\nu_{s+1} = 0$ the expected time path of $\tilde{z}_t$ is given by

$$\mathrm{E}_0\,\tilde{z}_t = (S_{11}^{-1}T_{11})^t Z_{11}^{-1}\tilde{w}_0 \tag{A15}$$

Premultiplying equation (A12) with Z_{11} and using $\tilde{w}_t = Z_{11}\tilde{z}_t$ to obtain the VAR(1) process

$$\tilde{w}_{t+1} = \Gamma\tilde{w}_t + \begin{pmatrix} A_{11}^{-1}D_1 \\ 0_{n_2\times r} \end{pmatrix}\nu_{t+1} \tag{A16}$$

where

$$\Gamma = Z_{11}(S_{11}^{-1}T_{11})Z_{11}^{-1} \tag{A17}$$

Then

$$\tilde{w}_t = \Gamma^t\tilde{w}_0 + \sum_{s=0}^{t-1}\Gamma^{t-s-1}\begin{pmatrix} A_{11}^{-1}D_1 \\ 0_{n_2\times r} \end{pmatrix}\nu_{s+1} \tag{A18}$$

and the expected future path of $\tilde{w}_t$ is given by

$$\mathrm{E}_0\,\tilde{w}_t = \Gamma^t\tilde{w}_0 = \Gamma^t\begin{pmatrix} A_{11}^{-1}D_1 \\ 0_{n_2\times r} \end{pmatrix}\nu_0 \tag{A19}$$

The solution to the forward-looking vector $\tilde{v}_t$ follows from

$$\tilde{v}_t = Z_{21}\tilde{z}_t = Z_{21}Z_{11}^{-1}\tilde{w}_t = N\tilde{w}_t \qquad (N = Z_{21}Z_{11}^{-1}) \tag{A20}$$

by inserting the solution time path of $\tilde{w}_t$.

In order to determine the minimum value of the loss function J_0, set

$$\varepsilon_{t+1} = \begin{pmatrix} A_{11}^{-1} D_1 \\ 0_{n_2 \times r} \end{pmatrix} \nu_{t+1} \tag{A21}$$

According to (48), (51), (54), and (A18) we then obtain

$$\begin{aligned} J_0 &= \frac{1}{2} \, \mathrm{E}_0 \sum_{i=0}^{\infty} \lambda^i \tilde{w}_i' H^* \tilde{w}_i \\ &= \frac{1}{2} \sum_{i=0}^{\infty} \lambda^i \Big\{ (\Gamma^i \tilde{w}_0)' H^* (\Gamma^i \tilde{w}_0) + 2 \, \mathrm{E}_0 (\Gamma^i \tilde{w}_0)' H^* \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1} \varepsilon_{s+1} \big) \\ &\quad + \mathrm{E}_0 \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1} \varepsilon_{s+1} \big)' H^* \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1} \varepsilon_{s+1} \big) \Big\} \\ &= \frac{1}{2} \tilde{w}_0' \big(\sum_{i=0}^{\infty} \lambda^i \Gamma^{i\prime} H^* \Gamma^i \big) \tilde{w}_0 \\ &\quad + \frac{1}{2} \sum_{i=0}^{\infty} \lambda^i \, \mathrm{E}_0 \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1} \varepsilon_{s+1} \big)' H^* \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1} \varepsilon_{s+1} \big) \end{aligned} \tag{A22}$$

where we have used $\mathrm{E}_0 \, \varepsilon_{s+1} = 0$. $V = \sum_{i=0}^{\infty} \lambda^i \Gamma^{i\prime} H^* \Gamma^i$ satisfies the matrix equation (cf. (74))

$$V = H^* + \lambda \Gamma' V \Gamma \tag{A23}$$

and

$$\frac{1}{2} \tilde{w}_0' \big(\sum_{i=0}^{\infty} \lambda^i \Gamma^{i\prime} H^* \Gamma^i \big) \tilde{w}_0 = \frac{1}{2} \tilde{w}_0' V \tilde{w}_0 = \frac{1}{2} \, \mathrm{trace}(V \tilde{w}_0 \tilde{w}_0') \tag{A24}$$

To calculate the infinite sum in (A22) note that

$$\begin{aligned} &\mathrm{E}_0 \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1} \varepsilon_{s+1} \big)' H^* \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1} \varepsilon_{s+1} \big) \\ &= \mathrm{E}_0 (\Gamma^{i-1} \varepsilon_1 + \Gamma^{i-2} \varepsilon_2 + ... + \Gamma^0 \varepsilon_i)' H^* (\Gamma^{i-1} \varepsilon_1 + \Gamma^{i-2} \varepsilon_2 + ... + \Gamma^0 \varepsilon_i) \\ &= \mathrm{E}_0 (\Gamma^{i-1} \varepsilon_1)' H^* (\Gamma^{i-1} \varepsilon_1) + \mathrm{E}_0 (\Gamma^{i-2} \varepsilon_2)' H^* (\Gamma^{i-2} \varepsilon_2) + ... + \mathrm{E}_0 (\Gamma^0 \varepsilon_i)' H^* (\Gamma^0 \varepsilon_i) \\ &= \mathrm{E}_0 \, \varepsilon_i' (\Gamma^{0\prime} H^* \Gamma^0 + \Gamma' H^* \Gamma + ... + \Gamma^{i-2\prime} H^* \Gamma^{i-2} + \Gamma^{i-1\prime} H^* \Gamma^{i-1}) \varepsilon_i \\ &= \mathrm{E}_0 \, \varepsilon_i' \big(\sum_{s=0}^{i-1} \Gamma^{i-s-1\prime} H^* \Gamma^{i-s-1} \big) \varepsilon_i \end{aligned} \tag{A25}$$

since $\mathrm{E}_0(\varepsilon_i' \varepsilon_j) = 0$ for $i \neq j$. The variance-covariance matrix

$$\mathrm{E}_0(\varepsilon_i \varepsilon_i') = \mathrm{E}_0(\varepsilon_j \varepsilon_j') = \Sigma_{\varepsilon\varepsilon} \tag{A26}$$

is independent of i and j. We then obtain

$$\begin{aligned}&\frac{1}{2}\sum_{i=0}^{\infty}\lambda^i\,\mathrm{E}_0\,\big(\sum_{s=0}^{i-1}\Gamma^{i-s-1}\varepsilon_{s+1}\big)'H^*\big(\sum_{s=0}^{i-1}\Gamma^{i-s-1}\varepsilon_{s+1}\big)\\&=\frac{1}{2}\frac{\lambda}{1-\lambda}\,\mathrm{trace}(V\Sigma_{\varepsilon\varepsilon})\end{aligned}\tag{A27}$$

with V defined as in equation (A23). The optimal value of the loss function J_0 in the stochastic case (with $T=0$) is then given by

$$J_0=\frac{1}{2}\,\mathrm{trace}(V\tilde{w}_0\tilde{w}_0')+\frac{1}{2}\frac{\lambda}{1-\lambda}\,\mathrm{trace}(V\Sigma_{\varepsilon\varepsilon})\tag{A28}$$

Note that (A28) is a generalization of equation (72) where we have assumed a deterministic shock in $t=0$ ($\Sigma_{\varepsilon\varepsilon}=0$). The formula (A28) holds for a discount factor λ with $0<\lambda<1$.[8] The right-hand side of (A28) is not defined in the special case $\lambda=1$. If the discount factor λ approaches unity we must scale the intertemporal loss function J_0 by the factor $(1-\lambda)$ (Rudebusch and Svensson, 1999). Equation (A28) then implies

$$(1-\lambda)J_0=\frac{1}{2}(1-\lambda)\,\mathrm{trace}(V\tilde{w}_0\tilde{w}_0')+\frac{1}{2}\lambda\,\mathrm{trace}(V\Sigma_{\varepsilon\varepsilon})\tag{A29}$$

The scaled intertemporal loss function $(1-\lambda)J_0$ converges as λ approaches unity. (A29) implies

$$\lim_{\lambda\to1}(1-\lambda)J_0=\frac{1}{2}\,\mathrm{trace}(V\Sigma_{\varepsilon\varepsilon})\tag{A30}$$

Note that in the case $T=0$ and $\lambda=1$ the RHS of (A30) equals the RHS of (72) provided $w_0w_0'=\Sigma_{\varepsilon\varepsilon}$. In this special case the stochastic and deterministic case are equivalent. If the off-diagonal elements of W_1 and W_2 in the loss function (3) are equal to zero, then the limit value of $(1-\lambda)J_0$ can be expressed as

$$\lim_{\lambda\to1}(1-\lambda)J_0=\frac{1}{2}\,\mathrm{E}(L_t)\tag{A31}$$

where $\mathrm{E}(L_t)$ is the unconditional mean of the period-loss-function

$$L_t=(s_t',u_t')\begin{pmatrix}W_1&0\\0&W_2\end{pmatrix}\begin{pmatrix}s_t\\u_t\end{pmatrix}=\sum_{i=1}^{n_3}w_{ii,1}s_{i,t}^2+\sum_{i=1}^{m}w_{ii,2}u_{i,t}^2\tag{A32}$$

Then

$$\mathrm{E}(L_t)=\sum_{i=1}^{n_3}w_{ii,1}\,\mathrm{Var}\,s_{i,t}+\sum_{i=1}^{m}w_{ii,2}\,\mathrm{Var}\,u_{i,t}\tag{A33}$$

The period-loss-function can also be written as

$$L_t=Y_t'HY_t\tag{A34}$$

where $Y_t'=(k_t',u_t')$ and H as defined in (49). Then the unconditional period loss also fulfills

$$\mathrm{E}(L_t)=\mathrm{E}(Y_t'HY_t)=\mathrm{trace}(H\Sigma_{YY})\tag{A35}$$

where Σ_{YY} is the unconditional variance-covariance matrix of the vector Y.

[8]In the deterministic case, where $\Sigma_{\varepsilon\varepsilon}=0$, (A28) also holds for $\lambda=1$.

References

Backus, D., Driffill, J. (1986), The Consistency of Optimal Policy in Stochastic Rational Expectations Models. Discussion Paper No. 124, CEPR, London.

Beaudry, P., Portier. F. (2006). Stock Prices, News, and Economic Fluctuations. *American Economic Review* 96, 1293–1307.

Calvo, G. (1983). Staggered Prices in a Utility Maximizing Framework. *Journal of Monetary Economics* 12, 383–398.

Casares, M. (2006). Monetary Policy Rules in a New Keynesian Euro Area Model. *Journal of Money, Credit, and Banking* 39, 875–900.

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

Davis, J. M. (2007). News and the Term Structure in General Equilibrium. Unpublished Manuscript.

Fujiwara, I., Hirose, Y., Shintani, M. (2008). Can News Be a Major Source of Aggregate Fluctuations? A Bayesian DSGE Approach. IMES Discussion Paper 2008-E-16. Institute for Monetary and Economic Studies, Bank of Japan.

Giannoni, M. P., Woodford, M. (2004). Optimal Inflation-Targeting Rules. In *The Inflation Targeting Debate*, edited by B. S. Bernanke and M. Woodford. University of Chicago Press.

Klein, P. (2000), Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics & Control* 24, 1405–1423.

Rudebusch, G., Svensson, L. E. O. (1999), Policy Rules for Inflation Targeting. In *Monetary Policy Rules*, edited by J. B. Taylor. University of Chicago Press.

Schmitt-Grohé, S., Uribe, M. (2008). What's News in Business Cycles. NBER Working Paper 14215, National Bureau of Economic Research.

Smets, F., Wouters, R. (2003), An Estimated Dynamic Stochastic General Equilibrium Model for the Euro Area. *Journal of the European Economic Association* 1, 1123–1175.

Söderlind, P. (1999), Solution and Estimation of RE Macromodels with Optimal Policy. *European Economic Review* 43, 813–823.

Wohltmann, H.-W., Winkler, R. (2008). On the Non-Optimality of Information: An Analysis of the Welfare Effects of Anticipated Shocks in the New Keynesian Model. Economics Working Paper 2008-21. Department of Economics, Christian-Albrechts-Universität Kiel.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.