Universität Hamburg

Department Economics and Politics

A Detailed Derivation of the Sticky Price and Sticky Information New Keynesian DSGE Model

Jan-Oliver Menz, Lena Vogel

DEP Discussion Papers
Macroeconomics and Finance Series
2/2009

Hamburg, 2009

A Detailed Derivation of the Sticky Price and Sticky Information New Keynesian DSGE Model*

Jan-Oliver Menz, Lena Vogel

July 14, 2009

Abstract

This paper aims at providing macroeconomists with a detailed exposition of the New Keynesian DSGE model. Both the sticky price version and the sticky information variant are derived mathematically. Moreover, we simulate the models, also including lagged terms in the sticky price version, and compare the implied impulse response functions. Finally, we present solution methods for DSGE models, and discuss three important theoretical assumptions.

Keywords: New Keynesian Model, Sticky Prices, Sticky Information, Solution Algorithms

JEL classification: E0, E20, C61, C62, C63

*We thank Jana Gettel and Torsten Schünemann for excellent research support, and Ingrid Größl, Oskar von dem Hagen and Ulrich Fritsche for very helpful comments and discussions. This research was supported by the Deutsche Forschungsgemeinschaft. We are responsible for all remaining errors.

Contents

1 The Sticky Price Model

The standard version of the New Keynesian Model is discussed in detail by Clarida et al. (1999), however, without giving a full derivation of the IS curve and the Phillips curve. This is included in Walsh (2003), page 232 onwards, whose presentation we adopt as well.

1.1 Households' Decisions

The first part of the model describes households' behavior with regard to consumption spending and utility maximization. Note that this decision problem consists of two parts: households minimize the costs of buying the composite consumption good C_t and maximize their lifetime utility depending on consumption, money holdings and leisure.

Households' Cost Minimization Problem:

The composite consumption good C_t consists of differentiated goods c_{jt} produced by firms j. It is defined as

$$C_t = \left[\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right]^{\frac{\theta}{\theta-1}}, \quad \theta > 1 \tag{1}$$

θ gives the price elasticity of demand for the individual goods. Households try to minimize the costs of achieving the level of the composite consumption good by finding the least expensive combination of individual goods c_{jt}. With p_{jt} as the prices of the individual goods, this can be written mathematically as

$$\min_{c_{jt}} \quad \int_0^1 p_{jt} c_{jt} dj \tag{2}$$

s.t.

$$\left[\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right]^{\frac{\theta}{\theta-1}} \geq C_t. \tag{3}$$

To solve this problem, we form a Lagrangian

$$L = \int_0^1 p_{jt} c_{jt} dj - \psi_t \left[\left(\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right)^{\frac{\theta}{\theta-1}} - C_t\right], \tag{4}$$

which gives the first order condition (FOC) for c_{jt}:

$$p_{jt} - \psi_t \left[\frac{\theta}{\theta-1}\left(\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right)^{\frac{\theta}{\theta-1}-1} \frac{\theta-1}{\theta}\left(c_{jt}^{\frac{\theta-1}{\theta}-1}\right)\right] = 0$$

$$p_{jt} - \psi_t \left[\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right]^{\frac{1}{\theta-1}} c_{jt}^{-\frac{1}{\theta}} = 0 \tag{5}$$

Rearranging, using $^{-\theta}$ on both sides and applying the definition for the composite consumption good in (1) yields:

$$\begin{aligned}
c_{jt}^{-\frac{1}{\theta}} &= \frac{p_{jt}}{\psi_t \left[\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right]^{\frac{1}{\theta-1}}} \\
c_{jt}^{-\frac{1}{\theta}} &= \frac{p_{jt}}{\psi_t} \left[\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right]^{-\frac{1}{\theta-1}} \\
c_{jt} &= \left(\frac{p_{jt}}{\psi_t}\right)^{-\theta} \left[\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right]^{\frac{\theta}{\theta-1}} \\
c_{jt} &= \left(\frac{p_{jt}}{\psi_t}\right)^{-\theta} C_t \qquad (6)
\end{aligned}$$

This reformulated FOC can then be substituted again into the equation for the composite consumption good (1). Solving for the Lagrangian multiplier ψ_t gives:

$$\begin{aligned}
C_t &= \left[\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj\right]^{\frac{\theta}{\theta-1}} \\
C_t &= \left[\int_0^1 \left[\left(\frac{p_{jt}}{\psi_t}\right)^{-\theta} C_t\right]^{\frac{\theta-1}{\theta}} dj\right]^{\frac{\theta}{\theta-1}} \\
C_t &= \left(\frac{1}{\psi_t}\right)^{-\theta} \left[\int_0^1 p_{jt}^{1-\theta} dj\right]^{\frac{\theta}{\theta-1}} C_t \\
\psi_t^{-\theta} &= \left[\int_0^1 p_{jt}^{1-\theta} dj\right]^{\frac{\theta}{\theta-1}} \\
\psi_t &= \left[\int_0^1 p_{jt}^{1-\theta} dj\right]^{\frac{1}{1-\theta}} \equiv P_t \qquad (7)
\end{aligned}$$

Thus, the Lagrangian multiplier gives the aggregate price index P_t for consumption as the integral over the prices of the individual goods. Using this definition of the aggregate price index in the reformulated FOC in (6) gives the demand for good j:

$$c_{jt} = \left(\frac{p_{jt}}{P_t}\right)^{-\theta} C_t \qquad (8)$$

Recall that θ stands for the price elasticity of demand for good j. As $\theta \to \infty$, individual goods become closer substitutes and individual firms have less market power.

Households' Utility Maximization Problem:

The second step of households' decisions consists of maximizing lifetime utility subject to a period-by-period budget constraint. Using a constant relative risk aversion utility function (CRRA), the representative household's lifetime utility can be written as

$$\max_{U} \quad U = E_t \sum_{i=0}^{\infty} \beta^i \left[\frac{C_{t+i}^{1-\sigma}}{1-\sigma} + \frac{\gamma}{1-b} \left(\frac{M_{t+i}}{P_{t+i}} \right)^{1-b} - \chi \frac{N_{t+i}^{1+\eta}}{1+\eta} \right], \tag{9}$$

where C_t is the composite consumption good as in equation (1), N_t is time devoted to employment, hence $1-N_t$ is leisure, β^i is the exponential discount factor, M_t/P_t are real money balances, and $\eta = \frac{1}{\psi}$, where ψ is the Frisch elasticity of labor supply. Since we use a CRRA utility function, the parameter σ gives the degree of relative risk aversion, and $1/\sigma$ the elasticity of intertemporal substitution.

The household's period-by-period budget constraint in real terms is given by:

$$C_t + \frac{M_t}{P_t} + \frac{B_t}{P_t} = \left(\frac{W_t}{P_t} \right) N_t + \frac{M_{t-1}}{P_t} + (1+i_{t-1}) \left(\frac{B_{t-1}}{P_t} \right) + \Pi_t \tag{10}$$

Thus, the household can use his wealth in each period for consumption C_t, real money holdings M_t/P_t or for buying bonds B_t/P_t. His wealth consists of real wages W_t/P_t earned from labor N_t, real money holdings from the previous period M_{t-1}/P_t, the nominal interest gain from bond holdings from the previous period, $(1+i_{t-1})(B_{t-1}/P_t)$, and from real profits received from firms, Π_t.

Then, maximizing (9) s.t. (10) by choosing C_t, M_t, B_t, and N_t via the Lagrangian

$$\begin{aligned} L &= E_t \sum_{i=0}^{\infty} \beta^i \left[\frac{C_{t+i}^{1-\sigma}}{1-\sigma} + \frac{\gamma}{1-b} \left(\frac{M_{t+i}}{P_{t+i}} \right)^{1-b} - \chi \frac{N_{t+i}^{1+\eta}}{1+\eta} \right] \\ &- \sum_{i=0}^{\infty} \lambda_t \left[C_t + \frac{M_t}{P_t} + \frac{B_t}{P_t} - \left(\frac{W_t}{P_t} \right) N_t - \frac{M_{t-1}}{P_t} - (1+i_{t-1}) \left(\frac{B_{t-1}}{P_t} \right) - \Pi_t \right] \end{aligned} \tag{11}$$

gives the following FOCs:

$$C_t: \quad C_t^{-\sigma} - \lambda_t = 0 \Leftrightarrow \lambda_t = C_t^{-\sigma} \tag{12}$$

$$C_{t+1}: \quad E_t \left[\beta C_{t+1}^{-\sigma} \right] - \lambda_{t+1} = 0 \Leftrightarrow \lambda_{t+1} = E_t \left[\beta C_{t+1}^{-\sigma} \right] \tag{13}$$

$$B_t: \quad -\frac{\lambda_t}{P_t} + E_t \left[\lambda_{t+1} \frac{(1+i_t)}{P_{t+1}} \right] = 0 \tag{14}$$

$$M_t: \quad \gamma \left(\frac{M_t}{P_t} \right)^{-b} - \frac{\lambda_t}{P_t} + E_t \left[\lambda_{t+1} \frac{1}{P_{t+1}} \right] = 0 \tag{15}$$

$$N_t: \quad -\chi N_t^{\eta} + \left(\frac{W_t}{P_t} \right) \lambda_t = 0 \tag{16}$$

These conditions can be simplified further. Using (12) and (13) in (14) gives the **Euler equation** for the optimal intertemporal allocation of consumption.:

$$\begin{aligned} &- \frac{C_t^{-\sigma}}{P_t} + E_t \left[\beta C_{t+1}^{-\sigma} \frac{(1+i_t)}{P_{t+1}} \right] = 0 \\ &\Leftrightarrow C_t^{-\sigma} = \beta (1+i_t) P_t E_t \left[\frac{1}{P_{t+1}} C_{t+1}^{-\sigma} \right] \end{aligned} \tag{17}$$

Next, note that the first line in the Euler equation in (17) can be rearranged to give the following expression:

$$\begin{aligned} -\frac{C_t^{-\sigma}}{P_t} + (1+i_t)E_t\left[\beta C_{t+1}^{-\sigma}\frac{1}{P_{t+1}}\right] &= 0 \\ E_t\left[\beta C_{t+1}^{-\sigma}\frac{1}{P_{t+1}}\right] &= \frac{1}{1+i_t}\frac{C_t^{-\sigma}}{P_t} \end{aligned} \tag{18}$$

The left hand side of (18) can then be substituted for $E_t\left[\lambda_{t+1}\frac{1}{P_{t+1}}\right]$ in (15) due to (13). Using this together with (12) in (15), and setting the price index $P_t = 1$, yields the intratemporal optimality condition setting the marginal rate of substitution between money and consumption equal to the opportunity cost of holding money:

$$\begin{aligned} \gamma\left(\frac{M_t}{P_t}\right)^{-b} - \frac{C_t^{-\sigma}}{P_t} + \frac{1}{1+i_t}\frac{C_t^{-\sigma}}{P_t} &= 0 \\ \gamma\left(\frac{M_t}{P_t}\right)^{-b} &= \frac{C_t^{-\sigma}}{P_t} - \frac{1}{1+i_t}\frac{C_t^{-\sigma}}{P_t} \\ \frac{\gamma\left(\frac{M_t}{P_t}\right)^{-b}}{C_t^{-\sigma}} &= \frac{1}{P_t} - \frac{1}{1+i_t}\frac{1}{P_t} \\ \frac{\gamma\left(\frac{M_t}{P_t}\right)^{-b}}{C_t^{-\sigma}} &= \left(1 - \frac{1}{1+i_t}\right)\frac{1}{P_t} \\ \frac{\gamma\left(\frac{M_t}{P_t}\right)^{-b}}{C_t^{-\sigma}} &= \left(\frac{i_t}{1+i_t}\right) \end{aligned} \tag{19}$$

Finally, using (12) in (16) gives the intratemporal optimality condition setting the marginal rate of substitution between leisure and consumption equal to the real wage:

$$\begin{aligned} -\chi N_t^{\eta} &= -\left(\frac{W_t}{P_t}\right)\lambda_t \\ \frac{\chi N_t^{\eta}}{\lambda_t} &= \left(\frac{W_t}{P_t}\right) \\ \frac{\chi N_t^{\eta}}{C_t^{-\sigma}} &= \left(\frac{W_t}{P_t}\right) \end{aligned} \tag{20}$$

1.2 Firms' Decisions

The second part of the model consists of firms' decisions. Firms try to minimize the costs of production and maximize profits.

Firms' Cost Minimization:

Assuming that labor is the only factor of production, firms j minimize costs by choosing the lowest possible level of labor subject to producing the firm specific good c_{jt}, which results from the production function. Mathematically, one has

$$\min_{N_t} \quad \left(\frac{W_t}{P_t}\right) N_t \tag{21}$$

s.t.

$$c_{jt} = Z_t N_{jt}, \tag{22}$$

where the variable Z_t in the production function (22) is aggregate productivity which is assumed to be stochastic with $E(Z_t) = 1$. Here, we follow Walsh (2003) and assume a constant returns to scale technology.

Using the Lagrangian

$$L = \left(\frac{W_t}{P_t}\right) N_t + \varphi_t \left(c_{jt} - Z_t N_{jt}\right) \tag{23}$$

gives the FOC:

$$\left(\frac{W_t}{P_t}\right) - \varphi_t Z_t = 0 \Leftrightarrow \varphi_t = \frac{\frac{W_t}{P_t}}{Z_t}, \tag{24}$$

where φ_t denote firms' real marginal costs. Thus, we find that firms' real marginal costs in a flexible price equilibrium equal the real wage divided by the marginal product of labor, Z_t.

Firms' Profit Maximization:

In a second step, firms maximize profits, given by income from selling the individual good c_{jt} minus the costs of producing this product, $\varphi_t c_{jt}$, by setting their prices p_{jt} for their individual goods subject to the demand curve for their individual good given by (8) and the assumption that prices are sticky. Following Calvo (1983), in each period, a fraction ω of firms is not able to change its price and has to stick to the price chosen in the previous period. Mathematically, one can express this profit maximization problem as

$$\max_{p_{jt}} E_t \sum_{i=0}^{\infty} \omega^i \Delta_{i,t+i} \left[\left(\frac{p_{jt}}{P_{t+i}}\right) c_{jt+i} - \varphi_{t+i} c_{jt+i} \right] \tag{25}$$

s.t.

$$c_{jt} = \left(\frac{p_{jt}}{P_t}\right)^{-\theta} C_t \tag{8}$$

and the assumption of Calvo pricing. Note that the appropriate discount factor in (25) is given by $\Delta_{i,t+i} = \beta^i \left[\frac{C_{t+i}}{C_t}\right]^{-\sigma}$, since firms have to take into account the future demand elasticities when setting prices.

Substituting the demand curve (8) in (25) and using Calvo pricing leads to

$$
\begin{aligned}
& \max_{p_{jt}} \quad E_t \sum_{i=0}^{\infty} \omega^i \Delta_{i,t+i} \left[\left(\frac{p_{jt}}{P_{t+i}} \right) \left(\frac{p_{jt}}{P_{t+i}} \right)^{-\theta} C_{t+i} - \varphi_{t+i} \left(\frac{p_{jt}}{P_{t+i}} \right)^{-\theta} C_{t+i} \right] \\
\Leftrightarrow \quad & \max_{p_{jt}} \quad E_t \sum_{i=0}^{\infty} \omega^i \Delta_{i,t+i} \left[\left(\frac{p_{jt}}{P_{t+i}} \right)^{1-\theta} - \varphi_{t+i} \left(\frac{p_{jt}}{P_{t+i}} \right)^{-\theta} \right] C_{t+i} \\
\Leftrightarrow \quad & \max_{p_{jt}} \quad E_t \sum_{i=0}^{\infty} \omega^i \Delta_{i,t+i} \left[\left(\frac{1}{P_{t+i}} \right)^{1-\theta} p_{jt}^{1-\theta} - \varphi_{t+i} \left(\frac{1}{P_{t+i}} \right)^{-\theta} p_{jt}^{-\theta} \right] C_{t+i} \qquad (26)
\end{aligned}
$$

Note that p_{jt} is not moved forward to p_{jt+i} since firms choose their price in the current period under the constraint that they might not be able to change this price in future periods. Calculating the FOC and denoting the optimal price $p_{jt} = p_t^*$ yields :

$$
\begin{aligned}
& E_t \textstyle\sum_{i=0}^{\infty} \quad \omega^i \Delta_{i,t+i} \left[(1-\theta) \left(\frac{1}{P_{t+i}} \right)^{1-\theta} (p_t^*)^{-\theta} - \varphi_{t+i}(-\theta) \left(\frac{1}{P_{t+i}} \right)^{-\theta} (p_t^*)^{-\theta-1} \right] C_{t+i} \\
\Leftrightarrow \quad & E_t \textstyle\sum_{i=0}^{\infty} \quad \omega^i \Delta_{i,t+i} \left[(1-\theta) \left(\frac{1}{P_{t+i}} \right)^{1-\theta} (p_t^*)^{-\theta} + \theta \varphi_{t+i} \left(\frac{1}{P_{t+i}} \right)^{-\theta} (p_t^*)^{-\theta} \frac{1}{p_t^*} \right] C_{t+i} \\
\Leftrightarrow \quad & E_t \textstyle\sum_{i=0}^{\infty} \quad \omega^i \Delta_{i,t+i} \left[(1-\theta) \left(\frac{1}{P_{t+i}} \right)^{1-\theta} p_t^* + \theta \varphi_{t+i} \left(\frac{1}{P_{t+i}} \right)^{-\theta} \right] \frac{1}{p_t^*} (p_t^*)^{-\theta} C_{t+i} \\
\Leftrightarrow \quad & E_t \textstyle\sum_{i=0}^{\infty} \quad \omega^i \Delta_{i,t+i} \left[(1-\theta) \frac{1}{P_{t+i}} \left(\frac{1}{P_{t+i}} \right)^{-\theta} p_t^* + \theta \varphi_{t+i} \left(\frac{1}{P_{t+i}} \right)^{-\theta} \right] \frac{1}{p_t^*} (p_t^*)^{-\theta} C_{t+i} \\
\Leftrightarrow \quad & E_t \textstyle\sum_{i=0}^{\infty} \quad \omega^i \Delta_{i,t+i} \left[(1-\theta) \left(\frac{p_t^*}{P_{t+i}} \right) + \theta \varphi_{t+i} \right] \left(\frac{p_t^*}{P_{t+i}} \right)^{-\theta} \left(\frac{1}{p_t^*} \right) C_{t+i} = 0 \qquad (27)
\end{aligned}
$$

Using $\Delta_{i,t+i} = \beta^i \left[\frac{C_{t+i}}{C_t} \right]^{-\sigma}$ and rearranging gives:

$$
\begin{aligned}
&E_t \quad \sum_{i=0}^{\infty} \omega^i \beta^i \left[\frac{C_{t+i}}{C_t}\right]^{-\sigma} \left[(1-\theta)\left(\frac{p_t^*}{P_{t+i}}\right) + \theta\varphi_{t+i}\right] \left(\frac{p_t^*}{P_{t+i}}\right)^{-\theta} \left(\frac{1}{p_t^*}\right) C_{t+i} = 0 \\
&E_t \quad \sum_{i=0}^{\infty} \omega^i \beta^i \left(\frac{C_{t+i}}{C_t}\right)^{-\sigma} C_{t+i} \left(\frac{1}{p_t^*}\right) \left(\frac{p_t^*}{P_{t+i}}\right)^{-\theta} \left[(1-\theta)\left(\frac{p_t^*}{P_{t+i}}\right) + \theta\varphi_{t+i}\right] = 0 \\
&\Leftrightarrow \quad E_t \sum_{i=0}^{\infty} \omega^i \beta^i \left(\frac{C_{t+i}^{-\sigma}}{C_t^{\sigma}}\right) C_{t+i} \left(\frac{1}{p_t^*}\right) \left(\frac{p_t^*}{P_{t+i}}\right)^{-\theta} \left[(1-\theta)\left(\frac{p_t^*}{P_{t+i}}\right)\right] \\
&= \quad -E_t \sum_{i=0}^{\infty} \omega^i \beta^i \left(\frac{C_{t+i}^{-\sigma}}{C_t^{\sigma}}\right) C_{t+i} \left(\frac{1}{p_t^*}\right) \left(\frac{p_t^*}{P_{t+i}}\right)^{-\theta} \theta\varphi_{t+i} \\
&\Leftrightarrow \quad (1-\theta) E_t \sum_{i=0}^{\infty} \omega^i \beta^i \left(\frac{C_{t+i}^{1-\sigma}}{C_t^{\sigma}}\right) \left(\frac{1}{p_t^*}\right) \left(\frac{p_t^*}{P_{t+i}}\right)^{1-\theta} \\
&= \quad -\theta E_t \sum_{i=0}^{\infty} \omega^i \beta^i \left(\frac{C_{t+i}^{1-\sigma}}{C_t^{\sigma}}\right) \left(\frac{1}{p_t^*}\right) \left(\frac{p_t^*}{P_{t+i}}\right)^{-\theta} \varphi_{t+i} \\
&\Leftrightarrow \quad (1-\theta) p_t^* 1 - \theta E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \left(\frac{1}{P_{t+i}}\right)^{1-\theta} = -\theta p_t^* - \theta E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \varphi_{t+i} \left(\frac{1}{P_{t+i}}\right)^{-\theta} \\
&\Leftrightarrow \quad p_t^* \left(\theta - 1\right) E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \left(P_{t+i}\right)^{\theta-1} = \theta E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \varphi_{t+i} \left(P_{t+i}\right)^{\theta} \\
&\Leftrightarrow \quad p_t^* = \left(\frac{\theta}{\theta-1}\right) \frac{E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \varphi_{t+i} \left(P_{t+i}\right)^{\theta}}{E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \left(P_{t+i}\right)^{\theta-1}} \\
&\Leftrightarrow \quad \left(\frac{p_t^*}{P_t}\right) = \left(\frac{\theta}{\theta-1}\right) \frac{E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \varphi_{t+i} \left(\frac{P_{t+i}}{P_t}\right)^{\theta}}{E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \left(\frac{P_{t+i}}{P_t}\right)^{\theta-1}}
\end{aligned} \tag{28}
$$

This is the optimal price setting rule for firms facing sticky prices. We thus find that firms optimally set their price according to the relation of discounted future costs and revenues, multiplied by the mark-up $\frac{\theta}{\theta-1}$.

1.3 Flexible price equilibrium output

If all firms can adjust prices in every period, i.e., if $\omega = 0$, (28) reduces to

$$\left(\frac{p_t^*}{P_t}\right) = \left(\frac{\theta}{\theta-1}\right) \varphi_t = \mu\varphi_t, \tag{29}$$

with μ as mark-up.

Under flexible prices, all firms charge the same price, thus $p_t^* = P_t$ and $\varphi_t = \frac{1}{\mu}$. Combining this with the definition of real marginal costs in (24) gives:

$$
\begin{aligned}
\frac{1}{\mu} &= \frac{\frac{W_t}{P_t}}{Z_t} \\
\frac{Z_t}{\mu} &= \frac{W_t}{P_t}
\end{aligned} \tag{30}
$$

Moreover, from households' optimality condition (20) with regard to leisure and the real wage, it must hold that (under flexible prices)

$$\frac{W_t}{P_t} = \frac{Z_t}{\mu} = \frac{\chi N_t^{\eta}}{c_t^{-\sigma}}. \tag{31}$$

In a flexible price equilibrium, the real wage thus equals its marginal product divided by the mark-up and is also equal to the marginal rate of substitution between leisure and consumption.

Next, define $\widehat{x}_t = \ln x_t - \ln\overline{x}$ as percentage deviation of a variable X_t from its steady state $\overline{X}$ and let superscript f denote the flexible price equilibrium. Sine the steady state is constant by assumption, its logarithm is zero and one gets $\widehat{x}_t = \ln x_t$. Then one can write (31) as approximation around the steady state, while suppressing the constants χ and μ:

$$\eta\widehat{n}_t^f + \sigma\widehat{c}_t^f = \widehat{z}_t \tag{32}$$

Similarly, one can write the production function (22) as log-linearized deviations as:

$$\widehat{c}_t^f = \widehat{n}_t^f + \widehat{z}_t \tag{33}$$

Since output is assumed to be equal to consumption, using $\widehat{y}_t^f = \widehat{c}_t^f$ together with (32) and (33) gives:

$$\begin{aligned} \eta\left[\widehat{y}_t^f - \widehat{z}_t\right] + \sigma\widehat{y}_t^f &= \widehat{z}_t \\ \eta\widehat{y}_t^f + \sigma\widehat{y}_t^f &= (1+\eta)\widehat{z}_t \\ \widehat{y}_t^f &= \frac{1+\eta}{\eta+\sigma}\widehat{z}_t \end{aligned} \tag{34}$$

This is the flexible price equilibrium output.

1.4 Derivation of the IS Curve

The standard New Keynesian IS curve is derived by log-linearizing the Euler equation in (17):

$$C_t^{-\sigma} = \beta(1+i_t)E_t\left[\frac{P_t}{P_{t+1}}C_{t+1}^{-\sigma}\right] \tag{17}$$

Again use log-linearization around the steady state, and note that the inflation rate is given as $\ln P_{t+1} - \ln P_t = \pi_{t+1}$. One then gets, using again the fact that the logarithm of a constant is zero:

$$\begin{aligned} -\sigma\ln C_t &= \ln(\beta(1+i_t)) + E_t\left(\ln P_t - \ln P_{t+1}\right) - \sigma E_t\ln C_{t+1} \\ \ln C_t &= -\frac{1}{\sigma}\ln i_t + \frac{1}{\sigma}E_t\pi_{t+1} + E_t\ln C_{t+1} \\ \widehat{c}_t &= -\frac{1}{\sigma}\widehat{i}_t + \frac{1}{\sigma}E_t\pi_{t+1} + E_t\widehat{c}_{t+1} \end{aligned} \tag{35}$$

Without investment, government expenditure and net exports, $C_t = Y_t$. Therefore, we can write the log-linearized Euler equation as an IS curve:

$$\widehat{y}_t = E_t\widehat{y}_{t+1} - \frac{1}{\sigma}\left(\widehat{i}_t - E_t\pi_{t+1}\right) \tag{36}$$

Furthermore, define the output gap as $x_t = \widehat{y}_t - \widehat{y}_t^f$, where $\widehat{y}_t = \ln y_t$ is the percentage output deviation from its steady state under Calvo pricing and $\widehat{y}_t^f = \ln y_t^f$ is the percentage output deviation from its steady state under flexible prices. Then:

$$x_t = E_t x_{t+1} - \frac{1}{\sigma}\left(\widehat{i}_t - E_t\pi_{t+1}\right) + u_t, \tag{37}$$

where $u_t \equiv E_t\widehat{y}_{t+1}^f - \widehat{y}_t^f$ is an exogenous shock driven by exogenous productivity shocks, since the flexible price equilibrium output is given in (34) as

$$\widehat{y}_t^f = \left(\frac{1+\eta}{\sigma+\eta}\right)\widehat{z}_t. \tag{34}$$

Hence, u_t is given by:

$$u_t = E_t\hat{y}_{t+1}^f - \hat{y}_t^f = E_t\left(\frac{1+\eta}{\sigma+\eta}\right)\widehat{z}_{t+1} - \left(\frac{1+\eta}{\sigma+\eta}\right)\widehat{z}_t = \left(\frac{1+\eta}{\sigma+\eta}\right)\Delta\widehat{z}_{t+1} \tag{38}$$

Note that in case when $C_t \neq Y_t$, one can add the other aggregate demand components as an additional shock, called demand shock.[1]

1.5 Derivation of the Phillips Curve

For the derivation of the New Keynesian Phillips curve in a general equilibrium framework, one uses firms' optimal price setting rule in (28):

$$\left(\frac{p_t^*}{P_t}\right) = \left(\frac{\theta}{\theta-1}\right)\frac{E_t\sum_{i=0}^{\infty}\omega^i\beta^i C_{t+i}^{1-\sigma}\varphi_{t+i}\left(\frac{P_{t+i}}{P_t}\right)^{\theta}}{E_t\sum_{i=0}^{\infty}\omega^i\beta^i C_{t+i}^{1-\sigma}\left(\frac{P_{t+i}}{P_t}\right)^{\theta-1}} \tag{28}$$

Additionally, from the definition of P_t in (7) and the assumption of Calvo pricing, note that one can write the price index as a weighted average of the newly set price p_t^* and the price index from the previous period, P_{t-1}:

$$P_t^{1-\theta} = (1-\omega)(p_t^*)^{1-\theta} + \omega P_{t-1}^{1-\theta} \tag{39}$$

Next, define $Q_t = \frac{p_t^*}{P_t}$ as the relative price chosen by all firms that adjust their price in period t and note that in the steady state, $\pi_t = \pi = 0$ and $Q_t = Q = 1$.

Dividing (39) by P_t, one gets:

$$1 = (1-\omega)Q_t^{1-\theta} + \omega\left(\frac{P_{t-1}}{P_t}\right)^{1-\theta} \tag{40}$$

Dividing (40) by $(1-\theta)$ and log-linearizing around the steady state yields:

[1] See for this Clarida et al. (1999).

$$0 = (1-\omega)\hat{q}_t - \omega\pi_t \tag{41}$$

which is equal to

$$\hat{q}_t = \left(\frac{\omega}{1-\omega}\right)\pi_t. \tag{42}$$

Next, remember from (29) that $\mu = \left(\frac{\theta}{\theta-1}\right)$ and $Q_t = \left(\frac{p_t^*}{P_t}\right)$, which together can be used to rewrite the firms' optimal price setting rule in (28) as:

$$\left[E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \left(\frac{P_{t+i}}{P_t}\right)^{\theta-1}\right] Q_t = \mu \left[E_t \sum_{i=0}^{\infty} \omega^i \beta^i C_{t+i}^{1-\sigma} \varphi_{t+1} \left(\frac{P_{t+i}}{P_t}\right)^{\theta}\right] \tag{43}$$

Both sides of this equation can now be approximated by a Taylor series. The general rule for this is:

$$\hat{x}_t\hat{y}_t\hat{z}_t = \bar{x}\bar{y}\bar{z} + \bar{x}\bar{y}(z_t - \bar{z}) + \bar{x}\bar{z}(y_t - \bar{y}) + \bar{y}\bar{z}(x_t - \bar{x}),$$

which gives in our case

$$\hat{q}_t\hat{p}_{t+i}\hat{p}_t\hat{c}_{t+i} = \bar{q}\bar{p}\bar{c} + \bar{p}\bar{c}(q_t - \bar{q}) + \bar{q}\bar{c}(p_{t+i} - \bar{p}) + \bar{q}\bar{c}(p_t - \bar{p}) + \bar{q}\bar{p}(c_{t+i} - \bar{c}),$$

where variables with a hat denote log-linear deviations from the steady state and variables with a bar denote the steady state variables. To apply this rule, we first rewrite the left-hand side of (43) in logarithmic exponentials of e as

$$\left[E_t \sum_{i=0}^{\infty} \omega^i \beta^i e^{(1-\sigma)C_{t+i}} e^{(\theta-1)(P_{t+i}-P_t)}\right] e^{q_t} \tag{44}$$

and take a Taylor approximation around C_t, P_t, and Q_t, taking into account that in the steady state, one has $\pi_t = \pi = 0$ and $q = 0$, since $Q = 1$. (Remember that $e^0 = 1$.)

$$\begin{aligned}
&\approx \quad E_t \sum_{i=0}^{\infty} \omega^i \beta^i e^{(1-\sigma)c} \times 1 \times 1 \\
&+ \quad E_t \sum_{i=0}^{\infty} \omega^i \beta^i e^{(1-\sigma)c} \times 1 \times (q_t - \bar{q}) \\
&+ \quad E_t \sum_{i=0}^{\infty} \omega^i \beta^i e^{(1-\sigma)c} \times 1 \times (\theta-1)(p_{t+i} - \bar{p}) \\
&- \quad E_t \sum_{i=0}^{\infty} \omega^i \beta^i e^{(1-\sigma)c} \times 1 \times (\theta-1)(p_t - \bar{p}) \\
&+ \quad E_t \sum_{i=0}^{\infty} \omega^i \beta^i (1-\sigma) e^{(1-\sigma)c} \times 1 \times (c_{t+i} - \bar{c})
\end{aligned} \tag{45}$$

Furthermore, approximating $\sum_{i=0}^{\infty} \omega^i \beta^i$ by $\frac{1}{1-\omega\beta}$ and collecting terms yields:

$$= \frac{e^{(1-\sigma)c}}{1-\omega\beta} + \frac{e^{(1-\sigma)c}}{1-\omega\beta}\hat{q}_t + e^{(1-\sigma)c}\sum_{i=0}^{\infty}\omega^i\beta^i\left[(1-\sigma)E_t\hat{c}_{t+i} + (\theta-1)(E_t\hat{p}_{t+i} - \hat{p}_t)\right] \quad (46)$$

Similarly, writing the right hand side of (43) in logarithmic exponentials of e and making use of a Taylor approximation around C_t, P_t, and φ_t gives:

$$\mu\left[E_t\sum_{i=0}^{\infty}\omega^i\beta^i e^{(1-\sigma)c_{t+i}}e^{\varphi_{t+i}}e^{\theta(p_{t+i}-p_t)}\right] \quad (47)$$

and

$$\begin{aligned}
\approx\ & \mu[E_t\sum_{i=0}^{\infty}\omega^i\beta^i e^{(1-\sigma)c}\varphi \\
+\ & E_t\sum_{i=0}^{\infty}\omega^i\beta^i\varphi(1-\sigma)e^{(1-\sigma)c}(c_{t+i}-\bar{c}) \\
+\ & E_t\sum_{i=0}^{\infty}\omega^i\beta^i e^{(1-\sigma)c}\varphi(\varphi_{t+i}-\bar{\varphi}) \\
+\ & E_t\sum_{i=0}^{\infty}\omega^i\beta^i e^{(1-\sigma)c}\varphi\theta(p_{t+i}-\bar{p}) \\
-\ & E_t\sum_{i=0}^{\infty}\omega^i\beta^i e^{(1-\sigma)c}\varphi\theta(p_t-\bar{p})]
\end{aligned} \quad (48)$$

Again approximating the infinite sum and collecting terms yields:

$$= \mu\left[\left(\frac{e^{(1-\sigma)c}}{1-\omega\beta}\right)\varphi + e^{(1-\sigma)c}\varphi\sum_{i=0}^{\infty}\omega^i\beta^i\left[E_t\hat{\varphi}_{t+i} + (1-\sigma)E_t\hat{c}_{t+i} + \theta(E_t\hat{p}_{t+i}-\hat{p}_t)\right]\right] \quad (49)$$

Setting (46) equal to (49)

$$\begin{aligned}
&\frac{e^{(1-\sigma)c}}{1-\omega\beta} + \frac{e^{(1-\sigma)c}}{1-\omega\beta}\hat{q}_t + e^{(1-\sigma)c}\sum_{i=0}^{\infty}\omega^i\beta^i\left[(1-\sigma)E_t\hat{c}_{t+i} + (\theta-1)(E_t\hat{p}_{t+i}-\hat{p}_t)\right] \\
= \mu&\left[\left(\frac{e^{(1-\sigma)c}}{1-\omega\beta}\right)\varphi + e^{(1-\sigma)c}\varphi\sum_{i=0}^{\infty}\omega^i\beta^i\left[E_t\hat{\varphi}_{t+i} + (1-\sigma)E_t\hat{c}_{t+i} + \theta(E_t\hat{p}_{t+i}-\hat{p}_t)\right]\right]
\end{aligned} \quad (50)$$

and noting that $\mu\varphi = Q = 1$, yields:

$$\begin{aligned}
&\left(\frac{1}{1-\omega\beta}\right)\hat{q}_t + \sum_{i=0}^{\infty}\omega^i\beta^i\left[(1-\sigma)E_t\hat{c}_{t+i} + (\theta-1)(E_t\hat{p}_{t+i}-\hat{p}_t)\right] \\
&\qquad = \sum_{i=0}^{\infty}\omega^i\beta^i\left[E_t\hat{\varphi}_{t+i} + (1-\sigma)E_t\hat{c}_{t+i} + \theta(E_t\hat{p}_{t+i}-\hat{p}_t)\right]
\end{aligned} \quad (51)$$

Canceling the terms that appear on both sides gives:

$$\left(\frac{1}{1-\omega\beta}\right)\hat{q}_t = \sum_{i=0}^{\infty}\omega^i\beta^i(E_t\hat{\varphi}_{t+i} + E_t\hat{p}_{t+i} - \hat{p}_t) \tag{52}$$

or, withdrawing $\hat{p}_t$ from the sum:

$$\left(\frac{1}{1-\omega\beta}\right)\hat{q}_t = \sum_{i=0}^{\infty}\omega^i\beta^i(E_t\hat{\varphi}_{t+i} + E_t\hat{p}_{t+i}) - \left(\frac{1}{1-\omega\beta}\right)\hat{p}_t \tag{53}$$

Multiplying both sides by $1-\omega\beta$ and adding $\hat{p}_t$ yields:

$$\hat{q}_t + \hat{p}_t = (1-\omega\beta)\sum_{i=0}^{\infty}\omega^i\beta^i(E_t\hat{\varphi}_{t+i} + E_t\hat{p}_{t+i}) \tag{54}$$

This expression states that the optimal price $\hat{p}_t^* \equiv \hat{q}_t + \hat{p}_t$ equals the expected discounted value of future nominal marginal costs, i.e. the right hand side of the equation. This can be written in a two period framework, where quadratic terms are dropped and $\hat{q}_{t+1}$ is substituted for $\hat{\varphi}_{t+1}$ due to the flexible price relation in (29).

$$\hat{q}_t + \hat{p}_t = (1-\omega\beta)(\hat{\varphi}_t + \hat{p}_t) + \omega\beta(E_t\hat{q}_{t+1} + E_t\hat{p}_{t+1}), \tag{55}$$

Rearranging results in

$$\begin{aligned}\hat{q}_t &= (1-\omega\beta)\hat{\varphi}_t + \omega\beta(E_t\hat{q}_{t+1} + E_t\hat{p}_{t+1} - \hat{p}_t) \\ &= (1-\omega\beta)\hat{\varphi}_t + \omega\beta(E_t\hat{q}_{t+1} + E_t\pi_{t+1}).\end{aligned} \tag{56}$$

Using the expression for the log-linearized price index in (42) for $\hat{q}_t$, one gets

$$\begin{aligned}\left(\frac{\omega}{1-\omega}\right)\pi_t &= (1-\omega\beta)\hat{\varphi}_t + \omega\beta\left[\left(\frac{\omega}{1-\omega}\right)E_t\pi_{t+1} + E_t\pi_{t+1}\right] \\ &= (1-\omega\beta)\hat{\varphi}_t + \frac{\omega^2\beta}{1-\omega}E_t\pi_{t+1} + \frac{\omega\beta(1-\omega)}{1-\omega}E_t\pi_{t+1} \\ &= (1-\omega\beta)\hat{\varphi}_t + \omega\beta\left(\frac{1}{1-\omega}\right)E_t\pi_{t+1}\end{aligned} \tag{57}$$

Then, multiplying both sides by $(1-\omega)/\omega$ yields the standard New Keynesian Phillips curve, where inflation is a function of real marginal costs and expected inflation:

$$\pi_t = \widetilde{\kappa}\hat{\varphi}_t + \beta E_t\pi_{t+1} \tag{58}$$

with

$$\widetilde{\kappa} \equiv \frac{(1-\omega)(1-\omega\beta)}{\omega} \tag{59}$$

So far, we have only derived an expression for inflation which depends on real marginal costs. Deriving the Phillips curve including the output gap can be done as follows. First, note that under flexible labor markets, the real wage is equal to the marginal rate of substitution between leisure and consumption, as it has been derived in (20):

$$\frac{\chi N_t^{\eta}}{C_t^{-\sigma}} = \left(\frac{W_t}{P_t}\right) \tag{20}$$

or in terms of percentage deviations around the steady state, using $C_t = Y_t$:

$$\hat{w}_t - \hat{p}_t = \eta \hat{n}_t + \sigma \hat{y}_t$$

Second, note that in a flexible equilibrium, firms' marginal costs are equal to the real wage divided by the marginal product of labor. Thus, expressing (24) in terms of percentage deviations around the steady state gives:

$$\hat{\varphi}_t = (\hat{w}_t - \hat{p}_t) - \hat{z}_t$$

Third, note that by using $Y_t = C_t$, $\hat{n}_t$ is given by (32) as:

$$\hat{n}_t = \frac{\hat{z}_t - \sigma \hat{y}_t}{\eta}$$

Then, one can write

$$\begin{aligned}
\hat{\varphi}_t &= (\hat{w}_t - \hat{p}_t) - \hat{z}_t \\
&= (\eta \hat{n}_t + \sigma \hat{y}_t) - (\hat{y}_t - \hat{n}_t) \\
&= \left(\eta \frac{\hat{z}_t - \sigma \hat{y}_t}{\eta} + \sigma \hat{y}_t\right) - \left(\hat{y}_t - \left(\frac{\hat{z}_t - \sigma \hat{y}_t}{\eta}\right)\right) \\
&= \hat{z}_t - \hat{y}_t + \frac{\hat{z}_t}{\eta} - \frac{\sigma \hat{y}_t}{\eta} \\
&= \left(-1 - \frac{\sigma}{\eta}\right) \hat{y}_t + \left(1 + \frac{1}{\eta}\right) \hat{z}_t \\
&= (\eta + \sigma)\hat{y}_t - (1 + \eta)\hat{z}_t \\
&= (\sigma + \eta) \left[\hat{y}_t - \left(\frac{1+\eta}{\sigma+\eta}\right) \hat{z}_t\right]
\end{aligned} \tag{60}$$

Since the flexible price equilibrium output is given in (34) as

$$\widehat{y}_t^f = \left(\frac{1+\eta}{\sigma+\eta}\right) \widehat{z}_t, \tag{34}$$

one can write (60) as

$$\hat{\varphi}_t = (\sigma + \eta)(\hat{y}_t - \hat{y}_t^f) = \gamma(\hat{y}_t - \hat{y}_t^f) \tag{61}$$

and the New Keynesian Phillips Curve in (58) becomes

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1}, \tag{62}$$

where $x_t \equiv \hat{y}_t - \hat{y}_t^f$ is the output gap between actual output and output under flexible prices, and

$$\kappa = \gamma \widetilde{\kappa} = \frac{(\sigma + \eta)(1 - \omega)(1 - \beta\omega)}{\omega} \tag{63}$$

A final step consists of allowing for a cost-push shock in the Phillips curve. Clarida et al. (2001) argue that such a shock can arise from a stochastic *wage mark-up* in imperfect labor markets leading to distortions in the optimality condition equalizing the marginal rate of substitution between leisure and consumption and the real wage. Then, (20) becomes:

$$\frac{\chi N_t^{\eta}}{C_t^{-\sigma}} e^{\mu_t^w} = \left(\frac{W_t}{P_t}\right) \tag{64}$$

Log-linearizing (64) leads to

$$\eta \hat{n}_t + \sigma \hat{c}_t + \mu_t^w = \hat{w}_t - \hat{p}_t \tag{65}$$

and to the resulting expression for real marginal costs, analog to (60)

$$\hat{\varphi}_t = (\eta \hat{n}_t + \sigma \hat{c}_t) - (\hat{y}_t - \hat{n}_t) + \mu_t^w \tag{66}$$

Then, the New Keynesian Phillips curve is given by

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + e_t, \tag{67}$$

where $e_t \equiv \tilde{\kappa} \mu_t^w$ gives the cost-push shock. Note that adding such a cost-push shock also affects the flexible price equilibrium output.[2] Alternatively, many authors simply add an additive cost-push-shock after having derived the Phillips curve in the standard way, interpreting it as an oil price shock.

2 The Sticky Information Model

In this section, we derive an alternative formulation of the New Keynesian DSGE model. Instead of sticky prices, it is now assumed that a fraction of firms and consumers do not update regularly their information on which they build expectations, due to costs of acquiring information. Microeconomic foundation for consumers acting under sticky information is given in Reis (2006a), and for firms in Reis (2006b). The general equilibrium model derived in this section builds on the working paper versions of the two articles by Mankiw and Reis (2006a,b,c, 2007). Sticky information exists for all agents and in all markets, meaning that consumers can be inattentive when planning total expenditure, firms can be inattentive when setting prices and workers can be inattentive when offering their labor to firms. However, consumers are assumed to always allocate their spending optimally across varieties of goods and firms always allocate their hiring optimally across varieties of labor. Both households and firms thus combine flexible and sticky behavior.

2.1 Households' Decisions

We assume a continuum of households who live forever and consist of both consumers and workers, distributed in the unit interval and indexed by j. Households equally own firms.

Their utility maximization is given by:

[2] See Walsh (2003), p.253.

$$\max_{U} \quad U(C_{t,j}, N_{t,j}) = E_t \sum_{i=0}^{\infty} \beta^i \left[\frac{C_{t+i,j}^{1-\sigma} - 1}{1-\sigma} - \chi \frac{N_{t+i,j}^{1+\eta}}{1+\eta} \right], \tag{68}$$

where the variables are the same as in the standard model with the only difference that the model now considers different types of households who differ according to the period when they last updated their information. Thus, $C_{t,j}$ is consumption in t by household j that differs from other household by his information set. As in the standard formulation, consumption by household j is defined as a Dixit-Stiglitz aggregator of consumption of varieties of goods indexed by i with the elasticity of substitution θ:

$$C_{t,j} = \left(\int_0^1 C_{t,j}(i)^{\frac{\theta-1}{\theta}} di \right)^{\frac{\theta}{\theta-1}} \tag{69}$$

This Dixit-Stiglitz aggregator of consumption has an associated static price index P_t from the minimization problem of households, analog to the standard model:

$$P_t = \left(\int_0^1 P_t(i)^{1-\theta} di \right)^{\frac{1}{1-\theta}} \tag{70}$$

Households face the following **budget constraint** in each period:

$$P_t C_{t,j} + B_{t,j} = W_{t,j} N_{t,j} + (1 + i_{t-1}) B_{t-1,j} + T_{t,j}, \tag{71}$$

where P_t - aggregate price level, $B_{t,j}$ - holdings of nominal bonds, $W_{t,j}$ - nominal wages, i_{t-1} - nominal net return at t on a bond purchased in $t-1$ and $T_{t,j}$ - lump-sum nominal transfers from profits of firms and insurance contracts to ensure that all households start with the same wealth each period.

The budget constraint can be expressed in terms of real wealth $A_{t,j}$ as follows:

$$A_{t,j} \equiv \frac{W_{t,j} N_{t,j} + (1 + i_{t-1}) B_{t-1,j} + T_{t,j}}{P_t}, \tag{72}$$

which yields for the series of budget constraints:

$$C_{t,j} + \frac{B_{t,j}}{P_t} = A_{t,j} \tag{73}$$

This optimization problem can be solved by using dynamic programming. The parameter δ gives the probability that households can update their information in any given period. Assuming no stickiness in information, thus $\delta = 1$ for all j households, equation (74) below would collapse to the standard Bellman equation giving the value function as the sum of current consumption and discounted expected consumption one period ahead.

Excursion: Dynamic Programming in a Standard DSGE Model

Derivations in this box follow the chapters 2 and 4 in Woodford (2003). Households face the following dynamic optimization problem:

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t, m_t) \right) \tag{1}$$

s.t. the period-by-period budget constraint

$$M_{t-1} + (1 + i_{t-1})B_{t-1} + P_t Y_t + T_t = P_t C_t + M_t + B_t \tag{2}$$

To solve this problem with dynamic programming, first rewrite the budget constraint in real terms as

$$a_{t-1} + y_t + \tau_t = c_t + \frac{i_t}{1+i_t} m_t + \frac{1}{1+r_t} a_t$$

where

$$\begin{aligned} a_{t-1} &\equiv \frac{M_{t-1} + (1+i_{t-1})B_{t-1}}{P_t} \\ \tau_t &\equiv \frac{T_t}{P_t} \\ 1 + r_t &\equiv (1+i_t)\frac{P_t}{\tilde{P}_{t+1}} \end{aligned}$$

With a_t as the state variable, one can calculate the value function $V_t(a_{t-1})$ as the maximized value of the household's expected lifetime utility, conditional on a given a_{t-1}.[1]
Setting up the Bellman equation as recursive solution, assuming that the last relationship is known, gives:

$$V(a_{t-1}) = \max_{c_t, m_t} \left(u(c_t, m_t) + \beta E_t(V_t(a_t))\right)$$

s.t. the budget constraint.
Calculating the FOC's, using the chain rule, gives:

$$\begin{aligned} u_c(c_t, m_t) \quad &+\beta E_t \left(V'_{t+1}(a_t)[-(1+r_t)] \right) = 0 \\ u_m(c_t, m_t) \quad &+\beta E_t \left(V'_{t+1}(a_t)[-(1+r_t) i_t (1+i_t)^{-1}] \right) = 0 \end{aligned}$$

Simplifying results in

$$\begin{aligned} u_c(c_t, m_t) &= \beta E_t \left(V'_{t+1}(a_t)[1+r_t] \right) \\ u_m(c_t, m_t) &= \beta i_t (1+i_t)^{-1} E_t \left(V'_{t+1}(a_t)[1+r_t] \right) \end{aligned}$$

[1] See McCandless (2008) for a formulation with C_t as the state variable.

Using the second equation in the first gives the static intratemporal optimality condition for holding money balances:

$$u_m(c_t, m_t) = i_t(1+i_t)^{-1}u_c(c_t, m_t)$$

In a second step, we have to differentiate $V_t(a_{t-1})$ in order to replace this derivative in in the FOCs. However, V_t not only depends on a_{t-1}, but also on $c_t$andm_t which themselves depend on a_{t-1}. But the envelope theorem states that the variables c_t, m_t are already chosen optimally, so that one can neglect their influence on V_t.
Thus, we get:

$$V_t'(a_{t-1}) = \beta E_t\left(V_{t+1}'(a_t)[1+r_t]\right)$$

Using this in the first FOC:

$$u_c(c_t, m_t) = V_t'(a_{t-1})$$

and substituting this equation again into the Bellman equation differentiated for a_{t-1} yields the intertemporal Euler equation:

$$u_c(c_t, m_t) = \beta E_t\left((1+r_t)u_c(c_{t+1}, m_{t+1})\right)$$

In the case of sticky information, equation (74) gives the value function for a household j who plans at period t. The first term in the bracket gives the expected discounted utility of the consumer who, with probability $(1-\delta)^i$, never updates his information in the subsequent periods. The second term gives the continuation value function for the case that the consumer can update again in the future, occurring each period with probability $\delta(1-\delta)^i$. It is important to note the change in the notation of consumption: $C_{t+i,i}$ denotes consumption at date i for a household that updates his information set in period i.

$$V(A_t) = \max_{U(C_{t+i,i})}\left\{\sum_{i=0}^{\infty}\beta^i(1-\delta)^i\frac{C_{t+i,i}^{-\sigma}-1}{1-\sigma} + \beta\delta\sum_{i=0}^{\infty}\beta^i(1-\delta)^i E_t[V(A_{t+1+i})]\right\} \tag{74}$$

s.t.

$$A_{t+1+i} = \frac{W_{t+1+i}N_{t+1+i} + T_{t+1+i}}{P_{t+1+i}} + (1+i_t+i)\frac{S_{t+i}}{P_{t+1+i}} \tag{75}$$

This budget constraints can be derived as follows. The households' budget in the second period consists of income from this period and saving from the previous period, including interest income. Note that from the definition of $A_{t,j}$ in the first period, one has

$$C_{t,j} + \frac{1}{P_t}B_{t,j} = A_{t,j}, \tag{76}$$

or

$$\frac{1}{P_t}B_{t,j} = A_{t,j} - C_{t,j}$$
$$B_{t,j} = (A_{t,j} - C_{t,j})P_t \equiv S_t \tag{77}$$

This can then be used in the forwarded version of (72) above to get:

$$A_{t+1+i} = \frac{W_{t+1+i}N_{t+1+i} + T_{t+1+i}}{P_{t+1+i}} + \frac{(1+i_{t+i})P_{t+i}}{P_{t+1+i}}(A_{t+i,j} - C_{t+i,j}) \tag{78}$$

or

$$A_{t+1+i} = \frac{W_{t+1+i,.}N_{t+1+i,.} + T_{t+1+i,.}}{P_{t+1+i}} + R_{t+i}(A_{t+i} - C_{t+i,.}), \tag{79}$$

where R_{t+i} denotes the real interest rate. It is important to emphasize the change in the subscripts in the final version of the budget constraint in equation (79). The index j used before denoted households and emphasized that consumers arrive with different resources from period j, when they last updated, in period t, which is when they update and decide again. Mankiw and Reis (2006a,c) then assume the existence of a perfect insurance market that guarantees that $A_{t,j} = A_t$, i.e., wealth is the same for all planners at the beginning of every period. They then write the budget constraint with subscripts , . to emphasize this point.

Plugging this reformulated budget constraint (79) into the Bellman equation in (74) gives the reformulated optimization problem as:

$$\begin{aligned} V(A_t) &= \max_{C_{t+i,i}}\{\sum_{i=0}^{\infty}\beta^i(1-\delta)^i\frac{C_{t+i,i}^{-\sigma}-1}{1-\sigma} \\ &+ \beta\delta\sum_{i=0}^{\infty}\beta^i(1-\delta)^i E_t\left[V\left(\frac{W_{t+1+i}N_{t+1+i}+T_{t+1+i}}{P_{t+1+i}} + R_{t+i}(A_{t+i}-C_{t+i})\right)\right]\} \end{aligned} \tag{80}$$

Calculating the first order condition with respect to $C_{t+i,i}$ yields:

$$\frac{\partial V(A_t)}{\partial C_{t+i,i}} : \beta^i(1-\delta)^i C_{t+i,i}^{-\sigma} = \beta\delta\sum_{k=i}^{\infty}\beta^k(1-\delta)^k\overline{R}_{t+i,t+1+k}E_t[V'(A_{t+1+k})] \tag{81}$$

for all $i = 0, 1, ...,$ and where $\overline{R}_{t+i,t+1+k} = \prod_{z=t+i}^{t+k} R_{z+1}$ is the compound return between two dates.

Note that one still lacks the derivative of the value function with respect to future real wealth A_{t+1+k}. This can be calculated using the envelope theorem.[3] Under the envelope condition, one can calculate $\partial V/\partial A_{t+1+k}$ as $\partial V/\partial A_t$, where one has to take into account, that one has a functional relationship such as $V(C(A))$, i.e. the value of utility depends on consumption which itself depends on wealth. Thus, one can derive the value function in (74) with respect to A_t by using the chain rule as:

[3] See for the following Bagliano and Bertola (2004), p.38 and McCandless (2008), p.50.

$$
\begin{aligned}
V_t'(A_t) &= \frac{\partial V}{\partial C_{t,i}}\frac{\partial C_{t,i}}{\partial A_t} + V_{t+1}'(A_{t+1}) - V_{t+1}'(A_{t+1})\frac{\partial C_{t,i}}{\partial A_t} \\
&= \left(\beta^i(1-\delta)^i C_{t+i,i}^{-\sigma} - \beta\delta\sum_{k=i}^{\infty}\beta^k(1-\delta)^k\overline{R}_{t+i,t+1+k}E_t[V'(A_{t+1+k})]\right)\frac{\partial C_{t,i}}{\partial A_t} \\
&+ \beta\delta\sum_{k=i}^{\infty}\beta^k(1-\delta)^k\overline{R}_{t+i,t+1+k}E_t[V'(A_{t+1+k})] \\
&= \beta\delta\sum_{k=i}^{\infty}\beta^k(1-\delta)^k\overline{R}_{t+i,t+1+k}E_t[V'(A_{t+1+k})] \qquad (82)
\end{aligned}
$$

since the first term in the brackets equals zero, which follows from the first order condition in (81). Then, for $i = 0$, setting (81), the FOC, equal to the envelope condition (82), gives

$$V'(A_t) = C_{t,0}^{-\sigma} \qquad (83)$$

These results can be used to derive two Euler equations. First, let $i = 0$ in the FOC (81), and use the forwarded equation (83), $V'(A_{t+1}) = C_{t+1,0}^{-\sigma}$, for A_{t+1}. Since this gives the Euler equation of the updating consumer, we have $\delta = 1$, and thus get the standard Euler equation as

$$C_{t,0}^{-\sigma} = \beta E_t[R_{t+1}C_{t+1,0}^{-\sigma}] \qquad (84)$$

Second, note that one can combine (81) and (82) to produce an Euler equation for the inattentive consumer. Hence, we write (81) for $i = j$ since households j differ with respect to the date when they last updated their information set:

$$\beta^j(1-\delta)^j C_{t+j,j}^{-\sigma} = \beta\delta\sum_{k=j}^{\infty}\beta^k(1-\delta)^k E_t\left[V'(A_{t+1+k})\bar{R}_{t+j,t+1+k}\right] \qquad (85)$$

and (82) for $t = t + j$

$$E_t[V'(A_{t+j})] = \beta\delta\sum_{k=0}^{\infty}\beta^k(1-\delta)^k E_{t+j}\left[V'(A_{t+j+1+k})\bar{R}_{t+j,t+j+1+k}\right] \qquad (86)$$

Note that the two terms on the right-hand side of both equations are essentially the same. This can be seen by rewriting the first equation as[4]

$$\beta^j(1-\delta)^j C_{t+j,j}^{-\sigma} = \beta^j(1-\delta)^j\left[\beta\delta\sum_{k=0}^{\infty}\beta^k(1-\delta)^k E_{t+k}\left(V'(A_{t+1+k+k})\bar{R}_{t+j,t+1+k+k}\right)\right] \qquad (87)$$

Then, note that the term in brackets equals the right-hand side of (86) (recalling that $k = j$), and hence can be set equal to $E_t[V'(A_{t+j})]$.

Using this in (83), i.e. forwarding this equation, one has

$$E_t V'(A_{t+j}) = E_t C_{t+j,0}^{-\sigma}, \qquad (88)$$

[4] See for this McCandless (2008), p.53.

which, when plugging into (87), and dividing both sides by $\beta^j(1-\delta)^j$ yields the second Euler equation:

$$C_{t+j,j}^{-\sigma} = E_t C_{t+j,0}^{-\sigma} \tag{89}$$

This is the Euler equation for an inattentive consumer who sets the marginal utility of consumption equal to his expectation of the marginal utility of the attentive consumer, when he last updated.

2.2 Derivation of the Sticky Information IS Curve

To solve for the sticky information IS curve, one can proceed as follows.

Log-linearizing the two Euler equations (84) and (89), backwarding (89) by j periods and expressing log-deviations with a hat, leads to:

$$\hat{c}_{t,0} = E_t(\hat{c}_{t+1,0} - \frac{1}{\sigma}\hat{r}_t) \tag{90}$$

and

$$\hat{c}_{t,j} = E_{t-j}(\hat{c}_{t,0}) \tag{91}$$

Then, note that total consumption in t is given by:

$$\hat{c}_t = \delta \sum_{j=0}^{\infty} (1-\delta)^j \hat{c}_{t,j} \tag{92}$$

and the log-linearized market clearing condition is given by:

$$\hat{y}_t = \hat{c}_t + \hat{u}_t, \tag{93}$$

with $\hat{u}_t$ as demand shock.

Finally, Mankiw and Reis (2006a) assume that in the limit, all consumers are fully informed. This gives

$$\lim_{i\to\infty} E_t(\hat{c}_{t+i,0}) = \lim_{i\to\infty} E_t[\hat{y}_{t+i}^f] \equiv \hat{y}_t^f \tag{94}$$

and

$$\lim_{i\to\infty} E_t(\hat{r}_{t+i}) = \lim_{i\to\infty} E_t(\hat{r}_{t+1}^n) = 0 \tag{95}$$

Then, these equations can be combined to arrive at the sticky information IS curve.

Start with (93) and use (92) for c_t:

$$\hat{y}_t = \delta \sum_{j=0}^{\infty} (1-\delta)^j \hat{c}_{t,j} + \hat{u}_t \tag{96}$$

Next, use (91)

$$\hat{y}_t = \delta \sum_{j=0}^{\infty} (1-\delta)^j E_{t-j}(\hat{c}_{t,0}) + \hat{u}_t \tag{97}$$

and (90)

$$\hat{y}_t = \delta \sum_{j=0}^{\infty} (1-\delta)^j E_{t-j}(E_t(\hat{c}_{t+1,0} - \frac{1}{\sigma}\hat{r}_t)) \tag{98}$$

Finally, using (94) and (95) gives the sticky information IS curve:

$$\begin{aligned} \hat{y}_t &= \delta \sum_{j=0}^{\infty} E_{t-j}(\hat{y}_t^f - \frac{1}{\sigma}\hat{r}_t) + \hat{u}_t \\ \hat{y}_t &= \delta \sum_{j=0}^{\infty} E_{t-j}(\hat{y}_t^f - \frac{1}{\sigma}R_t) + \hat{u}_t, \end{aligned} \tag{99}$$

where the long run real interest rate is given as $R_t = E_t(\sum_{i=0}^{\infty} r_{t+i})$. In the sticky information model, output is thus explained by past expectations of the natural or flexible price output $\hat{y}_t^f$ and the long run real interest rate R_t and can be disturbed by a demand shock $\hat{u}_t$.

2.3 Firms' Decision

Firms maximize profits given that prices are flexible, i.e., we now have a static optimization problem. Under the assumption that each period, a fraction λ of firms updates its information set, a firm that last updated j periods ago sets its price $p_{t,j}$ according to

$$\max_{p_{t,j}} E_{t-j}\left[\frac{p_{t,j}Y_{t,j}}{P_t} - \frac{W_t N_{t,j}}{P_t}\right] \tag{100}$$

s.t. the Cobb-Douglas production function with decreasing returns to scale, assuming that $\alpha < 1$,

$$Y_{t,j} = Z_t N_{t,j}^{\alpha} \tag{101}$$

and s.t. the demand for the single product $Y_{t,i}$ produced by each firm, where $\overline{C}_t$ is composite consumption, i.e. the consumption spending of all consumers, and G_t is consumption demand of the government.

$$Y_{t,i} = \left(\frac{P_t(i)}{P_t}\right)^{-\theta} \overline{C}_t G_t = \left(\frac{P_t(i)}{P_t}\right)^{-\theta} \left(\int_0^1 C_{t,j}dj\right) G_t \tag{102}$$

Using the market clearing condition, one gets for the quantity produced by each firm:

$$Y_{t,j} = \left(\frac{p_{t,j}}{P_t}\right)^{-\theta} \overline{C}_t G_t \tag{103}$$

Then, using (103) and (101) in (100), where (103) was used before in $N_{t,j} = Y_{t,j}^{\frac{1}{\alpha}} Z_t^{-\frac{1}{\alpha}}$, one can rewrite the maximization problem of firms as

$$\max_{p_{t,j}} E_{t-j}\left[\frac{p_{t,j}}{P_t}\left(\frac{p_{t,j}}{P_t}\right)^{-\theta}\overline{C}_t G_t - \frac{W_t}{P_t}\left(\left(\left(\left(\frac{p_{t,j}}{P_t}\right)^{-\theta}\overline{C}_t G_t\right)^{\frac{1}{\alpha}} Z_t^{-\frac{1}{\alpha}}\right)\right)\right] \tag{104}$$

Then, we maximize (104) by using the product rule on the first summand and the chain rule on the second summand. This yields:

$$E_{t-j}\left[\frac{Y_{t,j}}{P_t} + \frac{p_{t,j}}{P_t}\left(-\theta\frac{p_{t,j}^{-\theta-1}}{P_t^{\theta}}\right) - \frac{W_t}{P_t}\left(Z_t^{-\frac{1}{\alpha}}\frac{1}{\alpha}Y_{t,j}^{\frac{1}{\alpha}-1}\right)(-\theta)\frac{p_{t,j}^{-\theta-1}}{P_t^{-\theta}}\right] = 0 \tag{105}$$

Next, note that the derivative of $Y_{t,j}$ with respect to $p_{t,j}$ can be written as

$$\frac{\partial Y_{t,j}}{\partial p_{t,j}} = -\theta\frac{p_{t,j}^{-\theta-1}}{P_t^{-\theta}} = -\theta\frac{p_{t,j}^{-\theta}}{p_{t,j}P_t^{-\theta}} = -\theta\frac{1}{p_{t,j}}Y_{t,j} \tag{106}$$

Using this in (105) gives

$$E_{t-j}\left[\frac{Y_{t,j}}{P_t} + \frac{p_{t,j}}{P_t}\left(-\theta\frac{Y_{t,j}}{p_{t,j}}\right) - \frac{W_t}{P_t}\frac{\left(Z_t^{-\frac{1}{\alpha}}\frac{1}{\alpha}Y_{t,j}^{\frac{1}{\alpha}-1}\right)(-\theta)Y_{t,j}}{p_{t,j}}\right] = 0 \tag{107}$$

which can be simplified to get the optimal price setting rule as

$$\begin{aligned} E_{t-j}\left[(1-\theta)\frac{Y_{t,j}}{P_t} + \frac{\theta\frac{1}{\alpha}\frac{W_t}{P_t}Z_t^{-\frac{1}{\alpha}}Y_{t,j}^{\frac{1}{\alpha}}}{p_{t,j}}\right] &= 0 \\ p_{t,j} = E_{t-j}&\left[\frac{(1-\theta)\frac{Y_{t,j}}{P_t}}{-\theta\frac{1}{\alpha}\frac{W_t}{P_t}Z_t^{-\frac{1}{\alpha}}Y_t^{\frac{1}{\alpha}}}\right] \\ p_{t,j} = \frac{\theta}{\theta-1}E_{t-j}&\left(\frac{W_t Y_t^{\frac{1}{\alpha}}Z_t^{-\frac{1}{\alpha}}}{\alpha Y_{t,j}}\right) \end{aligned} \tag{108}$$

2.4 Derivation of the Sticky Information Phillips Curve

To derive the sticky information Phillips curve from this optimal price setting rule of firms in (108), one proceeds as follows. Note that the log-linearized price index is given as

$$\hat{p}_t = \lambda\sum_{t=0}^{\infty}(1-\lambda)^j\hat{p}_{t,j}, \tag{109}$$

the log-linearized version of (103) as

$$\hat{y}_{t,j} = \hat{y}_t - \theta(\hat{p}_{t,j} - \hat{p}_t), \tag{110}$$

and the log-linearized version of the price setting rule (108) as

$$\hat{p}_{t,j} = E_{t-j}\left[\hat{w}_t + \left(\frac{1}{\alpha} - 1\right)\hat{y}_{t,j} - \frac{\hat{z}_t}{\alpha}\right] \tag{111}$$

Start with plugging (110) into (111) and rearrange:

$$\begin{aligned}
\hat{p}_{t,j} &= E_{t-j}\left[\hat{w}_t + \left(\frac{1}{\alpha} - 1\right)[\hat{y}_t - \theta(\hat{p}_{t,j} - \hat{p}_t)] - \frac{\hat{z}_t}{\alpha}\right] \\
&= E_{t-j}\left[\hat{w}_t + \frac{1}{\alpha}[\hat{y}_t - \theta\hat{p}_{t,j} + \theta\hat{p}_t] - \hat{y}_t + \theta\hat{p}_{t,j} - \theta\hat{p}_t - \frac{\hat{z}_t}{\alpha}\right] \\
\hat{p}_{t,j} &+ \frac{1}{\alpha}\theta\hat{p}_{t,j} - \theta\hat{p}_{t,j} = E_{t-j}\left[\hat{w}_t + \frac{1}{\alpha}\hat{y}_t + \frac{1}{\alpha}\theta\hat{p}_t - \hat{y}_t - \theta\hat{p}_t - \frac{\hat{z}_t}{\alpha}\right] \\
\hat{p}_{t,j} &= E_{t-j}\left[\frac{\hat{w}_t + \left(\frac{1}{\alpha} - 1\right)\hat{y}_t + \theta\left(\frac{1}{\alpha} - 1\right)\hat{p}_t - \frac{\hat{z}_t}{\alpha}}{1 + \theta\left(\frac{1}{\alpha} - 1\right)}\right] \\
&= E_{t-j}\left[\frac{\hat{w}_t\alpha + (1-\alpha)\hat{y}_t + \theta(1-\alpha)\hat{p}_t - \hat{z}_t}{\alpha + \theta(1-\alpha)}\right] \\
&= E_{t-j}\left[\frac{\hat{w}_t\alpha + (1-\alpha)\hat{y}_t + \theta(1-\alpha)\hat{p}_t - \hat{z}_t + \alpha\hat{p}_t - \alpha\hat{p}_t}{\alpha + \theta(1-\alpha)}\right] \\
&= E_{t-j}\left[\frac{\hat{w}_t\alpha + \hat{p}_t(\theta(1-\alpha) + \alpha) - \alpha\hat{p}_t + (1-\alpha)\hat{y}_t - \hat{z}_t}{\alpha + \theta(1-\alpha)}\right] \\
&= E_{t-j}\left[\hat{p}_t + \frac{\alpha(\hat{w}_t - \hat{p}_t) + (1-\alpha)\hat{y}_t - \hat{z}_t}{\alpha + \theta(1-\alpha)}\right]
\end{aligned} \tag{112}$$

Finally, using this expression in (109) gives an expression for the price level under sticky information linking the current price level to the price level and real marginal costs $\hat{\varphi}_t$ expected by firms that have only updated their information in the past:

$$\begin{aligned}
\hat{p}_t &= \lambda\sum_{j=0}^{\infty}(1-\lambda)^j E_{t-j}\left[\hat{p}_t + \frac{\alpha(\hat{w}_t - \hat{p}_t) + (1-\alpha)\hat{y}_t - \hat{z}_t}{\alpha + \theta(1-\alpha)}\right] \\
&= \lambda\sum_{j=0}^{\infty}(1-\lambda)^j E_{t-j}\left[\hat{p}_t + \hat{\varphi}_t\right]
\end{aligned} \tag{113}$$

As in the case of the sticky price version, one can express the Phillips curve in terms of the output gap and add a cost push shock. This can be done by following Ball et al. (2003, 2005). Recall equation (29), the mark-up price setting of firms in the case of flexible prices and full information:

$$\left(\frac{p_t^*}{P_t}\right) = \left(\frac{\theta}{\theta - 1}\right)\varphi_t = \mu\varphi_t, \tag{29}$$

with μ as mark-up.

Using (31), one can write for the real marginal costs $\hat{\varphi}_t$:

$$\left(\frac{p_t^*}{P_t}\right) = \left(\frac{\theta}{\theta - 1}\right)\frac{C_t^\sigma N_{it}^\eta}{Z_t} \tag{114}$$

Next, using the market clearing condition $C_t = Y_t$ and the production function (22) for N_{it} gives

$$\left(\frac{p_t^*}{P_t}\right) = \left(\frac{\theta}{\theta-1}\right) \frac{Y_t^{\sigma}\left(\frac{Y_{it}}{Z_t}\right)^{\eta}}{Z_t} \tag{115}$$

Finally, using the demand equation (8) for Y_{it} yields

$$\left(\frac{p_t^*}{P_t}\right) = \left(\frac{\theta}{\theta-1}\right) \frac{Y_t^{\sigma}\left(\frac{\left(\frac{p_{it}}{P_t}\right)^{-\theta} Y_t}{Z_t}\right)^{\eta}}{Z_t} \tag{116}$$

Taking logarithms gives

$$\begin{aligned}
p_{it}^* &= p_t + \sigma y_t - \theta\eta(p_{it} - p_t) + \eta y_t - \eta z_t - z_t + \frac{\ln\left(\frac{\theta}{\theta-1}\right)}{1+\eta\theta} \\
p_{it}^* &- p_t + \theta\eta(p_{it} - p_t) = (\sigma+\eta)y_t - (1+\eta)z_t + \frac{\ln\left(\frac{\theta}{\theta-1}\right)}{1+\eta\theta} \\
p_{it}^* &= p_t + \frac{\sigma+\eta}{1+\eta\theta} y_t - \frac{(1+\eta)z_t}{1+\eta\theta} + \frac{\ln\left(\frac{\theta}{\theta-1}\right)}{1+\eta\theta} \qquad (117)
\end{aligned}$$

If markets are fully competitive, all firms set the same price, i.e. $p_{it}^* = p_t$, which gives the natural output as

$$\begin{aligned}
\frac{\sigma+\eta}{1+\eta\theta} y_t &= \frac{(1+\eta)z_t}{1+\eta\theta} - \frac{\ln\left(\frac{\theta}{\theta-1}\right)}{1+\eta\theta} \\
y_t^n &= \frac{(1+\eta)z_t - \ln\left(\frac{\theta}{\theta-1}\right)}{\sigma+\eta} \qquad (118)
\end{aligned}$$

Rearranging (118) for z_t,

$$z_t = \frac{(\sigma+\eta)y_t^n + \ln\left(\frac{\theta}{\theta-1}\right)}{1+\eta} \tag{119}$$

and substituting into (117) gives

$$\begin{aligned}
p_{it}^* &= p_t + \frac{\sigma+\eta}{1+\eta\theta} y_t - \frac{(1+\eta)}{1+\eta\theta}\left(\frac{(\sigma+\eta)y_t^n + \ln\left(\frac{\theta}{\theta-1}\right)}{1+\eta}\right) + \frac{\ln\left(\frac{\theta}{\theta-1}\right)}{1+\eta\theta} \\
p_{it}^* &= p_t + \frac{\sigma+\eta}{1+\eta\theta}(y_t - y_t^n) \qquad (119)
\end{aligned}$$

Then, using this expression in (109) gives the sticky information Phillips curve in terms of the output gap:

$$\hat{p}_t = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t], \tag{120}$$

where $e_t \equiv \tilde{\kappa}_t \mu_t^w$ and $\gamma = (\sigma+\eta)/(1+\eta\theta)$.

To transform this equation for the price level into an expression for the inflation rate, one proceeds as follows:

First, note that one can rewrite equation (120) as

$$\hat{p}_t = \lambda(p_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t) + \lambda \sum_{j=0}^{\infty}(1-\lambda)^{j+1} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] \quad (121)$$

Second, rewrite equation (120) for period $t-1$:

$$\hat{p}_{t-1} = \lambda \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_{t-1} + \gamma(\hat{y}_{t-1} - \hat{y}_{t-1}^f) + e_{t-1}] \quad (122)$$

Then, subtracting equation (122) from (121) gives

$$\begin{aligned}\hat{p}_t - \hat{p}_{t-1} &= \lambda(\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t) + \lambda \sum_{j=0}^{\infty}(1-\lambda)^{j+1} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] \\ &- \lambda \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_{t-1} + \gamma(\hat{y}_{t-1} - \hat{y}_{t-1}^f) + e_{t-1}] \quad (123)\end{aligned}$$

and, extracting the term $(1-\lambda)$,

$$\begin{aligned}\pi_t &= \lambda(\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t) + \lambda \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] \\ &- \lambda^2 \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] \\ &- \lambda \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_{t-1} + \gamma(\hat{y}_{t-1} - \hat{y}_{t-1}^f) + e_{t-1}] \\ &= \lambda(\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t) + \lambda \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\pi_t + \gamma(\Delta\hat{y}_t - \Delta\hat{y}_t^f) + \Delta e_t] \\ &- \lambda^2 \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] \quad (124)\end{aligned}$$

Next, multiply (120) by λ and rearrange

$$\begin{aligned}\lambda\hat{p}_t &= \lambda^2(\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t) + \lambda^2 \sum_{j=0}^{\infty}(1-\lambda)^{j+1} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] \\ \Leftrightarrow\ & (1-\lambda)\lambda^2 \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] = \lambda\hat{p}_t - \lambda^2\hat{p}_t - \lambda^2\gamma(\hat{y}_t - \hat{y}_t^f) - \lambda^2 e_t \\ \Leftrightarrow\ & \ldots = \lambda[(1-\lambda)\hat{p}_t - \lambda\gamma(\hat{y}_t - \hat{y}_t^f) - \lambda e_t] \\ \Leftrightarrow\ & \lambda^2 \sum_{j=0}^{\infty}(1-\lambda)^{j} E_{t-j-1}[\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t] = \lambda\hat{p}_t - \frac{\lambda^2}{1-\lambda}(\gamma(\hat{y}_t - \hat{y}_t^f) + e_t) \quad (125)\end{aligned}$$

Finally, using (125) in (124) yields the Sticky Information Phillips curve:

$$
\begin{aligned}
\pi_t &= \lambda(\hat{p}_t + \gamma(\hat{y}_t - \hat{y}_t^f) + e_t) + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j-1}[\pi_t + \gamma(\Delta\hat{y}_t - \Delta\hat{y}_t^f) + \Delta e_t] \\
&- \lambda\hat{p}_t - \frac{\lambda^2}{1-\lambda}(\gamma(\hat{y}_t - \hat{y}_t^f) + e_t) \\
&= \frac{\lambda\gamma(1-\lambda) + \lambda^2}{1-\lambda}(\hat{y}_t - \hat{y}_t^f) + \frac{\lambda(1-\lambda) + \lambda^2}{1-\lambda} e_t \\
&+ \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j-1}[\pi_t + \gamma(\Delta\hat{y}_t - \Delta\hat{y}_t^f) + \Delta e_t] \\
&= \frac{\lambda\gamma}{1-\lambda}(\hat{y}_t - \hat{y}_t^f) + \frac{\lambda}{1-\lambda} e_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j-1}[\pi_t + \gamma(\Delta\hat{y}_t - \Delta\hat{y}_t^f) + \Delta e_t]
\end{aligned}
\tag{126}
$$

The sticky information Phillips curve thus gives inflation as a function of the current output gap and a cost-push shock as well as lagged expectations of current inflation, current changes in output gap and current changes in the cost-push shock.

3 Simulating the Model With Different Specifications

In the following section, we will simulate both the sticky price New Keynesian model, including a habit formation version with lagged terms in the IS curve and the Phillips curve, and the sticky information model. It is not our aim to provide a detailed sensitivity analysis evaluating the effects of different parameter values, but simply to compare the results of the three models. To begin with, it is worth noting that a macroeconomic model should be able to replicate the following stylized facts:

1. The change in inflation is procyclical, i.e. one should find a positive correlation between the inflation rate and the output gap detrended with the HP filter (Mankiw and Reis (2006a)).

2. The impulse responses to shocks typically have a hump-shaped form, i.e. the full impact of shocks only materializes some periods after the initial occurrence of the shock (Mankiw and Reis (2006a)).

3. The simulated series for output and inflation should exhibit pronounced cycles (De Grauwe (2008)).

4. Following a monetary policy shock, inflation reacts more sluggishly than output (De Grauwe (2008)).

5. The simulated series for output and inflation are very persistent, with output being more persistent than inflation (De Grauwe (2008)).

3.1 Simulated Models

Following McCallum (2001), the Taylor rule is the same for all models, namely:

$$\hat{i}_t = \mu_\pi \Delta p_t + \mu_y x_t + \mu_R \hat{i}_{t-1} + v_t, \tag{127}$$

3.1.1 The Sticky Price Model

The sticky price IS curve is given as:

$$x_t = E_t x_{t+1} - \frac{1}{\sigma}\left(\widehat{i}_t - E_t \pi_{t+1}\right) + u_t \tag{37}$$

And the sticky price Phillips curve as:

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + e_t \tag{67}$$

Note that despite of the possibility of deriving the shocks in the IS curve and the Phillips curve from a mark-up shock, we simply add them additively to all of the three models, as it is mostly done in the literature.

3.1.2 Habit Formation and Lagged Inflation

In order to be able to replicate the high persistence in both the output gap and the inflation rate, the most common way to augment the standard model consists of including habit formation for the consumption part (Fuhrer, 2000), and a lagged term for the Phillips curve (Fuhrer and Moore, 1995).

This then gives for the hybrid IS curve

$$x_t = -\frac{1}{\sigma}(\hat{i}_t - E_t \Delta p_{t+1}) + (1-\rho) E_t x_{t+1} + \rho x_{t-1} + u_t \tag{128}$$

and for the hybrid Phillips curve:

$$\Delta p_t = (1-\iota)\beta E_t \Delta p_{t+1} + \iota \Delta p_{t-1} + \kappa x_t + e_t \tag{129}$$

3.1.3 The Sticky Information Model

The sticky information IS curve is given by:

$$\hat{y}_t = \delta \sum_{j=0}^{\infty} E_{t-j}(\hat{y}_t^f - \frac{1}{\sigma}\hat{R}_t) + \hat{u}_t \tag{99}$$

The sticky information Phillips curve was derived as:

$$\pi_t = \frac{\lambda\gamma}{1-\lambda}(\hat{y}_t - \hat{y}_t^f) + \frac{\lambda}{1-\lambda} e_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-1-j}[\pi_t + \gamma(\Delta\hat{y}_t - \Delta\hat{y}_t^f) + \Delta e_t] \tag{125}$$

3.2 Shocks and Parameter Values

The models have four different shocks.

The shock to the IS curve (demand shock):

$$u_t = \alpha_1 u_{t-1} + \epsilon_t \tag{130}$$

The shock to the Phillips curve (cost-push-shock):

$$e_t = \alpha_2 e_{t-1} + \phi_t \tag{131}$$

The monetary policy shock:

$$v_t = \alpha_3 v_{t-1} + \psi_t \tag{132}$$

The technology shock:

$$\hat{z}_t = \alpha_4 \hat{z}_{t-1} + \eta_t \tag{133}$$

For the parameter values, we take as many values as possible from McCallum (2001). Thus, we set σ, the coefficient of relative risk aversion to 2.5 and the exponential discount factor β to 0.99. McCallum (2001) uses 0.03 for the coefficient of the output gap in the Phillips Curve, which corresponds to κ in our theoretical derivation of the sticky price model and to $\delta\gamma/(1-\delta)$ in the sticky information model.

Note that κ is given by

$$\kappa = \frac{(\sigma+\eta)(1-\omega)(1-\omega\beta)}{\omega}$$

and γ in the sticky information model as

$$\gamma = \frac{\eta+\sigma}{1+\eta\theta}$$

Hence, to get a coefficient of 0.03 in both models, given the values for σ and β, we set η, the Frisch elasticity of labor supply to 1, which lies in between the very wage elastic labor supply of 0.25 used by Mankiw and Reis (2006a) and the higher value of 1.5 used by Trabandt (2007). For θ, the elasticity of supply between different goods, we use a value of 35, which gives a mark-up of 3%, and is roughly the same as the mark-up of 5% assumed by Mankiw and Reis (2006a). It then remains to set ω, the fraction of firms that cannot change prices in each period, to 0.913, to get a value for the output coefficient in both Phillip Curves which is roughly equal to 0.032. For the parameters in the Taylor rule, we use the values from Mankiw and Reis (2006a), i.e. $\mu_\pi = 1.24$, $\mu_y = 0.33$, and $\mu_R = 0.92$. As in McCallum (2001), we use 0.95 for the AR term of the technology shock and assume that the other shocks are serially uncorrelated. The standard deviations of the four shocks are also taken from McCallum (2001) and are set to 0.03, 0.002, 0.0017, and 0.007, respectively. For the habit formation model, we set ι, the share of backward looking behavior in the Phillips Curve equal to 0.5, and ρ, the corresponding term in the IS curve equal to 0.5, also following McCallum (2001). Finally, for the share of consumers and firms, that update their information set in every period, i.e. δ and λ, respectively, we use 1/4 as in Mankiw and Reis (2006a).

The parameter values are summarized again in the following table:

Table 1: Parameter Values used for Simulation

Symbol	Name	Value	Source
α_1	AR term of shock to IS curve	0	McCallum 2001
α_2	AR term of shock to Phillips curve	0	McCallum 2001
α_3	AR term of shock to Taylor rule	0	McCallum 2001
α_4	AR term of shock to natural output	0.95	McCallum 2001
σ	Coefficient of relative risk aversion	2.5	McCallum 2001
β	Subjective discount factor	0.99	McCallum 2001
ω	Fraction of firms that cannot adjust prices	0.913	Own assumption
η	Frisch elasticity of labor supply	1	Own assumption
θ	Elasticity of substitution between different goods	35	Own assumption
κ	Output elasticity in sticky price PC	0.032	McCallum 2001
γ	Output elasticity in sticky information PC	0.032	Ball et al. 2005
μ_π	Weight of deviation from inflation target in Taylor rule	1.24	Mankiw Reis 2006a
μ_y	Weight of output gap in Taylor rule	0.33	Mankiw Reis 2006a
μ_R	Weight of interest rate smoothing in Taylor rule	0.92	Mankiw Reis 2006a
τ_1	Standard deviation of IS shock	0.03	McCallum 2001
τ_2	Standard deviation of cost-push-shock	0.002	McCallum 2001
τ_3	Standard deviation of monetary policy shock	0.0017	McCallum 2001
τ_4	Standard deviation of technology shock	0.007	McCallum 2001
ρ	Lagged term in IS curve	0.5	McCallum 2001
ι	Lagged term in PC curve	0.5	McCallum 2001
δ	Share of updating consumers	0.25	Mankiw Reis 2006a
λ	Share of updating firms	0.25	Mankiw Reis 2006a

3.3 Results: Impulse Response Functions

The simulations are carried out by using the solution algorithm of Meyer-Gohde (2007) which we describe in the next section. We modified some of his Matlab files available on his homepage[5].

Comparing the impulse response functions of the three different models, while keeping in mind the stylized facts mentioned earlier, we note the following:

1. Looking at the the impulse responses of the output gap (figure 1) and the inflation rate (figure 2) to a monetary policy shock, we note that the sticky information model is able to replicate the hump-shaped behavior found in real data, as it is true for the habit formation model. However, its worth noting that the extent of the hump-shaped behavior of the inflation rate depends on the coefficients of the Taylor rule. With the parameter values used by McCallum (2001), the response of the inflation rate becomes much less hump-shaped. Finally, the sticky information model is not able to replicate the stylized fact that the inflation rate reacts more sluggishly than the output gap to a monetary policy intervention.

2. With respect to the IS shock and the PC shock, it is interesting to see that the impulse responses of the inflation rate and the output gap in the sticky information do not differ much from those in the standard model. This stems from the fact that we did not allow the shocks to be autocorrelated, doing so would give impulse responses very similar to the ones in Mankiw and Reis (2006a) and Arslan (2007). However, we followed McCallum (2001) who argued that the theoretical rationale to introduce persistence into the model via the shock term is quite weak.

[5] http://anna.ww.tu-berlin.de/~makro/Meyer-Gohde.html

Figure 1: Impulse Responses of Output Gap

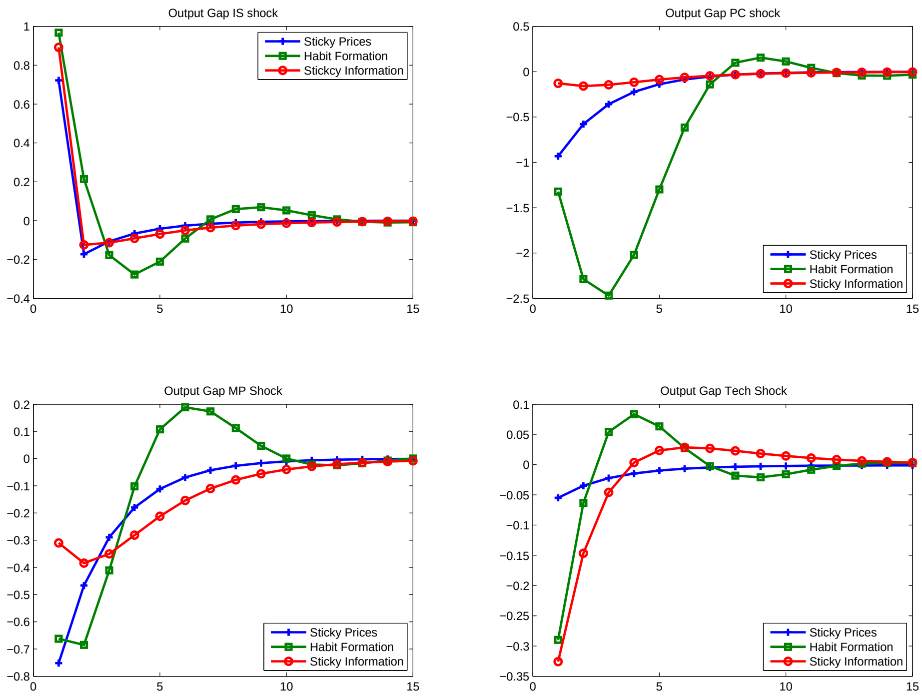


Figure 2: Impulse Responses of Inflation Rate

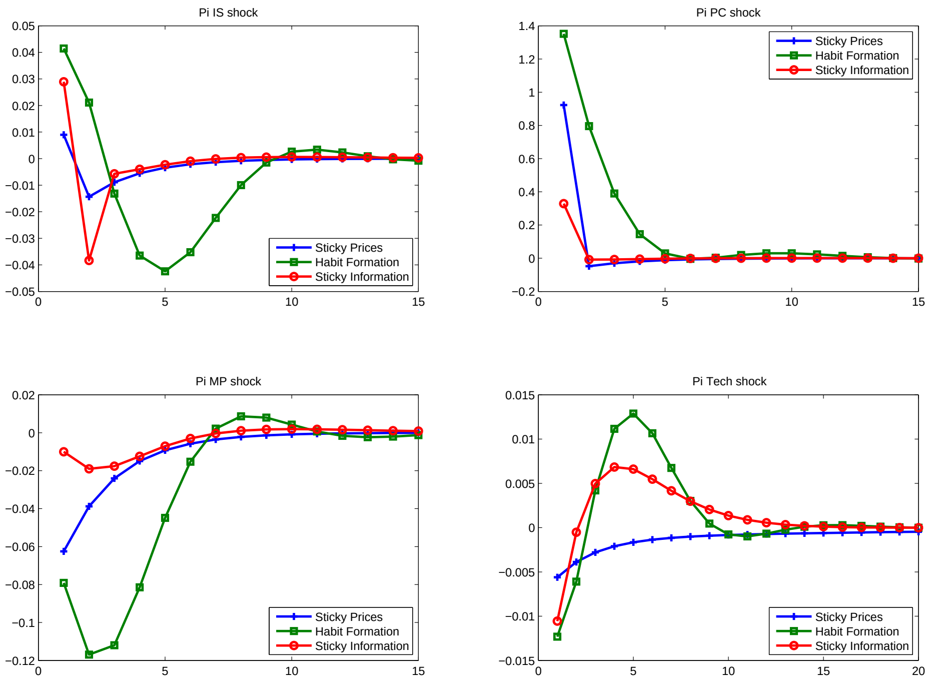
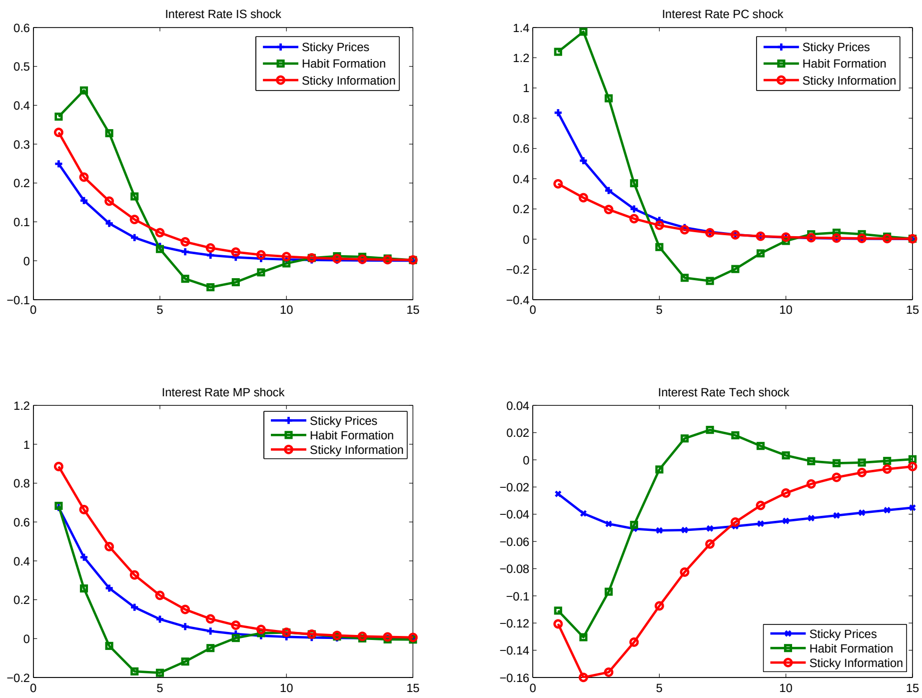


Figure 3: Impulse Responses of Interest Rate



4 Solution Methods

The three equation systems derived so far are not easy to solve because of their inclusion of dynamic and forward-looking expectational variables. In case of the sticky information model, one even has to deal with past expectations of current variables. Moreover, one faces the problem of the existence of multiple equilibria and thus has to evaluate the conditions which guarantee the existence of a single and stable steady state equilibrium. In order to solve the models and evaluate stability conditions, the IS curve, the Phillips curve, and the Taylor rule are rewritten in matrix form and rearranged into vectors of endogenous and predetermined variables. Various solution methods have been developed to solve such equation systems efficiently.

This overview gives a short review of solution methods for linear rational expectations models with forward-looking variables, both for single equations and systems of equations. Most methods can be implemented in Matlab.

4.1 Repeated Substitution

Consider a very simple expectational first-order difference equation, where an endogenous variable y_t depends on expectations of y_{t+1} held at time t and some exogenous variable x_t:[6]

[6]This section is based on Blanchard and Fischer (1989), section 5.1, pp. 214 onwards.

$$y_t = aE[y_{t+1}|t] + cx_t, \tag{134}$$

where $E[y_{t+1}|t]$ denotes the expectations of y_{t+1} held at time t.

Assumptions with regard to expectations:

1. Assumption of *rational expectations*: Expectations of y_{t+1} held at time t equal the *mathematical expectation* of y_{t+1} based on all relevant information available at t.

2. Individuals have full knowledge of the relevant economic model, i.e. the equation (134) and the parameters a and c.

3. All individuals have the same information set at time t, thus ruling out asymmetric information.

Hence, expectations are defined as follows:

$$\begin{aligned} E[y_{t+1}|t] &= E[y_{t+1}|I_t], \text{ where} \\ I_t &= [y_{t-i}, x_{t-i}, z_{t-i}, \ i = 0, ..., \infty] \end{aligned} \tag{135}$$

Note that this definition, in addition to the assumption of complete knowledge of the model, also implies no loss of memory and introduces an additional variable z_t, that does not feature in the model, but might help predict future values of y and x.

To solve (134) with repeated substitution, we make use of the *law of iterated expectations*: For any x with information set I_{t+1} and subset I_t we have:

$$E[E[x|I_{t+1}]|I_t] = E[x|I_t] \tag{136}$$

Note that the law of iterated expectations implies that there is *no systematic bias* in agents' expectations. Applying the law of iterated expectations, we forward equation (134) by one period and take expectations:

$$E[y_{t+1}|I_t] = aE[y_{t+2}|I_t] + cE[x_{t+1}|I_t] \tag{137}$$

Substituting the expression for $E[y_{t+1}|I_t]$ in (134) we get:

$$y_t = a^2E[y_{t+2}|I_t] + acE[x_{t+1}|I_t] + cx_t \tag{138}$$

Finally, substituting recursively up until period T, we get the recursive solution:

$$y_t = c\sum_{i=0}^{T} a^iE[x_{t+i}|I_t] + a^{T+1}E[y_{t+T+1}|I_t] \tag{139}$$

As T goes towards infinity, the existence and number of possible solutions depends on the behavior of the two summands in the recursive solution (139).

Possible Solutions to Equation (139):

1. $|a| < 1$:[7] The infinite sum $c\sum_{i=0}^{\infty} a^i E[x_{t+i}|I_t]$ converges and we will get a fundamental and a bubble solution to (134):

   (a) $|a| < 1$ **and** $\lim_{T\to\infty} a^{T+1} E[y_{t+T+1}|I_t] = 0$:

$$y_t^* = c\sum_{i=0}^{\infty} a^i E[x_{t+i}|I_t] \tag{140}$$

   is a solution to (134) and if we specify an expected path for x, we can solve for y explicitly.

   (b) $|a| < 1$ **without** $\lim_{T\to\infty} a^{T+1} E[y_{t+T+1}|I_t] = 0$: There can be additional bubble solutions:

$$y_t = y_t^* + b_t \tag{141}$$

2. $|a| > 1$: The infinite sum is unlikely to converge, and the solution will be an infinite set of stable bubbles:

$$\begin{gathered} y_t = (1-a)^{-1}c + b_t, \text{ where} \\ b_t = a^{-1}b_{t-1} + e_t, \;\; E[e_t|I_{t-1}] = 0 \end{gathered} \tag{142}$$

In the following, we will give a few examples for possible fundamental and bubble solutions, in order to further clarify the concepts:

4.1.1 The Fundamental Solution

Assuming the conditions in (1a) hold, by specifying a process for x, its fundamental solution determines the process for y, making it possible to solve for y explicitly: Assume, for instance, that x is announced at time t_0 to be increased from x_0 to x_T at time $T > t_0$. From (140), this will result in the following path for y:

$$\begin{aligned} y_t &= (1-a)^{-1}cx_0, && \text{for } t < t_0, \\ &= (1-a)^{-1}cx_0 + a^{T-t}(1-a)^{-1}c(x_T - x_0), && \text{for } t_0 \le t > T, \\ &= (1-a)^{-1}cx_T, && \text{for } t \ge T. \end{aligned} \tag{143}$$

By means of illustration, if for instance equation (134) is used to model the logarithmic price level as a function of the current nominal money stock and the expected price level, the path for the price level in response to an announced increase in the money stock in the future, given by equation (140), suggests that the price level increases already today. Intuitively, this stems from the fact that, due to the assumption of rational expectations, agents know that an increase in the money stock will induce a higher price level and therefore attempt to reduce their real money balances immediately. This causes the price level to go up asymptotically directly after the announcement before the increase in the money stock actually takes place.

[7] An additional condition for convergence of the infinite sum is that the expectation of x do not grow at a rate faster than $(1/a) - 1$. Any constant exponential growth of levels of x, implying constant linear growth of its logarithm, will satisfy this condition, so that we can assume it to be generally satisfied.

4.1.2 The Set of Bubble Solutions

If $\lim_{T\to\infty} a^{T+1}E[y_{t+T+1}|I_t] = 0$ does not hold, in addition to the fundamental solution there can exist a bubble solution, such that equation (141) will be a solution to the original equation (134). In order for this to be the case, the following conditions have to hold for the bubble solution b_t: Forwarding (141) one period, taking expectations and substituting for y_t and $E[yt+1|I_t]$ in (134) gives:

$$y_t^* + b_t = aE[y_{t+1}^*|I_t] + aE[b_{t+1}|I_t] + cx_t, \tag{144}$$

which, by definition of the fundamental solution in (140), reduces to

$$\begin{aligned} b_t &= aE[b_{t+1}|I_t] \\ E[b_{t+1}|I_t] &= a^{-1}b_t. \end{aligned} \tag{145}$$

For any bubble b_t that satisfies equation (145), the combined solution in (141) will be a solution to equation (134). Since we assume that the parameter a is less than 1, b_t explodes in expected value as expectations move towards infinity:

$$\lim_{i\to\infty} E[b_{t+i}|I_t] = a^{-i}b_t = \begin{cases} +\infty, & \text{if} b_t > 0, \\ -\infty, & \text{if} b_t < 0. \end{cases} \tag{146}$$

Hence, b_t embodies the notion of speculative bubbles, and can be modeled both as an ever expanding bubble with a constant time trend or as a bursting bubble, underlying a certain probability that it will burst each period, and that a new bubble gets started. Although in principle, with $|a| < 1$, a bubble solution can only be ruled out under the condition that expectations do not explode too fast in (1a), there are additional economic criteria that can rule out bubble solutions or make their appearance less likely. These criteria could, for instance, be the finiteness of the economy, a terminal condition for y_t or the availability of a close substitute with infinitely elastic supply.

4.2 The Method of Undetermined Coefficients

4.2.1 Solution of a Single Equation

Suppose, we want to solve a difference equation with both a lagged dependent variable and lagged as well as current expectations of that variable, say p_t:[8]

$$p_t = a_0E[p_{t+1}|I_t] + a_1p_{t-1} + a_2E[p_t|I_{t-1}] + x_t, \tag{147}$$

where x_t is some exogenous variable.

The *method of undetermined coefficients* consists in guessing a form of the solution for the original equation in (147) and solving for the coefficients. As a guess, it is assumed here that the solution for p_t contains a lagged dependent variable, as well as current and once-lagged expectations of once-lagged current and future values of the exogenous variable x:

$$p_t = \lambda p_{t-1} + \sum_{i=0}^{\infty} c_iE[x_{t+i}|t] + \sum_{i=0}^{\infty} d_iE[x_{t+i-1}|t-1]. \tag{148}$$

[8]This section is based on Blanchard and Fischer (1989), appendix of section 5, pp. 262 onwards.

We now have to find values for λ, c_i and d_i such that (148) is a solution to (147). As a first step, we derive expressions for $E[p_t|I_{t-1}]$ and $E[p_{t+1}|I_t]$ in (148) by taking expectations both at time t and $t-1$. Using the *law of iterated expectations*, we then find:

$$E[p_t|t-1] = \lambda p_{t-1} + \sum_{i=0}^{\infty} c_i E[x_{t+i}|t-1] + \sum_{i=0}^{\infty} d_i E[x_{t+i-1}|t-1] \tag{149}$$

$$E[p_{t+1}|t] = \lambda p_t + \sum_{i=0}^{\infty} c_i E[x_{t+i+1}|t] + \sum_{i=0}^{\infty} d_i E[x_{t+i}|t] \tag{150}$$

Substituting expressions in (149) and (150) into the original equation in (147) and simplifying we get:

$$\begin{aligned} p_t &= (1-a_0\lambda)^{-1}\{a_0\left(\sum_{i=0}^{\infty} c_i E[x_{t+i+1}|t] + \sum_{i=0}^{\infty} d_i E[x_{t+i}|t]\right) + (a_1 + a_2\lambda)p_{t-1} \\ &+ a_2\left(\sum_{i=0}^{\infty} c_i E[x_{t+i}]|t-1] + \sum_{i=0}^{\infty} d_i E[x_{t+i-1}|t-1]\right) + x_t\} \end{aligned} \tag{151}$$

Now, in order for our 'guess' in equation (148) to be a solution to (147), equations (148) and (151) must be identical. Therefore, we can equate the coefficients for each variable and then solve accordingly. Starting with p_{t-1}, we get from the coefficients:

$$\lambda = (1-a_0\lambda)^{-1}(a_1 + a_2\lambda), \tag{152}$$

which gives a quadratic function in λ:

$$a_0\lambda^2 + (a_2 - 1)\lambda + a_1 = 0 \tag{153}$$

The equation in (153) can be solved easily, usually resulting in two solutions for λ:

$$\lambda_{1,2} = \frac{-(a_2-1) \pm \sqrt{(a_2-1)^2 - 4a_0a_1}}{2a_0} \tag{154}$$

If the model in equation (147) satisfies the condition $|a| < 1$, then the solution to (154) will give us one root smaller than one in absolute value and one root greater than one. Thus, the model is saddle point (un)stable. By choosing the smaller root as the coefficient on p_{t-1}, we will automatically choose the stable solution. Let λ_1 be the root that is less in absolute value, and λ_2 be the root that is greater in absolute value. Note that from (154) we get that $\lambda_1\lambda_2 = a_1/a_0$ and $\lambda_1 + \lambda_2 = (1-a_2)/a_0$. Using this in euqation (151), we can now solve for c_i and d_i, assuming that $\lambda = \lambda_1$:

$$\begin{aligned} x_t: \quad & c_0 = (1-a_0\lambda_1)^{-1}[1 + a_0d_0], \\ E[x_{t+1}|t]: \quad & c_1 = (1-a_0\lambda_1)^{-1}[a_0(c_0 + d_1)], \\ E[x_{t+i}|t]: \quad & c_i = (1-a_0\lambda_1)^{-1}[a_0(c_{i-1} + d_i)], \\ x_{t-1}: \quad & d_0 = (1-a_0\lambda_1)^{-1}[a_2d_0], \\ E[x_t|t-1]: \quad & d_1 = (1-a_0\lambda_1)^{-1}[a_2(c_0 + d_1)], \\ E[x_{t+i}|t-1]: \quad & d_{i+1} = (1-a_0\lambda_1)^{-1}[a_2(c_i + d_{i+1})] \end{aligned}$$

Noting that $d_0 = 0$ and simplifying[9], we get for c_0, c_i and d_i:

$$\begin{aligned} c_0 &= (1 - a_0\lambda_1)^{-1}, && && (155) \\ c_i &= \left(\frac{\lambda_1 a_0}{a_1}\right) c_{i-1} = \lambda_2^{-1} c_{i-1} && \text{for } i = 1, \ldots, && (156) \\ d_i &= \left(\frac{a_2}{a_0}\right) c_i, && \text{for } i = 1, \ldots. && (157) \end{aligned}$$

Keeping in mind that λ_2 is greater than one in absolute value, both c_i and d_i converge towards zero as i goes towards infinity. Thus, we have identified p_t as a function of lagged p_t and current and once-lagged expectations of current and future values of x_t, with declining weights as expectations lie farther in the future. Supposing the process for x was known, we could identify the process for p explicitly.

Potential problems with the method of undetermined coefficients:

1. The solution depends on the initial guess, which may inadvertently exclude a possible solution or discard other solutions.
2. The method shows only indirectly whether the model possesses the desired saddle point property.
3. For large models, the method can become somewhat unwieldy.

4.2.2 Systems of Equations: Generalized Schur Decomposition

The method of undetermined coefficients can also be implemented for the solution of a system of expectational difference equations with rational expectations. Building on McCallum (1983) and an earlier version of Klein (2000), McCallum (1998) shows how the method of undetermined coefficients can be used to easily solve a system of linear difference equations with rational expectations, making use of the generalized Schur (QZ) decomposition.

Assume the model consists of a $M \times 1$ vector of non-predetermined endogenous variables y_t, a $K \times 1$ vector of predetermined variables k_t, and a $N \times 1$ vector of exogenous variables u_t. Furthermore, assume that the exogenous variables u_t follow a first-order autoregressive process with white-noise process ϵ_t (also $N \times 1$).[10] The model can then be written as follows:

$$A_{11} E_t y_{t+1} = B_{11} y_t + B_{12} k_t + C_1 u_t \tag{158}$$

$$u_t = R u_{t-1} + \epsilon_t \tag{159}$$

[9] To solve for d_i, backward the expression for d_{i+1} and solve for c_{i-1}. Then, using the result for c_{t-1} in the equation for c_i yields the solution for d_i. Next, to solve for c_i, use the result for d_i in the equation for c_i, plug in $a_2 = 1 - (\lambda_1 + \lambda_2)a_0$ and $\lambda_2 = a_1/a_0\lambda_1$, and rearrange to obtain the solution for c_i.

[10] A predetermined variable is a function only of variables included in the information set Ω_t at time t, so that k_{t+1} depends only on Ω_t, but not on Ω_{t+1}. By contrast, a non-predetermined variable can depend on any variable in the information set at time $t+1$, Ω_{t+1}. Thus, there is uncertainty surrounding the expectation of y_{t+1} held at time t, captured in the expression $y_{t+1} = E(y_{t+1}|\Omega_t)$.

To complete the model, we also assume a path for k_t:

$$k_{t+1} = B_{21}y_t + B_{22}k_t + C_2u_t \tag{160}$$

Note that the matrices A_{11}, B_{21} and B_{22} may be singular. A solution to the model with the method of undetermined coefficients will be of the general form defining the variables' *law of motion*

$$y_t = \Omega k_t + \Gamma u_t, \tag{161}$$

$$k_{t+1} = \Pi_1 k_t + \Pi_2 u_t, \tag{162}$$

where the matrices Ω, Γ, Π_1 and Π_2 are real.[11] Therefore, $E_t y_{t+1} = \Omega E_t k_{t+1} + \Gamma E_t u_{t+1} = \Omega(\Pi_1 k_t + \Pi_2 u_t) + \Gamma R u_t$. Substituting this result in (158) and (160), we then get:

$$A_{11}[\Omega(\Pi_1 k_t + \Pi_2 u_t) + \Gamma R u_t] = B_{11}[\Omega k_t + \Gamma u_t] + B_{12}k_t + C_1 u_t \text{[12]} \tag{163}$$

$$(\Pi_1 k_t + \Pi_2 u_t) = B_{21}(\Omega k_t + \Gamma u_t) + B_{22}k_t + C_2 u_t \text{[13]} \tag{164}$$

In order for (163) and (164) to be a solution to the model in (158) and (160), coefficients on k_t and u_t in (163) and (164) must be equal. We therefore get, by collecting terms for k_t:

$$\begin{bmatrix} A_{11} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \Omega\Pi_1 \\ \Pi_1 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \Omega \\ I \end{bmatrix}, \tag{165}$$

whereas the terms in u_t imply[14]

$$A_{11}\Omega\Pi_2 + A_{11}\Gamma R = B_{11}\Gamma + C_1 \tag{166}$$

$$\Pi_2 = B_{21}\Gamma + C_2. \tag{167}$$

Denoting the two square matrices in (165) with A and B, we assume that $|B - \lambda A|$ is nonzero for some complex number λ. For this condition to hold, the model must be well formulated with restrictions on some endogenous variables, however, it will hold even if the matrices A_{11} and B_{21} and B_{22} are singular.

Theorem 1 *Then, the Generalized Schur Decomposition Theorem guarantees the existence of unitary, invertible matrices Q and Z such that $QAZ = S$ and $QBZ = T$, where S and T are triangular.*

[11] Similar to equations (184) and (185) in Uhlig's toolkit that we will present in the next section.
[12] Similar to equation (192) in Uhlig's toolkit.
[13] Similar to equation (191) in Uhlig's toolkit.
[14] In Uhlig's toolkit: $\Omega = P$, $\Pi_1 = R$, $\Gamma = Q$ and $\Pi_2 = S$.

The ratios t_{ii}/s_{ii} are then generalized eigenvalues of the matrix expression $|B-\lambda A|$ and can be rearranged without violating the Generalized Schur Decomposition Theorem. A rearrangement of generalized eigenvalues (or columns of Q and Z) corresponds to selecting different solutions of the method of undetermined coefficients (see discussion above and below). Here, we assume that the generalized eigenvalues are ordered according to the moduli, with the largest values first.
Premultiplying (165) by Q and noting that $QA = SH$ and $QB = TH$, where $H \equiv Z^{-1}$, we get:

$$\begin{bmatrix} S_{11} & 0 \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \Omega\Pi_1 \\ \Pi_1 \end{bmatrix} = \begin{bmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \Omega \\ I \end{bmatrix} \tag{168}$$

The first row of (168), accordingly, can be written as

$$S_{11}(H_{11}\Omega + H_{12})\Pi_1 = T_{11}(H_{11}\Omega + H_{12}). \tag{169}$$

This will be satisfied for Ω such that

$$\Omega = -H_{11}^{-1}H_{12}. \tag{170}$$

In order to express the solution for Ω in (170) in terms of the matrix Z, recall that we have $H \equiv Z^{-1}$. This gives us:

$$\begin{aligned} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} &= \begin{bmatrix} H_{11}Z_{11} + H_{12}Z_{21} & H_{11}Z_{12} + H_{12}Z_{22} \\ H_{21}Z_{11} + H_{22}Z_{21} & H_{21}Z_{12} + H_{22}Z_{22} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \end{aligned} \tag{171}$$

Position (2,2) of the matrix HZ in (171) results in the following relation:[15]

$$\begin{aligned} H_{21}Z_{12} + H_{22}Z_{22} = I \quad \Leftrightarrow \quad H_{21}Z_{12} &= I - H_{22}Z_{22} \\ H_{21}Z_{12}Z_{22}^{-1} &= Z_{22}^{-1} - H_{22} \\ H_{21}Z_{12}Z_{22}^{-1} + H_{22} &= Z_{22}^{-1}. \end{aligned} \tag{172}$$

Similarly, position (2,1) of the matrix HZ gives us:[16]

$$\begin{aligned} H_{11}Z_{12} = -H_{12}Z_{22} \quad \Leftrightarrow \quad Z_{12} &= -H_{11}^{-1}H_{12}Z_{22} \\ Z_{12}Z_{22}^{-1} &= -H_{11}^{-1}H_{12} \\ H_{11}Z_{12}Z_{22}^{-1} + H_{12} &= 0. \end{aligned} \tag{173}$$

Now, using the second equality of equation (173), we can express the solution for Ω as follows:

$$\Omega = -H_{11}^{-1}H_{12} = Z_{12}Z_{22}^{-1}. \tag{174}$$

[15] Provided that Z_{22}^{-1} exists.
[16] Provided that H_{11}^{-1} and Z_{22}^{-1} exist.

We thus find a solution for Ω in (170), provided that Z_{22}^{-1} exits.

Similarly, for the second row of (168) we get:

$$S_{21}(H_{11}\Omega+H_{12})\Pi_1+S_{22}(H_{21}\Omega+H_{22})\Pi_1 = T_{21}(H_{11}\Omega+H_{12})+T_{22}(H_{21}\Omega+H_{22}). \quad (175)$$

From equation (174) and the last equality in equation (173) we have that the expression in brackets of the first summands on either side of (175) must be equal to zero. Equally, from the last equality in equation (172) we get that the expression in brackets of the second summands equals Z_{22}^{-1}. (175) can thus be simplified to:

$$S_{22}Z_{22}^{-1}\Pi_1 = T_{22}Z_{22}^{-1}. \quad (176)$$

Since by arrangement of the generalized eigenvalues, S_{22} has no zero elements on the diagonal and is triangular, we know that S_{22}^{-1} exists by construction and can thus be used to reformulate (176), resulting in a solution for Π_1:

$$\Pi_1 = Z_{22}S_{22}^{-1}T_{22}Z_{22}^{-1}. \quad (177)$$

Finally, we need to find solutions for Γ and Π_2 to be able to characterize our solution for y_t and k_{t+1}. Combining (166) and (167), we have

$$G\Gamma + A_{11}\Gamma R = F, \quad (178)$$

where $G \equiv A_{11}\Omega B_{21} - B_{11}$ and $F \equiv C_1 - A_{11}\Omega C_2$. If G^{-1} exists, which it typically will with nonsingular B_{11}, (178) becomes

$$\Gamma + G^{-1}A_{11}\Gamma R = G^{-1}F. \quad (179)$$

The latter can be solved for Γ with the following formula:

$$vec(\Gamma) = [I + R' \otimes G^{-1}]^{-1}vec(G^{-1}F).\text{[17]} \quad (180)$$

Finally, the solution for Π_2 can be obtained from equation (167).
To sum up, the solutions of the method of undetermined coefficients for the model in equations (158) - (160) for a given ordering of the eigenvalues is obtained sequentially from equations (170), (177), (180) and (167.).

4.2.3 Uhlig's Toolkit

Uhlig (1995) shows how to find a solution to a system of linear rational expectations equations. Two approaches can be distinguished. In the *brute force* method, the vector of lagged endogenous variables x_{t-1} is considered as being predetermined and thus exogenous. However, he argues in favor of using a more *sensitive* approach, since it keeps the original structure of the system without having to apply several transformations before starting the solution process. Another advantage is that one has one equation without expectations. It is the second approach that will be presented in this subsection.

[17] Equation (180) makes use of the identity that if A, B and C are real conformable matrices, $vec(ABC) = (C'\otimes A)vec(B)$. Applying this to (179) gives $vec(\Gamma)-(R'\otimes G^{-1})vec(\Gamma) = vec(G^{-1}F)$, which is reformulated to give (180)

To start, define

x_t, size $m \times 1$: vector of endogenous state variables
y_t, size $n \times 1$: vector of other endogenous variables (*jump variables*)
z_t, size $k \times 1$: vector of exogenous stochastic processes[18]

The equilibrium relationships between these variables are:

$$0 = Ax_t + Bx_{t-1} + Cy_t + Dz_t \tag{181}$$

$$0 = E_t \left[Fx_{t+1} + Gx_t + Hx_{t-1} + Jy_{t+1} + Ky_t + Lz_{t+1} + Mz_t \right] \tag{182}$$

$$z_{t+1} = Nz_t + \epsilon_{t+1}; E_t[\epsilon_{t+1}] = 0 \tag{183}$$

The following assumptions are used:

1. C is of size $l \times n$, $l \geq n$ and $rank(C) = n$
2. F is of size $(m + n - l) \times n$
3. N has only stable eigenvalues

Note that in the *brute force* method, one would have written (181) in form of (182).
Then, one writes the *recursive equilibrium law of motion* as

$$x_t = Px_{t-1} + Qz_t \tag{184}$$

$$y_t = Rx_{t-1} + Sz_t \tag{185}$$

The solution is characterized in the following theorem:

Theorem 2 *If there is a recursive equilibrium law of motion solving equations (181), (182), and (183), then the coefficient matrices can be found as follows.*

To solve for the coefficient matrices, define two matrices as follows:

Since C is not quadratic, C^+ is the *pseudo-inverse* of C, such that $C^+ = C^+CC^+$ and $CC^+C = C$, with C^+ $n \times l$. Since by assumption $rank(C) \geq n$, one gets $C^+ = (C'C)^{-1}C'$.[19]

and

C^0 is a matrix with $(l - n) \times l$, whose rows form a basis of the null space of C'. C^0

[18] Here, the vector of endogenous state variables x_t is defined similar to the non-predetermined variables in y_t in the Schur decomposition and the jump variables in y_t are defined similar to the vector of predetermined variables k_t in the Schur decomposition. As Uhlig notes: "Fundamentally, there is no difference" Uhlig (1995), p.40.

[19] The *pseudo-inverse* can be computed in MATLAB with $pinv(C)$.

can be found via the *singular value decomposition*[20] of C'.

Then, the solution can be written as:

1. P satisfies the matrix quadratic equations:

$$0 = C^0 AP + C^0 B \tag{186}$$
$$0 = (F - JC^+ A)P^2 - (JC^+ B - G + KC^+ A)P - KC^+ B + H \tag{187}$$

   This solution is stable iff all eigenvalues of P are smaller than unity in absolute value.

2. R is given by

$$R = -C^+(AP + B) \tag{188}$$

3. Given P and R, let V be the matrix

$$V = \left[\begin{array}{cc} I_k \otimes A, & I_k \otimes C \\ N' \otimes F + I_k \otimes (FP + JR + G), & N' \otimes J + I_k \otimes K \end{array} \right] \tag{189}$$

   where I_k is the identity matrix of size $k \times k$. Then

$$V \left[\begin{array}{c} vec(Q) \\ vec(S) \end{array} \right] = - \left[\begin{array}{c} vec(D) \\ vec(LN + M) \end{array} \right] \tag{190}$$

   where $vec(\cdot)$ denotes vectorization.[21]

Proof

Plug the law of motion (184), (185) into (181).

$$A(Px_{t-1} + Qz_t) + Bx_{t-1} + C(Rx_{t-1} + Sz_t) + Dz_t = 0$$

[20]Any $n \times m$ matrix A can be written as: $A = UDV'$, where
$U =$ eigenvectors of AA'
$D = \sqrt{diag(eig(AA'))}$
$V =$ eigenvectors of $A'A$

[21]By making use of the vec-operator, a matrix A is transformed into a vector, by arranging the column vectors of A to *one* column vector $vec(A)$. Be a_i the i-the column of A:
$$A = (a_1, a_2, ..., a_i) \Longrightarrow vec(A) := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \end{pmatrix}$$
The Kronecker-product is defined as follows:
$$A \otimes B := \begin{pmatrix} a_{11}B, & a_{12}B & \ldots, & a_{1n}B \\ \vdots, & \vdots, & & \vdots \\ a_{m1}B, & a_{m2}B, & \ldots, & a_{mn}B \end{pmatrix} = (a_{ij}B)$$
Note that $A \otimes B \neq B \otimes A$. See for this Rinne (2004).

$$(AP + CR + B)x_{t-1} + (AQ + CS + D)z_t = 0 \tag{191}$$

This has to hold for arbitrary x_{t-1} and z_t, hence the coefficient matrices for x_{t-1} and z_t in (191) are zero. Next, plugging in the law of motion (184), (185) into equation (182) twice

$$\begin{aligned} 0 &= E_t[F(Px_t + Qz_{t+1}) + G(Px_{t-1} + Qz_t) + Hx_{t-1} + J(Rx_t + Sz_{t+1}) \\ &+ K(Rx_{t-1} + Sz_t) + Lz_{t+1} + Mz_t] \end{aligned}$$

$$\begin{aligned} 0 &= E_t[F(P(x_{t-1} + Qz_t) + Qz_{t+1}) + G(Px_{t-1} + Qz_t) + Hx_{t-1} \\ &+ J(R(Px_{t-1} + Sz_t) + Sz_{t+1}) + K(Rx_{t-1} + Sz_t) + Lz_{t+1} + Mz_t] \end{aligned}$$

and using (183) and taking expectations yields

$$\begin{aligned} 0 &= F(P(x_{t-1} + Qz_t) + QNz_t) + G(Px_{t-1} + Qz_t) + Hx_{t-1} \\ &+ J(R(Px_{t-1} + Qz_t) + SNz_t) + K(Rx_{t-1} + Sz_t) + LNz_t + Mz_t \end{aligned}$$

which can be rearranged to give:

$$\begin{aligned} 0 &= ((FP + JR + G)P + KR + H)x_{t-1} + ((RQ + JS + L)N \\ &+ (FP + JR + G)Q + KS + M)z_t \end{aligned} \tag{192}$$

Again, the coefficient matrices on x_{t-1} and z_t have to be zero. One takes the column by column vectorization of the coefficient matrices of z_t in (191) and (192), collects terms and gets the formula for Q and S in (190), by using the matrix V. Next, to find P and thus R, one rewrites the coefficient matrix on x_{t-1} in (191) as:

$$R = -C^+(AP + B) \tag{193}$$

using the *pseudo-inverse* C^+ for rectangular matrices. Then, by using the matrix C^0 and noting that $C^0C = 0$, one gets:

$$0 = C^0AP + C^0B \tag{194}$$

Then, using (193) to replace R in the coefficient matrix on x_{t-1} in (192) gives:

$$0 = (FP + J(-C^+(AP + B)) + G)P + K(-C^+(AP + B)) + H, \tag{195}$$

hence the solution P in (187). Thus, one has a formula for all the the four matrices P, Q, R, S in the equilibrium law of motion given by (184) and (185). With regard to the question of stability, note that this is determined by the stability of P, since N, the matrix of the stochastic process z_{t+i} has stable roots by assumption.

Excursion: Solving matrix quadratic equations

Solving the matrix quadratic equations in (186) and (187) can be done in the following way:
First, note that (187) can be written generally as

$$\Psi P^2 - \Gamma P - \Theta = 0 \tag{1}$$

For equations (186) and (187), define

$$\Psi = \begin{bmatrix} 0_{l-n,m} \\ F - JC^{+}A \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} C^{0}A \\ JC^{+}B - G + KC^{+}A \end{bmatrix}$$

$$\Theta = \begin{bmatrix} C^{0}B \\ KC^{+}B - H \end{bmatrix}$$

where $0_{l-n,m}$ is a matrix with only zero entries. Equation (1) can be solved by turning it into a generalized eigenvalue and eigenvector problem.[2] A generalized eigenvalue λ and eigenvector s of a matrix Ξ with respect to a matrix Δ are defined to be a vector and a value satisfying

$$\lambda \Delta s = \Xi s \tag{2}$$

The solution for (1) is then characterized by the following theorem:

Theorem 3 *Solution of quadratic matrix equations*

To solve (1), define the $2m \times 2m$ matrices Ξ and Δ as, given the $m \times m$ matrices P, Γ, Θ:

$$\Xi = \begin{bmatrix} \Gamma & \Theta \\ I_m & 0_{m,m} \end{bmatrix}$$

$$\Delta = \begin{bmatrix} \Psi & 0_{m,m} \\ 0_{m,m} & I_m \end{bmatrix}$$

where I_m is the identity matrix of size m, and $0_{m,m}$ is the $m \times m$ matrix with only zero entries.
Then, for m generalized eigenvalues $\lambda_1, ..., \lambda_m$ and m eigenvectors $s_1, ..., s_m$, written as $s_i' 0[\lambda_i x_i', x_i']$ for some $x_i \in \Re^m$ and if $(x_1, ..., x_m)$ is linearly independent, then a solution to (1) is given by

$$P = \Omega \Delta \Omega^{-1} \tag{3}$$

where $\Omega = (x_1, ..., x_m)$ and $\Delta = diag(\lambda_1, ..., \lambda_m)$. The solution P is stable if $|\lambda_i| < 1 \forall i = 1, ..., m$

[2] In MATLAB: $eig(\Xi, \Delta)$.

4.2.4 Stability Analysis and Solution Selection Criteria

As already shown in section 4.1, a system of linear rational expectations equations, in addition to the stable and unique fundamental solution, may yield an infinity of stable bubble solutions or exploding, thus non-stable, bubble solutions.[22] Blanchard and Kahn (1980) show how to check the conditions for a unique stable solution for a linear rational expectations model. The method described below can be directly applied to the generalized Schur decomposition or Uhlig's toolkit.

Suppose we have a general model of the following form, where X is an $(n \times 1)$ vector of **predetermined** variables, P is an $(m \times 1)$ vector of **non-predetermined** variables and Z is an $(k \times 1)$ vector of exogenous variables. The model consists of the following three equations:

$$\begin{bmatrix} X_{t+1} \\ E_t P_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ P_t \end{bmatrix} + \gamma Z_t, \quad X_{t=0} = X_0, \tag{196}$$

$$E_t P_{t+1} = E(P_{t+1} | \Omega_t), \tag{197}$$

$$\begin{gathered} \forall t \; \exists \bar{Z}_t \in \Re^k, \quad \theta_t \in \Re \qquad \text{such that} \\ -(1+i)^{\theta_t} \bar{Z}_t \leq E(Z_{t+i} | \Omega_t) \leq (1+i)^{\theta_t} \bar{Z}_t \quad \forall i \geq 0. \end{gathered} \tag{198}$$

Equation (196) describes the structural model, equation (197) defines rational expectations and equation (198) requires that the exogenous variables in Z do not grow too fast, by essentially ruling out exponential growth of the expectations of Z_{t+i} held at time t. Furthermore, the authors also assume that the expectations of X_t and P_t do not explode, thereby ruling out bubble solutions.
Similar to the solution methods described above, the model is solved by transforming the coefficient matrix A and deriving solutions in the transformed matrix. Blanchard and Kahn (1980) make use of the *Jordan canonical form*:

$$A = C^{-1} J C \tag{199}$$

The elements on the diagonal of the matrix J are the eigenvalues of A, which are ordered by *increasing absolute value*.[23] J is then further decomposed as

$$J = \begin{bmatrix} \underset{(\bar{n} \times \bar{n})}{J_1} & 0 \\ 0 & \underset{(\bar{m} \times \bar{m})}{J_2} \end{bmatrix}, \tag{200}$$

such that all eigenvalues of J_1 are on or inside the unit circle and all eigenvalues of J_2 are outside the unit circle. Then the stability conditions of the model can be stated as follows:

[22] This section is based on Walsh (2003), Section 5.4.3, pp. 245 - 247, Blanchard and Kahn (1980) and McCallum (1998).

[23] Note that a different ordering will yield different solutions. The MSV criterion described below is an example of a different ordering of the eigenvalues of A.

1. If $\bar{m} = m$, i.e. if the number of eigenvalues of A **outside** the unit circle is **equal** to the number of **non-predetermined** variables, then there exists a **unique solution** which is forward-looking in the sense that the non-predetermined variables P_t depend on the past only through their effect on current predetermined variables X_t.

2. If $\bar{m} > m$, i.e. if the number of eigenvalues outside the unit circle **exceeds** the number of non-predetermined variables, there will be **no solution** satisfying both (196) and (198).

3. If $\bar{m} < m$, i.e. if the number of eigenvalues outside the unit circle **is less** than the number of non-predetermined variables, there is an **infinity of solutions**.

We thus have explicit stability conditions for the eigenvalues of the matrix A in order for there to be a unique stable solution. Note that the condition $\bar{m} = m$ is equivalent to the condition that A must have the strict saddle point property, as the resulting solution will be a saddle point.

McCallum (1983) develops another stability criterion, the so-called minimal-state-variable (MSV) procedure. McCallum (1998) notes that this procedure will generally choose the same solution as the Blanchard-Kahn criterion, if there is exactly one stable solution, thus $\bar{m} = m$ holds. However, in the case of $\bar{m} > m$, the MSV criterion will choose a single explosive solution, whereas in the case of $\bar{m} < m$ the MSV procedure will yield the single stable solution that is bubble-free. Thus, if $\bar{m} = m$ is violated, the MSV procedure may give alternative solutions that may be of specific scientific interest.

4.3 Solving Expectations Models with Lagged Expectations

The solution methods presented above can solve systems of equations with rational expectations easily, making use of the method of undetermined coefficients. However, these solution algorithms do not account for lagged expectations of variables, such as can be found in models with sticky information, e.g. Mankiw and Reis (2006b, 2007). Because these models entail both a sum to $+\infty$ of variables due to rational expectations, and a sum to $-\infty$ of lagged expectations due to sticky information, a new solution algorithm is needed.

Meyer-Gohde (2007, 2009) presents a solution method to linear rational expectations models with a (potentially infinite) sum of lagged expectations, that encompasses linear rational expectations models without lagged expectations as a special case. The solution method thus builds on those by McCallum (1983), McCallum (1998), Uhlig (1995) and Klein (2000), but extends the analysis to account for the additional backward-looking dimension.

To begin with, the model (consisting of a system of expectational difference equations that are linear in the percentage deviations of variables from their respective steady states) is characterized by the following equations:

$$
\begin{aligned}
0 &= \sum_{i=0}^{I} A_i E_{t-i}[Y_{t+1}] + \sum_{i=0}^{I} B_i E_{t-i}[Y_t] + \sum_{i=0}^{I} C_i E_{t-i}[Y_{t-1}] \\
&+ \sum_{i=0}^{I} F_i E_{t-i}[W_{t+1}] + \sum_{i=0}^{I} G_i E_{t-i}[W_t] \qquad (201)
\end{aligned}
$$

$$
W_t = \sum_{j=0}^{\infty} N_j \varepsilon_{t-j}, \quad \varepsilon_t \sim i.i.d.N(0,\Omega) \qquad (202)
$$

$$
\lim_{j\to\infty} \xi^{-j} E_t[Y_{t+j}] = 0, \quad \forall\, \xi \in \Re \; s.t. \; \xi > g^u, \quad \text{where } g^u \geq 1, \qquad (203)
$$

where Y_t is a $k \times 1$ vector of endogenous variables, W_t denotes an $n \times 1$ vector of exogenous variables following an $MA(\infty)$ process with coefficients $\{N_j\}_{j=0}^{\infty}$ and where $I \in \aleph_0$. The system is specified such that there are as many equations in the matrix system in (201) as there are endogenous variables, namely k. The first equation thus gives the log-linearized equilibrium equations of the model in question, the second equation specifies the exogenous variables as an infinite moving-average process of stochastic shocks ε and the third equation gives a transversality condition stating that the endogenous variables in Y may not grow faster than their maximum growth rate g^u.

Using the method of undetermined coefficients, we guess that the solution of the system with respect to the endogenous variables in Y_t will take the following form:

$$
Y_t = \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j} \qquad (204)
$$

Inserting the solution in (204) into (201) gives:

$$
\begin{aligned}
0 &= \sum_{j=0}^{\infty} \left(\sum_{i=0}^{min(I,j)} A_i \right) \Theta_{j+1} \varepsilon_{t-j} + \sum_{j=0}^{\infty} \left(\sum_{i=0}^{min(I,j)} B_i \right) \Theta_j \varepsilon_{t-j} \\
&+ \sum_{j=0}^{\infty} \left(\sum_{i=0}^{min(I,j+1)} C_i \right) \Theta_j \varepsilon_{t-j-1} + \sum_{j=0}^{\infty} \left(\sum_{i=0}^{min(I,j)} F_i \right) N_{j+1} \varepsilon_{t-j} \\
&+ \sum_{j=0}^{\infty} \left(\sum_{i=0}^{min(I,j)} G_i \right) N_j \varepsilon_{t-j} \qquad (205)
\end{aligned}
$$

Defining

$$
\tilde{M}_j = \sum_{i=0}^{min(I,j)} M_i, \quad \text{for } M = A, B, C, F, G, \qquad (206)
$$

we can simplifiy (205) to

$$
\begin{aligned}
0 \;=\; & \sum_{j=0}^{\infty} \tilde{A}_j \Theta_{j+1} \varepsilon_{t-j} + \sum_{j=0}^{\infty} \tilde{B}_j \Theta_j \varepsilon_{t-j} + \sum_{j=0}^{\infty} \tilde{C}_{j+1} \Theta_j \varepsilon_{t-j-1} \\
+ & \sum_{j=0}^{\infty} \tilde{F}_j N_{j+1} \varepsilon_{t-j} + \sum_{j=0}^{\infty} \tilde{G}_j N_j \varepsilon_{t-j}.
\end{aligned} \tag{207}
$$

Comparing coefficients for ε_{t-j} in (207) gives the non-stochastic linear recursion

$$
0 = \tilde{A}_j \Theta_{j+1} + \tilde{B}_j \Theta_j + \tilde{C}_j \Theta_{j-1} + \tilde{F}_j N_{j+1} + \tilde{G}_j N_j \tag{208}
$$

with initial conditions

$$
\Theta_{-1} = 0 \tag{209}
$$

and terminal conditions from the transversality condition in (203):

$$
\lim_{j \to \infty} \xi^{-j} \Theta_j = 0 \tag{210}
$$

The recursion in (208) that characterizes our solution can be solved with the Generalized Schur decomposition, whereby Meyer-Gohde (2007) distinguishes between three cases:

1. $I = 0$
2. $0 < I < \infty$
3. $I \to \infty$

In the first case, the model contains no lagged expectations and thus collapses to the standard case, in the second case there is a finite number of lagged expectations in the model and in the third case the number of lagged expectations converges towards infinity. The first case is covered in the previous section and for ease of exposition we will here only describe the solution to the third case, as it encompasses the second case and is also the most relevant, since sticky information models typically include an infinite sum of lagged expectations.

Solution for $I \to \infty$:
First, define the following matrices of limiting coefficients, as $j \to \infty$:

$$
\lim_{j \to \infty} \left(\tilde{M}_j \right)_{l,m} = \left(\tilde{M}_\infty \right)_{l,m}, \quad \text{for } M = A, B, C, F, G, \tag{211}
$$

where l denotes row and m denotes column. Assuming that the limit in (211) exists and is finite, by definition there exists some $I(\delta)_{M,l,m}$ for each M, l and m such that

$$
\forall\, \delta > 0, \quad \exists I(\delta)_{M,l,m} \text{ s.t. } n > I(\delta)_{M,l,m} \Rightarrow | \left(\tilde{M}_n \right)_{l,m} - \left(\tilde{M}_\infty \right)_{l,m} | < \delta \tag{212}
$$

By the same argument, there exists some upper bound $I(\delta)_{max} = max\{I(\delta)_{M,l,m}\}$:

$$\forall\delta > 0, \quad \exists I(\delta)_{max} \text{ s.t. } n > I(\delta)_{max} \Rightarrow |\left(\tilde{M}_n\right)_{l,m} - \left(\tilde{M}_\infty\right)_{l,m}| < \delta; \quad \forall M, l, m \tag{213}$$

We can use this result to approximate the non-stochastic recursion with time-varying coefficients in (208) for $I \rightarrow \infty$ as

$$0 = \tilde{A}_j\Theta_{j+1} + \tilde{B}_j\Theta_j + \tilde{C}_j\Theta_{j-1} + \tilde{F}_jN_{j+1} + \tilde{G}_jN_j, \quad 0 \leq j \leq I(\delta)_{max} \tag{214}$$

and

$$0 = \tilde{A}_\infty\Theta_{j+1} + \tilde{B}_\infty\Theta_j + \tilde{C}_\infty\Theta_{j-1} + \tilde{F}_\infty N_{j+1} + \tilde{G}_\infty N_j, \quad j \geq I(\delta)_{max}. \tag{215}$$

For $j \geq I(\delta)_{max}$, the system thus reduces to a recursive equation with *constant* coefficients $\tilde{M}_\infty$ as in (215). The system of equations is thus divided into a non-autonomous part with time-varying coefficients $\tilde{M}_j$ for $0 \leq j \leq I(\delta)_{max}$ and an autonomous part with constant coefficients $\tilde{M}_\infty$ for $j \geq I(\delta)_{max}$. This autonomous system can be re-written in first-order form:

$$\begin{pmatrix} 0 & -\tilde{A}_\infty \\ I & 0 \end{pmatrix}\begin{pmatrix} \Theta_j \\ \Theta_{j+1} \end{pmatrix} = \begin{pmatrix} \tilde{C}_\infty & \tilde{B}_\infty \\ 0 & I \end{pmatrix}\begin{pmatrix} \Theta_{j-1} \\ \Theta_j \end{pmatrix} + \begin{pmatrix} \tilde{F}_\infty N_{j+1} + \tilde{G}_\infty N_j \\ 0 \end{pmatrix} \tag{216}$$

Applying the QZ method to look for a Generalized Schur Decomposition, the coefficient matrices are decomposed such that

$$Q\begin{pmatrix} 0 & -\tilde{A}_\infty \\ I & 0 \end{pmatrix}Z = S$$

and

$$Q\begin{pmatrix} \tilde{C}_\infty & \tilde{B}_\infty \\ 0 & I \end{pmatrix}Z = T,$$

where Q and Z are unitary matrices and S and T are upper-triangular.[24] The solution of the method of undetermined coefficients for the coefficients in $\tilde{M}_\infty$ is then given by the generalized eigenvalues λ_i:

$$\lambda_i = \begin{cases} \frac{T_{i,i}}{S_{i,i}} & \text{if } S_{i,i} \neq 0 \\ \infty & \text{otherwise} \end{cases} \tag{217}$$

As suggested by Blanchard and Kahn (1980), the $2k$ generalized eigenvalues are ordered such that the first s eigenvalues are those less than or equal to the maximal growth rate g^u (thus satisfying the transversality condition in (203) with the remaining $2k-s$ greater than g^u). If $s = k$ (k = number of endogenous variables), the solution to the system will be unique, if $s > k$ it will be indeterminate (no solution) and if $s < k$, the solution will be explosive (an infinity of solutions).

[24] Note that the matrix Z is usually solved numerically, even if it could be found analytically as well.

In the following, we assume that $s = k$, thus, there exists a unique solution to the system of equations. Furthermore, we assume that the upper-left $k \times k$ block of Z is invertible. Then, defining $\begin{pmatrix} \Xi_j^s \\ \Xi_j^u \end{pmatrix} = Z^+ \begin{pmatrix} \Theta_{j-1} \\ \Theta_j \end{pmatrix}$ [25], we can rewrite (216) as

$$\begin{pmatrix} 0 & -\tilde{A}_\infty \\ I & 0 \end{pmatrix} Z \begin{pmatrix} \Xi_{j+1}^s \\ \Xi_{j+1}^u \end{pmatrix} = \begin{pmatrix} \tilde{C}_\infty & \tilde{B}_\infty \\ 0 & 0 \end{pmatrix} Z \begin{pmatrix} \Xi_j^s \\ \Xi_j^u \end{pmatrix} + \begin{pmatrix} \tilde{F}_\infty N_{j+1} + \tilde{G}_\infty N_j \\ 0 \end{pmatrix} \tag{218}$$

Multiplying through by Q and using the definitions given above then yields:

$$\begin{pmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{pmatrix} \begin{pmatrix} \Xi_{j+1}^s \\ \Xi_{j+1}^u \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix} \begin{pmatrix} \Xi_j^s \\ \Xi_j^u \end{pmatrix} + \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \begin{pmatrix} \tilde{F}_\infty N_{j+1} + \tilde{G}_\infty N_j \\ 0 \end{pmatrix} \tag{219}$$

The second row of (219) can be written in terms of Ξ_j^u as follows:

$$\Xi_j^u = T_{22}^{-1} S_{22} \Xi_{j+1}^u - T_{22}^{-1} Q_2 \begin{bmatrix} \tilde{F}_\infty N_{j+1} + \tilde{G}_\infty N_j \\ 0 \end{bmatrix} \tag{220}$$

We can solve the first order difference equation in (220) 'forward' with the simple formula derived for the fundamental solution of recursive substitution in Section 4.1, equation (140). This gives the following solution for Ξ_j^u:

$$\Xi_j^u = -T_{22}^{-1} \sum_{k=0}^{\infty} [T_{22}^{-1} S_{22}]^k Q_2 \begin{bmatrix} \tilde{F}_\infty N_{j+1+k} + \tilde{G}_\infty N_{j+k} \\ 0 \end{bmatrix} \tag{221}$$

as long as it holds that

$$\lim_{k \to \infty} [T_{22}^{-1} S_{22}]^k Q_2 \begin{bmatrix} \tilde{F}_\infty N_{j+1+k} + \tilde{G}_\infty N_{j+k} \\ 0 \end{bmatrix} = 0 \tag{222}$$

Defining

$$M_j^{I_{\max}} = -T_{22}^{-1} \sum_{k=0}^{\infty} [T_{22}^{-1} S_{22}]^k Q_2 \begin{bmatrix} \tilde{F}_\infty N_{j+1+k} + \tilde{G}_\infty N_{j+k} \\ 0 \end{bmatrix} = \Xi_j^u, \tag{223}$$

the recursive solution for Θ_j is then given by:[26]

$$\Theta_j = (Z_{21} Z_{11}^{-1}) \Theta_{j-1} + (Z_{22} - Z_{21} Z_{11}^{-1} Z_{12}) M_j^{I_{max}}, \quad \forall j \geq I(\delta)_{max} \tag{224}$$

The proof for this result goes as follows:

Recall the definitions $\begin{pmatrix} \Xi_j^s \\ \Xi_j^u \end{pmatrix} = Z^+ \begin{pmatrix} \Theta_{j-1} \\ \Theta_j \end{pmatrix}$ and $Z^+ Z = I$, which can be partitioned into

$$\begin{bmatrix} Z_{11}^+ & Z_{12}^+ \\ Z_{21}^+ & Z_{22}^+ \end{bmatrix} \begin{bmatrix} \Theta_{j-1} \\ \Theta_j \end{bmatrix} = \begin{bmatrix} \Xi_j^s \\ \Xi_j^u \end{bmatrix} \tag{225}$$

and

[25] Z^+ denotes the Hermitian transpose of Z.
[26] See Klein (2000), theorem 5.1, pp. 1417-1418.

$$\begin{bmatrix} Z^+_{11} & Z^+_{12} \\ Z^+_{21} & Z^+_{22} \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{226}$$

From the second row of equation (225) and the definition in equation (223) we get that

$$Z^+_{21}\Theta_{j-1} + Z^+_{22}\Theta_j = \Xi^u_j = M^{I_{max}}_j \tag{227}$$

In order to show that this expression is equivalent to the solution in equation (224), we substitute (224) into (227) and prove that this gives an identity. Substitution yields:

$$[Z^+_{21} + Z^+_{22}Z_{21}Z^{-1}_{11}]\Theta_{j-1} + [Z^+_{22}(Z_{22} - Z_{21}Z^{-1}_{11}Z_{12})]M^{I_{max}}_j = M^{I_{max}}_j \tag{228}$$

For equation (228) to be an identity, we thus need that

1. $Z^+_{21} + Z^+_{22}Z_{21}Z^{-1}_{11} = 0$ and that
2. $Z^+_{22}(Z_{22} - Z_{21}Z^{-1}_{11}Z_{12}) = I$.

From the definition $Z^+Z = I$ in equation (226) we get the following relations when writing out elements for the second row, first column element and the second row, second column element:

$$Z^+_{21}Z_{11} + Z^+_{22}Z_{21} = 0, \tag{229}$$

and

$$Z^+_{21}Z_{12} + Z^+_{22}Z_{22} = I \tag{230}$$

Using equation (229) above and the fact that Z^{-1}_{11} is invertible, the first condition for (228) to be an identity can be reformulated and simplified as follows:

$$\begin{aligned} Z^+_{21} + Z^+_{22}Z_{21}Z^{-1}_{11} &= 0 \\ Z^+_{21}Z_{11} + Z^+_{22}Z_{21} &= 0 \\ 0 &= 0 \end{aligned} \tag{231}$$

Finally, making use of both equations (229) and (230), the left-hand side of the second condition can be reformulated as:

$$\begin{aligned} Z^+_{22}Z_{22} - Z^+_{22}Z_{21}Z^{-1}_{11}Z_{12} &= I \\ I - Z^+_{21}Z_{12} + Z^+_{21}Z_{11}Z^{-1}_{11}Z_{12} &= I \\ I - Z^+_{21}Z_{12} + Z^+_{21}IZ_{12} &= I \\ I &= I \end{aligned} \tag{232}$$

We have thus shown that the solution in equation (224) retains the identity in the definition equation in (227) and must thus be a valid expression for Θ_j. Equation (224) thus gives a recursive solution for all MA-coefficients of Y_t from $I(\delta)_{max}$

onwards, together with the initial condition in (209). The remaining coefficients for $j < I(\delta)_{max}$ can then be obtained as solutions to the system

$$\begin{pmatrix} \tilde{B}_0 & \tilde{A}_0 & 0 & & \cdots & & 0 \\ \tilde{C}_1 & \tilde{B}_1 & \tilde{A}_1 & 0 & \cdots & & 0 \\ 0 & \tilde{C}_2 & \tilde{B}_2 & \tilde{A}_2 & 0 \quad \cdots & & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & & 0 & \tilde{C}_{I_{\max}-1} & \tilde{B}_{I_{\max}-1} & \tilde{A}_{I_{\max}-1} \\ 0 & & \cdots & & 0 & -(Z_{21}Z_{11}^{-1}) & I_{\max} \end{pmatrix} \begin{pmatrix} \Theta_0 \\ \Theta_1 \\ \Theta_2 \\ \vdots \\ \Theta_{I_{\max}-1} \\ \Theta_{I_{\max}} \end{pmatrix}$$
$$= \begin{pmatrix} \tilde{F}_0 N_1 + \tilde{G}_0 N_0 \\ \tilde{F}_1 N_2 + \tilde{G}_1 N_1 \\ \tilde{F}_2 N_3 + \tilde{G}_2 N_2 \\ \vdots \\ \tilde{F}_{I_{\max}-1} N_{I_{\max}} + \tilde{G}_{I_{\max}-1} N_{I_{\max}-1} \\ (Z_{22} - Z_{21}Z_{11}^{-1}Z_{12}) M_{I_{\max}}^{I_{\max}} \end{pmatrix} \tag{233}$$

5 Appendix: Theoretical Discussion

5.1 Multiple Equilibria

An important issue of the New Keynesian model, either in its sticky price or its sticky information variant, is the question of the existence of one or more stable and unstable equilibria. Three conditions are imposed in this context in order to receive a unique and stable steady state.

First, it is mostly assumed that the household's utility is additively separable in real money balances and consumption, i.e. $U(c, m) = u(c) + v(m)$. Obstfeld (1984) shows that under more general preferences, one gets several convergent equilibria.

Second, the following transversality condition is imposed:

$$\lim_{T \to \infty} \left(\frac{1}{1+r} \right)^t A_t = 0, \tag{234}$$

where A_t is the stock of financial wealth and r is the return on financial wealth. This condition, also called "No-Ponzi"-condition, ensures that the household cannot borrow infinitely in order to consumer more than his lifetime resources. Hence, the transversality condition rules out the possibility of explosive paths in the model.[27]

Third, Inada conditions are imposed, in an explicit way for example in Obstfeld and Rogoff (1983) and McCallum and Nelson (1999):

$$\begin{aligned} \lim_{c \to 0} u'(c) &= \infty \\ \lim_{c \to \infty} u'(c) &= 0 \\ \lim_{m \to 0} v'(m) &= \infty \\ \lim_{m \to \infty} v'(m) &= 0 \end{aligned} \tag{235}$$

[27] See Obstfeld and Rogoff (1983).

Importantly, this means that if real money balances m go to infinity, marginal utility of real balances goes to zero, i.e., the desire for money is assumed to be satiable. This condition ensures that the model has a unique stable equilibrium with positive real money balances. As Ono (2001) shows, if prices are fully flexible, the Inada conditions guarantee that a fully competitive neoclassical equilibrium exists, whereas under sticky prices as in the New Keynesian case, temporary deviations in the short run are possible. However, even if the Inada conditions hold, there is also the possibility of a second stable equilibrium with zero real money balances. This equilibrium can only be ruled out if the "infeasibility condition" (Obstfeld and Rogoff (1983))

$$\lim_{m \to 0} v'(m)m \geq 0 \tag{236}$$

holds with inequality. This condition states that real money balances cannot become negative. Obstfeld and Rogoff (1983) show that (236) with inequality implies $\lim_{m \to 0} v(m) = -\infty$. This means that the utility of real money balances must go to minus infinity if real money balances go to zero. Otherwise, if the infeasibility condition holds with equality, a stable steady state with zero real money balances and hyperinflation exists.

It is important to emphasize the implications of the Inada conditions which is done in detail by Ono (2001). He shows the following: If $\lim_{m \to \infty} v'(m) = \varphi > 0$, i.e. if the household's desire to hold money is insatiable, two different outcomes are possible. Under flexible prices, one does not get an equilibrium at all, since the resulting implosive paths again imply negative real money balances which has been ruled out by the infeasibility condition. On the contrary, under sluggish prices, one gets a unique steady state equilibrium that also satisfies the transversality condition. However, this equilibrium exhibits a lack of demand as a persistent feature, and not as a temporary phenomenon as in the standard New Keynesian case of fulfilled Inada conditions and sluggish prices.[28]

5.2 The Role of Money

One important theoretical question concerns the role money should play in macroeconomic models. Note that we have both used money as an argument in the utility function when deriving the standard model (See equation (9)), and left it out when deriving the sticky information model (See equation (68)). This highlights well the blurry approach of New Keynesian macroeconomics when it comes to dealing with money. Neoclassical and also New Keynesian economists mostly see money only as a medium of exchange, which is used to facilitate the purchase of consumption goods. This explains the difficulties to derive a long-run equilibrium with positive demand for money. As it has been put by Walsh (2003)), p.46: "it should seem strange" that in the money-in-the-utility-function approach (Sidrauski (1967)), "even though the money holdings are never used to purchase consumption, they yield utility". An alternative approach, the cash-in-advance models (Clower (1967)), yields the

[28] Ono (2001) quotes sociological evidence suggesting that people's marginal utility of real money balances does not go to infinity since individuals accumulate money for its own sake, due to reasons of status comparisons for example. Whether the Inada conditions hold or not is ultimately an empirical question, Ono (2001) provides some evidence for Japan suggesting an insatiable desire for money holdings.

same result. The third strand of models, which designs money a role for allocating resources intertemporally in overlapping-generation models (Samuelson (1958)), comes closer to the original Keynesian approach of seeing money also as a store of value. Since in the New Keynesian approach, monetary policy is conducted by the interest rate setting of the central bank, the money supply only enters via the optimality condition which has been derived in (19): The marginal rate of substitution between money and consumption equalizes the opportunity costs of holding money, which then gives an interest-rate elastic money demand function. It is worth emphasizing that this implies that the money supply is endogenous. Since the central bank controls the interest rate, the representative household adjusts its money demand according to the optimality condition, while the central bank passively fulfills this demand. However, money does not affect the equilibrium output and prices in any way. Of course, one could add money into the utility function, however, if one assumes additive separability, it again drops out during the optimization process. If one avoids this assumption, the Euler equation would contain real money balances. Whereas McCallum and Nelson (1999) have argued that empirically, this does not change the results very much, Reis (2007) has recently shown that even in a pure neoclassical framework, money is usually not neutral in the steady state.

In this sense, New Keynesian macroeconomics is purely "real analysis" in the term used by Schumpeter (1954), p.277 who argued that "real analysis proceeds from the principle that all the essential phenomena of economic life are capable of being described in terms of goods and services, of decisions about them, and of relations between them. Money enters the picture only in the modest role of a technical device that has been adopted in order to facilitate transactions". Even if Schumpeter (rightly) further points out that "it has to be recognized that essential features of the capitalist process may depend upon the 'veil'[i.e. money] and that the 'face behind it' is incomplete without it", this appeal is mostly ignored in New Keynesian macroeconomics. A further simplification concerning monetary issues is introduced with respect to financial assets. Assuming one nominal interest rate, New Keynesian models work with one interest-bearing asset, government bonds, in addition to a non-interest bearing asset, money. It is worth mentioning that this overly simplistic treatment of financial issues has been subject to a fundamental critique by Greenwald and Stiglitz (1993); Stiglitz and Greenwald (2003) and also Bernanke et al. (1996, 1998). These authors have argued that the non-neutrality of money does not stem from sticky prices or wages, but from the special characteristics of the credit market. Under conditions of asymmetric information in this market, making prices and wages more flexible can even worsen an economic downturn.[29] Summing up, the treatment of monetary issues in the New Keynesian model is highly incomplete and should thus be subject to further research, a first new starting point has recently been proposed by Christiano et al. (2007).

5.3 The Role of Capital Accumulation

Generally, capital accumulation is not included in the standard New Keynesian Model, as was also the case in the Hicksian IS-LM-model. In the case of the latter, this was explained by its focus on a short-run time horizon. This, however, is not the case in the New Keynesian Model with its mostly adopted infinite time

[29] See for a similar argument Ono (1994), Hahn and Solow (1997) and Keynes (1936).

horizon. In this new model, neglecting capital accumulation is justified by the fact that "adding investment and capital to the model (...) does not change the fundamental qualitative aspects: output demand still depends inversely on the real rate and positively on expected future output."[30] Whereas in the seminal paper by McCallum and Nelson (1999), the absence of capital accumulation is justified by analytical simplicity, given the empirical finding that there is not much cyclical variation of the capital stock, King (2000) sees this neglect very critically, since he claims that inflation shocks cannot be understood properly without explicitly modeling investment behavior. Although Woodford (2003), p.357, has provided an analysis of price setting under an endogenous capital stock, his way of tackling the problem has been subject to criticism by Sveen and Weinke (2004). They argue that capital accumulation affects both inflation and output dynamics by its effects on firms' marginal costs.

Moreover, introducing capital accumulation into the analysis does not only have effects on firms' price setting but also on the role of consumption smoothing. This has gained new prominence due to the introduction of rule-of-thumb consumers into the standard model. The argument behind this goes as follows.

In neoclassical consumption theory, consumers smooth their consumption with regard to the expected growth of future income. However, this future income is considered to be non-risky, i.e., the risk of getting unemployed in the future and thus experiencing an income of zero is either not modeled at all or assumed to be idiosyncratic and thus diversifiable.[31] If labor income is not treated as diversifiable and thus becomes an aggregate risk, this makes consumption smoothing less prevalent. If a higher expected income growth comes along with a higher variance of future income, this leads to precautionary saving, i.e., the consumer tries to insure himself against this additional risk by consuming less today. This then works as a self-imposed credit restriction: The consumer does not borrow against his expected future income due to its riskiness.[32]

In the New Keynesian approach, this topic has been dealt with in two ways: In the models by Woodford[33], perfect financial markets have been assumed, implying that labor income risk is diversifiable. In models where this assumption has not been made explicitly, a first-order Taylor approximation around the Euler equation has been taken in order to derive the New Keynesian IS-curve.[34] By contrast, introducing rule-of-thumb consumers, Galí et al. (2004) claim that this makes it necessary to introduce capital accumulation in order to have an explicit distinction between a rational, optimizing and consumption smoothing agent when facing risk, and a non-optimizing, non-consumption smoothing rule-of-thumb consumer, who consumes his current income in every period.[35] To sum up, in order to get an

[30] Clarida et al. (1999) p.1666.

[31] In the seminal paper by Hall (1978), this assumption is made explicit by using a quadratic utility function which makes the variance of future income drop out during the optimization process. The assumption of a quadratic utility function has been highly criticized by later authors (Blanchard and Mankiw (1988)) and has been replaced by CARA or CRRA utility functions with more adequate properties.

[32] See Carroll and Kimball (2008).

[33] E.g. Woodford (1996) and Rotemberg and Woodford (1997)

[34] However, Kimball (1990) has shown that one needs a second-order approximation in order to model non-diversifiable risk.

[35] "Notice that in the absence of capital accumulation, the only difference in behavior across household types would be a consequence of the fact that Ricardian households obtain some dividend

important role for consumption smoothing, one either has to assume perfect financial markets, i.e. to assume that future labor income is diversifiable, or one has to introduce capital accumulation in order to have a means of consumption smoothing.

References

Arslan, M. M. (2007, August). Dynamics of Sticky Information and Sticky Price Models in a New Keynesian DSGE Framework. *MPRA Paper 5269.*

Bagliano, F.-C. and G. Bertola (2004, February). *Models for Dynamic Macroeconomics.* Oxford: Oxford University Press.

Ball, L., G. Mankiw, and R. Reis (2003, February). Monetary Policy for Inattentive Economies. *Harvard Institute of Economic Research Discussion Paper 1997.*

Ball, L., N. G. Mankiw, and R. Reis (2005, May). Monetary Policy for Inattentive Economies. *Journal of Monetary Economics 52*(4), 703–725.

Bernanke, B., M. Gertler, and S. Gilchrist (1996, February). The Financial Accelerator and the Flight to Quality. *Review of Economics and Statistics 78*(1), 1–15.

Bernanke, B., M. Gertler, and S. Gilchrist (1998, March). The Financial Accelerator in a Quantitative Business Cycle Framework. *NBER Working Paper 6455.*

Blanchard, O. and G. Mankiw (1988, May). Consumption: Beyond Certainty Equivalence. *American Economic Review 78*(2), 173–178.

Blanchard, O. J. and S. Fischer (1989, March). *Lectures on Macroeconomics.* Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Blanchard, O. J. and C. M. Kahn (1980, July). The Solution of Linear Difference Models under Rational Expectation. *Econometrica 48*(5), 1305–1311.

Calvo, G. (1983, September). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics 12*(3), 383–398.

Carroll, C. D. and M. Kimball (2008, May). Precautionary Saving and Precautionary Wealth. In S. N. Durlauf and L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics* (2nd ed.). Houndmills, Hampshire: Palgrave Macmillan.

Christiano, L., M. Trabandt, and K. Walentin (2007, November). Introducing Financial Frictions and Unemployment into a Small Open Economy Model. *Sveriges Riksbank Working Paper 214.*

Clarida, R., J. Galí, and M. Gertler (1999, December). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature 37*(4), 1661–1707.

income from the ownership of firms. Yet, the existence of a market for firms' stocks would not provide a means to smooth consumption, since, in equilibrium, all shares would have to be held by Ricardian consumers in equal proportions." (Galí et al. (2004))

Clarida, R., J. Galí, and M. Gertler (2001, May). Optimal Monetary Policy in Open vs. Closed Economies. *American Economic Review 91*(2), 248–252.

Clower, R. (1967, December). A Reconsideration of the Microfoundations of Monetary Theory. *Western Economic Journal 6*(1), 1–9.

De Grauwe, P. (2008, May). DSGE-Modelling When Agents are Imperfectly Informed. *ECB Working Paper 897*.

Fuhrer, J. (2000, June). Habit Formation in Consumption and Its Implications for Monetary-Policy Models. *American Economic Review 90*(3), 367–390.

Fuhrer, J. and G. Moore (1995, February). Inflation Persistence. *Quarterly Journal of Economics 110*(1), 127–159.

Galí, J., J. Lopez-Salido, and J. Valles (2004, August). Rule-of-Thumb Consumers and the Design of Interest Rate Rules. *Journal of Money, Credit, and Banking 36*(4), 739–763.

Greenwald, B. and J. Stiglitz (1993, Winter). New and Old Keynesians. *Journal of Economic Perspectives 7*(1), 23–44.

Hahn, F. and R. M. Solow (1997, August). *A Critical Essay on Modern Macroeconomic Theory*. Massachusetts: MIT Press.

Hall, R. (1978, December). Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy 86*(6), 971–987.

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. London: Palgrave Macmillan.

Kimball, M. (1990, January). Precautionary Saving in the Small and in the Large. *Econometrica 58*(1), 53–73.

King, R. G. (2000, Summer). The New IS-LM Model: Language, Logic, and Limits. *Federal Reserve Bank of Richmond Economic Quarterly 86*(3), 45–103.

Klein, P. (2000, September). Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics & Control 24*(10), 1405 – 1423.

Mankiw, N. G. and R. Reis (2006a, February). Pervasive Stickiness. *NBER Working Paper 12024*.

Mankiw, N. G. and R. Reis (2006b, May). Pervasive Stickiness. *American Economic Review 96*(2), 164–169.

Mankiw, N. G. and R. Reis (2006c, October). Sticky Information in General Equilibrium. *NBER Working Paper 12605*.

Mankiw, N. G. and R. Reis (2007, April-May). Sticky Information in General Equilibrium. *Journal of the European Economic Association 5*(2-3), 603–613.

McCallum, B. T. (1983, November). On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective. *Journal of Monetary Economics 11*(2), 139–168.

McCallum, B. T. (1998, November). Solutions to Linear Rational Expectations Models: a Compact Exposition. *Economics Letters 61*(2), 143–147.

McCallum, B. T. (2001, May). Should Monetary Policy Respond Strongly to Output Gaps? *American Economic Review 91*(2 (Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association)), 258–262.

McCallum, B. T. and E. Nelson (1999, August). An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis. *Journal of Money, Credit, and Banking 31*(3, Part 1), 296–316.

McCandless, G. (2008, March). *The ABCs of RBCs - An Introduction to Dynamic Macroeconomic Models*. Harvard: Harvard University Press.

Meyer-Gohde, A. (2007, December). Solving Linear Rational Expectations Models with Lagged Expectations Quickly and Easily. *Technische Universität Berlin Department of Economics Section Macroeconomics*.

Meyer-Gohde, A. (2009, April). Linear Rational Expectations Models with Lagged Expectations: A Synthetic Method. *Technische Universität Berlin Department of Economics Section Macroeconomics*.

Obstfeld, M. (1984, January). Multiple Stable Equilibria in an Optimizing Perfect-Foresight Model. *Econometrica 52*(1), 223–228.

Obstfeld, M. and K. Rogoff (1983, August). Speculative Hyperinflations in Maximizing Models: Can We Rule Them Out? *Journal of Political Economy 91*(4), 675–687.

Ono, Y. (1994, December). *Money, Interest and Stagnation: Dynamic Theory and Keynes's Economics*. Oxford: Oxford University Press.

Ono, Y. (2001, February). A Reinterpretation of Chapter 17 of Keynes's General Theory: Effective Demand Shortage under Dynamic Optimization. *International Economic Review 42*(1), 207–236.

Reis, R. (2006a, November). Inattentive Consumers. *Journal of Monetary Economics 53*(8), 1761–1800.

Reis, R. (2006b, July). Inattentive Producers. *Review of Economic Studies 73*(3), 793–821.

Reis, R. (2007, January). The Analytics of Monetary Non-Neutrality in the Sidrauski Model. *Economics Letters 94*(1), 129–135.

Rinne, H. (2004, June). *Ökonometrie*. Munich: Vahlen.

Rotemberg, J. J. and M. Woodford (1997, January). An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy. *NBER Macroeconomics Annual 12*, 297–346.

Samuelson, P. A. (1958, December). An Exact Consumption-Loan Model of Interest With or Without the Social Contrivance of Money. *Journal of Political Economy 66*(6), 467–482.

Schumpeter, J. (1954). *History of Economic Analysis*. Oxford: Oxford University Press.

Sidrauski, M. (1967, May). Rational Choice and Patterns of Growth in a Monetary Economy. *American Economic Review 57*(2 (Papers and Proceedings of the Seventy-ninth Annual Meeting of the American Economic Association)), 534–544.

Stiglitz, J. and B. Greenwald (2003, October). *Towards a New Paradigm in Monetary Economics*. Raffaele Mattioli Lectures. Cambridge: Cambridge University Press.

Sveen, T. and L. Weinke (2004, September). Pitfalls in the Modeling of Forward-Lookig Price Setting and Investment Decisions. *Universitat Pompeu Fabra Economics Working Paper 773*.

Trabandt, M. (2007, July). Sticky Information vs. Sticky Prices: A Horse Race in a DSGE Framework. *Kiel Institute for the World Economy Working Paper 1369*.

Uhlig, H. (1995). A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily. *Tilburg University - Center for Economic Research - Discussion Paper 9597*.

Walsh, C. E. (2003, May). *Monetary Theory and Policy* (2 ed.). Cambridge and London: MIT Press.

Woodford, M. (1996, July). Control of the Public Debt: A Requirement for Price Stability? *NBER Working Paper 5684*.

Woodford, M. (2003, August). *Interest and Prices: Foundations of a Theory of Monetary Policy* (1 ed.). Princeton, New Jersey: Princeton University Press.