Chapter 2

Perturbation Methods

Overview. In the previous chapter we have seen that the solution of a DGE model with a representative agent is given by a set of policy functions that relate the agent's choice variables to the state variables that characterize the agent's economic environment. In this chapter we explore methods that use local information to obtain either a linear or a quadratic approximation of the agent's policy function. To see what this means, remember from elementary calculus that a straight line that is tangent to a function $y = f(x)$ at x^* locally approximates f: according to Taylor's theorem (see Section 11.2.1) we may write

$$f(x^* + h) = \underbrace{f(x^*) + f'(x^*)h}_{\text{linear function in } h} + \phi(h),$$

where the error $\phi(h)$ has the property

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{\phi(h)}{h} = 0.$$

Thus, close to x^*, f equals a slightly perturbed linear function. To set up the linear function, we only need to know (i) the value of f at x^* and (ii) the value of its first derivative f' at the same point.

Probably less well known is the following result. If $x_t = f(x_{t-1})$ is a non-linear difference equation and $\bar{x}_t = f'(x^*)\bar{x}_{t-1}$, $\bar{x}_t = x_t - x^*$ its linear approximation at x^* defined by $x^* = f(x^*)$, then the solution of the linear model provides a local approximation of the solution of the non-linear equation.[1] Perturbation methods

[1] See Section 12.1 on difference equations, if you are unfamiliar with this subject.

rest on these observations. As we will see, they are not limited to linear approximations. If f is n-times continuously differentiable, we may use a polynomial in h of degree up to $n-1$ to build a local model of f.

In this chapter we mainly consider linear approximations. They are the most frequently used solutions in applied research and are easy to apply. As you will see in later chapters they also provide a first guess for more advanced, non-local methods.

The next section considers deterministic models. In this context it is relatively easy to demonstrate by means of an example (the Ramsey model of Section 1.2) that we can get linear approximations to the policy functions by either solving the linearized system of Euler equations or by applying the implicit function theorem to the steady state conditions of the model. We use this result to provide a procedure that computes the solution of an arbitrary deterministic model with n variables from the linearized system of Euler equations.

Before we turn to the solution of stochastic DGE models in Sections 2.3 and 2.4, we consider a model where the linear policy functions provide an exact solution. This is the linear-quadratic (LQ) model outlined in Section 2.2. Two different approximation methods derive from the LQ problem. The first approach, considered in Section 2.3, approximates a given model so that its return function is quadratic and the law of motion is linear and solves the approximate model by value function iterations. The second approach, taken up in Section 2.4, relies on a linear approximation of the model's Euler equations and solves the ensuing system of linear stochastic difference equations.

We close the methodological part of this chapter in Section 2.5 with the quadratic approximation of the policy functions of an arbitrary stochastic DGE model. The bottom line of Sections 2.3 through 2.5 are three programs: `SolveLA` and `SolveLQA` compute linear approximations to deterministic as well as stochastic DGE models. The difference between the two programs is the way you must set up your model. `SolveLA` is a general purpose routine, while `SolveLQA` is limited to models whose solution can be obtained by solving a central planing problem. Yet, in some kinds of problems it is much easier to cast your model into the framework of `SolveLQA`. The third program, `SolveQA`, computes quadratic approximations of the policy functions of an arbitrary DGE model. Various applications illustrate the use of these programs in Section 2.6.

2.1 Linear Solutions for Deterministic Models

This Section applies two tools. The implicit function theorem, sketched in Section 11.2.2, allows us to compute the derivatives of a system of policy functions that is implicitly determined by a system of non-linear Euler equations. The close relation between the local solution of a system of non-linear, first-order difference equations and the solution of the related linearized system, outlined in Section 12.1, provides a second route to compute linear approximations of a model's policy functions. If you are unfamiliar with any of these tools, you might consider reading the respective sections before proceeding.

We use the deterministic growth model from Section 1.2 to illustrate both techniques before we turn to the general approach. We begin with the solution of the system of non-linear difference equations that governs the model's dynamics.

Approximate Computation of the Saddle Path. Consider equations (1.17) that characterize the solution of the Ramsey problem (1.8) from Section 1.2:

$$K_{t+1} - f(K_t) + C_t =: g^1(K_t, C_t, K_{t+1}, C_{t+1}) = 0, \tag{2.1a}$$

$$u'(C_t) - \beta u'(C_{t+1}) f'(K_{t+1}) =: g^2(K_t, C_t, K_{t+1}, C_{t+1}) = 0. \tag{2.1b}$$

Equation (2.1a) is the farmer's resource constraint.[2] It states that seed available for the next period K_{t+1} equals production $f(K_t)$

[2] Remember, that in the notation of Section 1.2 $f(K) := (1-\delta)K + F(N, K)$, where N are the farmer's exogenously given working hours.

minus consumption C_t. The first-order condition with respect to the next-period stock of capital K_{t+1} is equation (2.1b). These two equations implicitly specify a non-linear system of difference equations $\mathbf{x}_{t+1} = \mathbf{\Psi}(\mathbf{x}_t)$ in the vector $\mathbf{x}_t := [K_t, C_t]'$:

$$\mathbf{g}(\mathbf{x}_t, \mathbf{\Psi}(\mathbf{x_t})) = \mathbf{0}_{2\times 1},\ \mathbf{g} = [g^1, g^2]'.$$

The stationary solution defined by

$$1 = \beta f'(K^*), \tag{2.2a}$$
$$K^* = f(K^*) - C^* \tag{2.2b}$$

is a fixed point of $\mathbf{\Psi}$. We obtain the linear approximation of $\mathbf{\Psi}$ at $\mathbf{x}^* = [K^*, C^*]'$ via equation (11.38):

$$\bar{\mathbf{x}}_{t+1} = J(\mathbf{x}^*)\bar{\mathbf{x}}_t, \quad \bar{\mathbf{x}}_t := \mathbf{x}_t - \mathbf{x}^*. \tag{2.3}$$

with the Jacobian matrix J determined by

$$J(\mathbf{x}^*) = \begin{bmatrix} \frac{\partial g^1(\mathbf{x}^*,\mathbf{x}^*)}{\partial K_{t+1}} & \frac{\partial g^1(\mathbf{x}^*,\mathbf{x}^*)}{\partial C_{t+1}} \\ \frac{\partial g^2(\mathbf{x}^*,\mathbf{x}^*)}{\partial K_{t+1}} & \frac{\partial g^2(\mathbf{x}^*,\mathbf{x}^*)}{\partial C_{t+1}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial g^1(\mathbf{x}^*,\mathbf{x}^*)}{\partial K_t} & \frac{\partial g^1(\mathbf{x}^*,\mathbf{x}^*)}{\partial C_t} \\ \frac{\partial g^2(\mathbf{x}^*,\mathbf{x}^*)}{\partial K_t} & \frac{\partial g^2(\mathbf{x}^*,\mathbf{x}^*)}{\partial C_t} \end{bmatrix}. \tag{2.4}$$

The derivatives of $\mathbf{g}$ at the fixed point are easily obtained from (2.1a) and (2.1b) (we suppress the arguments of the functions and write f' instead of $f'(K^*)$ and so forth):

$$J(\mathbf{x}^*) = -\begin{bmatrix} 1 & 0 \\ -\beta u' f'' & -u'' \end{bmatrix}^{-1} \begin{bmatrix} -\frac{1}{\beta} & 1 \\ 0 & u'' \end{bmatrix} = \begin{bmatrix} \frac{1}{\beta} & -1 \\ -\frac{u'f''}{u''} & 1 + \frac{\beta u' f''}{u''} \end{bmatrix}.$$

In computing the matrix on the rhs of this equation we used the definition of the inverse matrix given in (11.14). The eigenvalues λ_1 and λ_2 of J satisfy (see (11.24)):

$$\det J = \frac{1}{\beta} = \lambda_1 \lambda_2,$$
$$\operatorname{tr} J = \underbrace{1 + \frac{1}{\beta} + \frac{\beta u' f''}{u''}}_{=:\Delta} = \lambda_1 + \lambda_2.$$

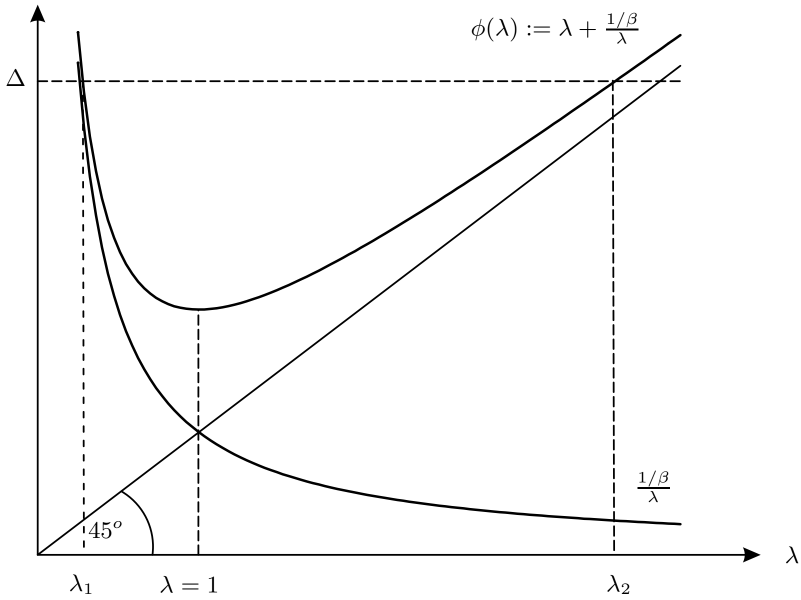

Figure 2.1: Eigenvalues of W

Therefore, they solve equation

$$\phi(\lambda) := \lambda + \frac{1/\beta}{\lambda} = \Delta.$$

The solutions are the points of intersection between the horizontal line through Δ and the hyperbola $\phi(\lambda)$ (see Figure 2.1). The graph of ϕ obtains a minimum at $\lambda_{min} = 1/\sqrt{\beta} > 1$, where $\phi'(\lambda_{min}) = 1 - (1/\beta)\lambda^{-2} = 0$.[3] Since $\phi(1) = 1 + (1/\beta) < \Delta$, there must be one intersection to the right of $\lambda = 1$ and one to the left, proving that J has one real eigenvalue $\lambda_1 < 1$ and another real eigenvalue $\lambda_2 > 1$.

Let $J = TST^{-1}$ with

$$S = \begin{bmatrix} \lambda_1 & s_{12} \\ 0 & \lambda_2 \end{bmatrix}$$

denote the Schur factorization of J (see (11.27) in Section 11.1.8). In the new variables (where $T^{-1} = (t^{ij})$)

[3] In Figure 2.1 λ_{min} is so close to $\lambda = 1$ that we do not show it.

$$\mathbf{y}_t = T^{-1}\bar{\mathbf{x}}_t \Leftrightarrow \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} t^{11} & t^{12} \\ t^{21} & t^{22} \end{bmatrix} \begin{bmatrix} K_t - K^* \\ C_t - C^* \end{bmatrix} \tag{2.5}$$

the system of equations (2.3) is given by

$$\mathbf{y}_{t+1} = S\mathbf{y}_t.$$

The second line of this matrix equation is

$$y_{2t+1} = \lambda_2 y_{2t}.$$

Since $\lambda_2 > 1$, the variable y_{2t} will diverge unless we set $y_{20} = 0$. This restricts the system to the stable eigenspace. Using $y_{2t} = 0$ in (2.5) implies

$$0 = t^{21}\bar{x}_{1t} + t^{22}\bar{x}_{2t}, \tag{2.6a}$$

$$y^{1t} = (t^{11} - t^{12}(t^{21}/t^{22}))x_{1t}. \tag{2.6b}$$

The first line is the linearized policy function for consumption:

$$C_t - C^* = -\frac{t^{21}}{t^{22}}\left[K_t - K^*\right]. \tag{2.7a}$$

The second line of (2.6) implies via $y_{1t+1} = \lambda_1 y_{1t}$ the linearized policy function for savings:

$$K_{t+1} - K^* = \lambda_1\left[K_t - K^*\right]. \tag{2.7b}$$

We illustrate these computations in the program `Ramsey2a.g`, where we use $u(C) = [C^{1-\eta} - 1]/(1-\eta)$ and $F(N, K) = K^\alpha$. In this program we show that it is not necessary to compute the Jacobian matrix analytically as we have done here. You may also write a procedure that receives the vector $[K_t, C_t, K_{t+1}, C_{t+1}]'$ as input and that returns the rhs of equations (2.1). This procedure can be passed to a routine that numerically evaluates the partial derivatives at the point (K^*, C^*, K^*, C^*). From the output of this procedure you can extract the matrices that appear on the rhs of equation (2.4).

Figure 2.2 compares the time path of the capital stock under the analytic solution $K_{t+1} = \alpha\beta K_t^\alpha$ (which requires $\eta = \delta = 1$)

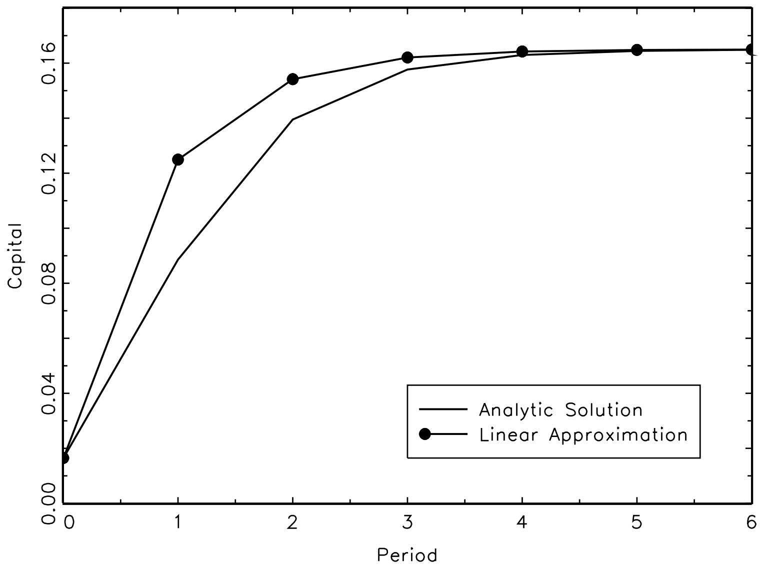

Figure 2.2: Approximate Time Path of the Capital Stock in the Deterministic Growth Model

with the path obtained from the approximate linear solution. The parameters are set equal to $\alpha = 0.27$ and $\beta = 0.994$, respectively. The initial capital stock equals one-tenth of the stationary capital stock. As we would have expected, far from the fixed point, the linear approximation is not that good. Yet, after about five iterations it is visually indistinguishable from the analytic solution.

Approximate Policy Functions. We now apply the implicit function theorem directly to find the linear approximation of the policy function for optimal savings. Let $K_{t+1} = h(K_t)$ denote this function. Since $K^* = h(K^*)$, its linear approximation at K^* is given by

$$K_{t+1} = h(K_t) \simeq K^* + h'(K^*)(K_t - K^*). \tag{2.8}$$

Substituting equation (2.1a) for $C_t = f(K_t) - h(K_t)$ into equation (2.1b) delivers:

$$\begin{aligned} g(K_t) := u'\left[(f(K_t) - h(K_t)\right] \\ - \beta u'\left[(f(h(K_t)) - h(h(K_t))\right] f'(h(K_t)). \end{aligned}$$

We know that $g(K^*) = 0$. Theorem 11.2.3 allows us to compute $h'(K^*)$ from $g'(K^*) = 0$. Differentiating with respect to K_t and evaluating the resulting expression at K^* provides the following quadratic equation in $h'(K^*)$ (we suppress the arguments of all functions):

$$(h')^2 - \underbrace{(1 + (1/\beta) + (\beta u' f'')/u'')}_{=:\Delta} h' + (1/\beta) = 0 \tag{2.9}$$

Let h'_1 and h'_2 denote the solutions. Since (by Viète's rule) $h'_1 + h'_2 = \Delta$ and $h'_1 h'_2 = 1/\beta$, the solutions of equation (2.9) equal the eigenvalues of the Jacobian matrix λ_1 and λ_2 obtained in the previous paragraph. The solution is, thus, given by $h'(K^*) = \lambda_1$ and the approximate policy function coincides with equation (2.7a). Note that we actually do not need to compute the approximate policy function for consumption: given the approximate savings function (2.7a) we obtain the solution for consumption directly from the resource constraint (2.1a).

Observe further that this way to compute $h'(K^*)$ is less readily implemented on a computer. In order to set up (2.9) we need software that is able to do symbolic differentiation. Our general procedure for non-linear, deterministic DGE models therefore relies on the approach considered in the previous paragraph.

The General Method. It is straightforward to generalize the method outlined above to compute the linear approximate solution of a non-linear system of difference equations implied by a deterministic DGE model. Suppose the map

$$\mathbf{g}(\mathbf{x}_t, \mathbf{x}_{t+1}) = \mathbf{0}_{n\times 1}, \quad \mathbf{x}_t \in \mathbb{R}^n$$

implicitly describes the model's dynamics. Assume, further, that n_1 of the elements in $\mathbf{x}_t$ have given initial conditions (as the capital stock in the deterministic growth model) and that $n_2 = n - n_1$ are jump variables (as consumption), whose initial conditions must be chosen in order to satisfy the model's transversality conditions. Let $\mathbf{x}^*$ denote the fixed point. Since the analytic computation of the Jacobian matrix is usually very cumbersome and failure prone, it is advisable to write a procedure that returns the rhs of $\mathbf{g}(\mathbf{x}_t, \mathbf{x}_{t+1})$. This procedure serves as input to a program that performs numeric differentiation. Given the matrices $A := \mathbf{g}_{\mathbf{x}_{t+1}}(\mathbf{x}^*, \mathbf{x}^*)$ and $B := \mathbf{g}_{\mathbf{x}_t}(\mathbf{x}^*, \mathbf{x}^*)$, the Jacobian matrix of the linearized system (2.3) is given by $J = A^{-1}B$. This matrix must have n_1 eigenvalues inside and n_2 eigenvalues outside the unit circle.

Let $\mathbf{y}_t := T^{-1}\bar{\mathbf{x}}_t$ with $J = TST^{-1}$ denote the new variables in which the system is decoupled

$$\begin{bmatrix} \mathbf{y}_{1t+1} \\ \mathbf{y}_{2t+1} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ 0_{n_2 \times n_1} & S_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y}_{1t} \\ \mathbf{y}_{2t} \end{bmatrix}.$$

Since all the eigenvalues on the main diagonal of S_{22} are outside the unitUnit circle circle, we must set $\mathbf{y}_{2t} = \mathbf{0}_{n_2 \times 1}$ to secure convergence. Thus, the second block of the matrix equation

$$\begin{bmatrix} \mathbf{y}_{1t} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix} = \begin{bmatrix} T^{11} & T^{12} \\ T^{21} & T^{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{x}}_{1t} \\ \bar{\mathbf{x}}_{2t} \end{bmatrix}$$

implies the policy function for the jump variables:

$$\bar{\mathbf{x}}_{2t} = -(T^{22})^{-1}T^{21}\bar{\mathbf{x}}_{1t}. \tag{2.10a}$$

Using this result to substitute for $\bar{\mathbf{x}}_{2t}$ in the first block of equations yields:

$$\mathbf{y}_{1t} = \left(T^{11} - T^{12}(T^{22})^{-1}T^{21}\right)\bar{\mathbf{x}}_{1t}.$$

Observe that the inverse of the matrix in parenthesis is T_{11} (apply the formula for the inverse of a partitioned matrix (11.15a) to the matrix T^{-1}). Thus,

$$\mathbf{y}_{1t+1} = (T_{11})^{-1}\bar{\mathbf{x}}_{1t+1} = S_{11}\mathbf{y}_{1t} = S_{11}T_{11}^{-1}\bar{\mathbf{x}}_{1t}$$

so that the policy function for $\bar{\mathbf{x}}_{1t+1}$ is given by

$$\bar{\mathbf{x}}_{1t+1} = T_{11}S_{11}T_{11}^{-1}\bar{\mathbf{x}}_{1t}. \tag{2.10b}$$

You will see in Section 2.4 that our procedure `SolveLA` that computes the linear approximate solution of stochastic DGE models provides the policy functions (2.10) as a special case.

2.2 The Stochastic Linear Quadratic Model

This section presents the stochastic linear quadratic model and derives some of its important properties. Since its main purpose is to provide a framework for both linear quadratic and linear approximation methods, we postpone detailed algorithms for the computation of the policy function until Section 2.3 and Section 2.4, respectively.

Description. Consider an economy governed by the following stochastic linear law of motion:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \boldsymbol{\epsilon}_t. \tag{2.11}$$

The n-dimensional column vector $\mathbf{x}_t$ holds those variables that are predetermined at period t. A fictitious social planner sets the values of the variables stacked in in the m-dimensional column vector $\mathbf{u}_t$. We refer to $\mathbf{x}$ as the state vector and to $\mathbf{u}$ as the control vector. $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are matrices. Due to the presence of shocks, the planner cannot control this economy perfectly. The n vector of shocks $\boldsymbol{\epsilon}$ has a multivariate normal distribution with $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and covariance matrix[4] $E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \Sigma$. The planner must choose $\mathbf{u}_t$ before he can realize the size of the shocks.

Given $\mathbf{x}_0$ the planner's objective is to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\mathbf{x}_t' Q\mathbf{x}_t + \mathbf{u}_t' R\mathbf{u}_t + 2\mathbf{u}_t' S\mathbf{x}_t\right], \quad \beta \in (0,1), \tag{2.12}$$

subject to (2.11). The current period objective function

$$g(\mathbf{x}_t, \mathbf{u}_t) := \left[\mathbf{x}_t', \quad \mathbf{u}_t'\right] \begin{bmatrix} Q & S' \\ S & R \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{u}_t \end{bmatrix} \tag{2.13}$$

is quadratic and concave in $(\mathbf{x}_t', \mathbf{u}_t')$. This requires that both the symmetric $n \times n$ matrix Q and the symmetric $m \times m$ matrix R are negative semidefinite.

Note that this specification encompasses non-stochastic state variables and first-order (vector) autoregressive processes.

[4] Remember that a prime denotes transposition, i.e., $\boldsymbol{\epsilon}'$ is a row vector and $\boldsymbol{\epsilon}$ a column vector.

Derivation of the Policy Function. The Bellman equation for the stochastic LQ problem is given by

$$\begin{aligned} v(\mathbf{x}) := \max_{\mathbf{u}} \quad & \mathbf{x}'Q\mathbf{x} + 2\mathbf{u}'S\mathbf{x} + \mathbf{u}'R\mathbf{u} \\ & + \beta E\left[v(A\mathbf{x} + B\mathbf{u} + \boldsymbol{\epsilon})\right], \end{aligned} \tag{2.14}$$

where we used (2.11) to replace next-period state variables in $Ev(\cdot)$ and where we dropped the time indices for convenience, because all variables refer to the same date t. Expectations are taken conditional on the information contained in the current state $\mathbf{x}$. We guess that the value function is given by $v(\mathbf{x}) := \mathbf{x}'P\mathbf{x} + d$, P being a n dimensional symmetric, negative semidefinite square matrix and $d \in \mathbb{R}$ an unknown constant.[5] Thus, we may write (2.14) as follows:[6]

$$\begin{aligned} &\mathbf{x}'P\mathbf{x} + d = \\ &\max_{\mathbf{u}} \quad \mathbf{x}'Q\mathbf{x} + 2\mathbf{u}'S\mathbf{x} + \mathbf{u}'R\mathbf{u} \\ &\qquad + \beta E\left[((A\mathbf{x} + B\mathbf{u} + \boldsymbol{\epsilon})'P(A\mathbf{x} + B\mathbf{u} + \boldsymbol{\epsilon}) + d)\right]. \end{aligned} \tag{2.15}$$

Evaluating the conditional expectations on the rhs of (2.15) yields:

$$\begin{aligned} &\mathbf{x}'P\mathbf{x} + d = \\ &\max_{\mathbf{u}} \quad \mathbf{x}'Q\mathbf{x} + 2\mathbf{u}S\mathbf{x} + \mathbf{u}'R\mathbf{u} \\ &\qquad + \beta\mathbf{x}'A'PA\mathbf{x} + 2\beta\mathbf{x}'A'PB\mathbf{u} + \beta\mathbf{u}'B'PB\mathbf{u} \\ &\qquad + \beta\,\mathrm{tr}(P\Sigma) + \beta d. \end{aligned} \tag{2.16}$$

In the next step we differentiate the rhs of (2.16) with respect to the control vector $\mathbf{u}$, set the result equal to the zero vector, and solve for $\mathbf{u}$. This provides the solution for the policy function:

---

[5] Note, since $\mathbf{x}_t'P\mathbf{x}_t$ is a quadratic form, it is not restrictive to assume that P is symmetric. Furthermore, since the value function of a well defined dynamic programming problem is strictly concave, P must be negative semidefinite.

[6] If you are unfamiliar with matrix algebra, you may find it helpful to consult Section 11.1. We present the details of the derivation of the policy function in Appendix 3.

$$\mathbf{u} = -\underbrace{(R + \beta B'PB)^{-1}(S + \beta B'PA)}_{F}\mathbf{x}. \tag{2.17}$$

To find the solution for the matrix P and the constant d, we eliminate $\mathbf{u}$ from the Bellman equation (2.16) and compare the quadratic forms and the constant terms on both sides. It turns out that P must satisfy the following implicit equation, known as algebraic matrix Riccati equation:

$$\begin{aligned} P = {} & Q + \beta A'PA \\ & - (S + \beta B'PA)' \left[R + \beta B'PB\right]^{-1} (S + \beta B'PA), \end{aligned} \tag{2.18}$$

and that d is given by:

$$d = \frac{\beta}{1-\beta}\,\mathrm{tr}(P\Sigma).$$

The solution of (2.18) can be obtained by iterating on the matrix Riccati difference equation

$$\begin{aligned} P_{s+1} = {} & Q + \beta A'P_sA \\ & - (S + \beta B'P_sA)' \left[R + \beta B'P_sB\right]^{-1} (S + \beta B'P_sA) \end{aligned}$$

starting with some initial negative definite matrix P_0.[7] Other methods to solve (2.18) rely on matrix factorizations. Since we will use iterations over the value function later on, we will not explore these methods any further. Once the solution for P has been computed, the dynamics of the model is governed by

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \boldsymbol{\epsilon}_{t+1} = (A - FB)\mathbf{x}_t + \boldsymbol{\epsilon}_t.$$

Certainty Equivalence. The solution of the stochastic LQ problem has a remarkable feature. Since the covariance matrix of the shocks Σ appears neither in equation (2.17) nor in equation (2.18), the optimal control is independent of the stochastic properties of the model summarized by Σ. Had we considered a deterministic

[7] For example $P_0 = -0.01 I_n$.

linear quadratic problem by assuming $\boldsymbol{\epsilon}_t = \mathbf{0} \forall t$, we would have found the same feedback rule (2.17). You may want to verify this claim by solving Problem 2.1. This property of the stochastic LQ problem is called certainty equivalence principle. It is important to note that if we use the LQ approximation to solve an arbitrary economic model we enforce the certainty equivalence principle on this solution. This may hide important properties of the model. For instance, consider two economies A and B which are identical in all respects except for the size of their productivity shocks. If economy's A shock has a much larger standard deviation than economy B's shock, it is hard to believe that the agents in both economies use the same feed-back rules.

Derivation of the Euler Equations. As we have seen in Chapter 1 an alternative way to derive the dynamic path of an optimizing model is to consider the model's Euler equations. It will be helpful for the approach taken in Section 2.4 to separate the state variables into two categories. Variables that have a given initial condition but are otherwise determined endogenously are stacked in the n dimensional vector $\mathbf{x}$. Purely exogenous shocks are summarized in the l dimensional vector $\mathbf{z}$. As in the previous subsection $\mathbf{u}$ is the m dimensional vector of controls. The planner's current period return function is the following quadratic form:

$$\begin{aligned} g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) := \mathbf{x}_t' A_{xx} \mathbf{x}_t + \mathbf{u}_t' A_{uu} \mathbf{u}_t + \mathbf{z}_t' A_{zz} \mathbf{z}_t \\ + 2\mathbf{u}_t' A_{ux} \mathbf{x}_t + 2\mathbf{u}_t' A_{uz} \mathbf{z}_t + 2\mathbf{x}_t' A_{xz} \mathbf{z}_t. \end{aligned} \tag{2.19}$$

$A_{ij}, i, j \in \{x, u, z\}$ are given matrices. The transition law of the endogenous state variables is

$$\mathbf{x}_{t+1} = B_x \mathbf{x}_t + B_u \mathbf{u}_t + B_z \mathbf{z}_t, \tag{2.20}$$

where $B_x \in \mathbb{R}^{n \times n}$, $B_u \in \mathbb{R}^{n \times m}$, and $B_z \in \mathbb{R}^{n \times l}$ are given matrices. The shocks follow a first-order vector autoregressive process

$$\mathbf{z}_{t+1} = \Pi \mathbf{z}_t + \boldsymbol{\epsilon}_{t+1}, \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \Sigma). \tag{2.21}$$

The eigenvalues of $\Pi \in \mathbb{R}^{l \times l}$ lie inside the unit circle. The planner maximizes

$$E_0 \sum_{t=0}^{\infty} \beta^t g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) \tag{2.22}$$

subject to (2.20) and (2.21).

Let $\boldsymbol{\lambda_t}$ denote the n vector of Lagrange multipliers. The Lagrangian of this LQ problem is

$$\mathscr{L} = E_0 \sum_{t=0}^{\infty} \beta^t \Big[g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) + 2\boldsymbol{\lambda}_t'(B_x \mathbf{x}_t + B_u \mathbf{u}_t + B_z \mathbf{z}_t - \mathbf{x}_{t+1}) \Big].$$

Differentiating this expression with respect to $\mathbf{u}_t$ and $\mathbf{x}_{t+1}$ provides the following set of first-order conditions:

$$\begin{aligned} 0 &= A_{uu}\mathbf{u}_t + A_{ux}\mathbf{x}_t + A_{uz}\mathbf{z}_t + B_u'\boldsymbol{\lambda}_t, \\ \boldsymbol{\lambda}_t &= \beta E_t \left[A_{xx}\mathbf{x}_{t+1} + A_{xz}\mathbf{z}_{t+1} + A_{ux}'\mathbf{u}_{t+1} + B_x'\boldsymbol{\lambda}_{t+1} \right]. \end{aligned}$$

The first of these equations may be rewritten as:

$$C_u \mathbf{u}_t = C_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + C_z \mathbf{z}_t, \tag{2.23a}$$

whereas the second equation and the transition law (2.20) can be summarized in the following matrix difference equation:

$$\begin{aligned} D_{x\lambda} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} + F_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} &= D_u E_t \mathbf{u}_{t+1} + F_u \mathbf{u}_t \\ &\quad + D_z E_t \mathbf{z}_{t+1} + F_z \mathbf{z}_t. \end{aligned} \tag{2.23b}$$

The matrices in these equations relate to those of the original problem as follows:

$$\begin{aligned} C_u &:= A_{uu}, & C_{x\lambda} &:= -\left[A_{ux}, B_u' \right], \\ C_z &:= -A_{uz}, \end{aligned}$$

$$\begin{aligned} D_{x\lambda} &:= \begin{bmatrix} \beta A_{xx} & \beta B_x' \\ I_n & 0_{n\times n} \end{bmatrix}, & F_{x\lambda} &:= \begin{bmatrix} 0_{n\times n} & -I_n \\ -B_x & 0_{n\times n} \end{bmatrix}, \\ D_u &:= \begin{bmatrix} -\beta A_{ux}' \\ 0_{n\times m} \end{bmatrix}, & F_u &:= \begin{bmatrix} 0_{n\times m} \\ B_u \end{bmatrix}, \\ D_z &:= \begin{bmatrix} -\beta A_{xz} \\ 0_{n\times l} \end{bmatrix}, & F_z &:= \begin{bmatrix} 0_{n\times l} \\ B_z \end{bmatrix}, \end{aligned}$$

where I_n and $0_{n\times m}$ denote the n dimensional identity matrix and the $n \times m$ zero matrix, respectively.

Equations (2.23) describe a system of stochastic linear difference equations in two parts. The first part (2.23a) determines the control variables as linear functions of the model's state variables, $\mathbf{x}_t$, exogenous shocks $\mathbf{z}_t$, and the vector of Lagrange multipliers $\boldsymbol{\lambda}_t$, often referred to as the vector of costate variables. The second part (2.23b) determines the dynamics of the vector of state and costate variables. In Section 2.4 equations (2.23) will serve as framework to study the approximate dynamics of non-linear models. Before we explore this subject and discuss the solution of (2.23), we consider the computation of the policy function via value function iterations in the next section.

2.3 LQ Approximation

This section provides the details of an algorithm proposed by Hansen and Prescott (1995). Their approach rests on a linear quadratic approximation of a given model and they device a simple to program iterative procedure to compute the policy function of the approximate model. In Section 2.3.2, we use the deterministic Ramsey model from Example 1.2.1 to illustrate the various steps. Section 2.3.3 outlines the general approach and its implementation in the Gauss program `SolveLQA`.

2.3.1 A Warning

Before we begin, we must warn you. As has been pointed out by Judd (1998), pp. 506-508 and, more recently, by Benigno and Woodford (2007), the method provides a correct linear approximation to the policy function only when the constraints are linear. A different policy function arises from maximizing a quadratic approximation of the objective function subject to linearized constraints. To see this, consider a simple static problem.

Maximize $U(x_1, x_2)$ subject to $x_2 = f(x_1, \epsilon)$, where ϵ is a parameter of the problem. Let $x_1 = h(\epsilon)$ denote the policy function that solves this problem and assume that a solution at $\epsilon = 0$ exists. This solution solves

$$\begin{aligned} g(\epsilon = 0) := U_1 \left[h(\epsilon), f(h(\epsilon), \epsilon)\right] \\ + U_2 \left[h(\epsilon), f(h(\epsilon), \epsilon)\right] f_1(h(\epsilon), \epsilon) = 0. \end{aligned}$$

The implicit function theorem 11.2.3 allows us to compute $h'(0)$ from $g'(0) = 0$. This provides[8]

$$h'(0) = -\frac{U_{12} f_2 + U_{22} f_1 f_2 + U_2 f_{12}}{U_{11} + 2U_{12} f_1 + U_{22} f_1^2 + U_2 f_{11}}. \tag{2.24}$$

The quadratic approximation of U at $x_1^* = h(0)$ and $x_2^* = f(x_1^*, 0)$ is obtained from applying equation (11.32) to U at (x_1^*, x_2^*):

$$U^Q = U(x_1^*, x_2^*) + U_1 \bar{x}_1 + U_2 \bar{x}_2 + \frac{1}{2} \begin{bmatrix} \bar{x}_1, & \bar{x}_2 \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}.$$

Maximizing this expression with respect to $\bar{x}_1 := x_1 - x_1^*$ subject to the linearized constraint

$$\bar{x}_2 = x_2 - x_2^* = f_1 \bar{x}_1 + f_2 \epsilon$$

provides (since $U_1 + U_2 f_1 = 0$)

$$\bar{x}_1 = -\frac{U_{12} f_2 + U_{22} f_1 f_2}{U_{11} + 2U_{12} f_1 + U_{22} f_1^2} \epsilon. \tag{2.25}$$

This solution differs from (2.24) with respect to the rightmost terms in the numerator and the denominator in the solution for $h'(0)$, $U_2 f_{12}$ and $U_2 f_{11}$, respectively. Both terms vanish, if the constraint is linear.

Benigno and Woodford (2007) propose to use the quadratic approximation of the constraint to replace the linear terms in U^Q. Indeed, if we replace $\bar{x}_2$ by

$$\bar{x}_2 = f_1 \bar{x}_1 + f_2 \epsilon + \frac{1}{2} \left[\bar{x}_1, \epsilon\right] \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \epsilon \end{bmatrix}$$

in the expression for U^Q and optimize this new function, we obtain the same linear policy function as given in equation (2.24).

[8] We used $U_{12} = U_{21}$, which holds, if U is twice continuously differentiable. See, e.g., Theorem 1.1 on p. 372 in Lang (1997).

2.3.2 An Illustrative Example

The Model. We know from Section 2.2 that the policy function of the LQ problem is independent of the second moments (and, a fortiori, of any higher moments) of the shocks. Therefore, nothing is lost but much is gained in notational simplicity, if we use the deterministic Ramsey model from example 1.2.1 to illustrate the approach of HANSEN and PRESCOTT (1995). In this example the farmer solves

$$\begin{aligned} \max_{\{C_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t \ln C_t, \quad \beta \in (0,1), \\ \text{s.t. } & K_{t+1} + C_t \leq \; K_t^{\alpha}, \; \alpha \in (0,1), \; t = 0, 1, \ldots, \\ & K_0 \text{ given.} \end{aligned}$$

C_t denotes consumption at time t, and K_t is the stock of capital. The dynamics of this model is determined by two equations:

$$1 = \beta \frac{C_t}{C_{t+1}} \alpha K_{t+1}^{\alpha-1}, \tag{2.26a}$$

$$K_{t+1} = K_t^{\alpha} - C_t. \tag{2.26b}$$

The first equation is a special case of the Euler equation (1.12) in the case of logarithmic preferences and a Cobb-Douglas production function. The second equation is the economy's resource constraint.

Approximation Step. We want to approximate this model by a linear quadratic problem. Towards this end we must look for a linear law of motion and put all remaining nonlinear relations into the current period return function $\ln C_t$. We achieve this by using investment expenditures $I_t = K_t^{\alpha} - C_t$ instead of consumption as a control variable. Remember, this model assumes 100 percent depreciation (i.e., $\delta = 1$), so that the linear transition law is:

$$K_{t+1} = I_t. \tag{2.27}$$

Let $g(K_t, I_t) := \ln(K_t^{\alpha} - I_t)$ denote the current period utility function. We approximate this function by a quadratic function in (K, I) at the point of the stationary solution of the model. This solution derives from equations (2.26) and (2.27) for $K_{t+1} = K_t = \bar{K}$ and $C_{t+1} = C_t = \bar{C}$. Thus,

$$\bar{K} = (\alpha\beta)^{(1/(1-\alpha))}, \tag{2.28a}$$
$$\bar{I} = \bar{K}. \tag{2.28b}$$

A second order Taylor series approximation of g yields:

$$\begin{aligned} g(K, I) \simeq\, & g(\bar{K}, \bar{I}) + g_K(K - \bar{K}) + g_I(I - \bar{I}) \\ & + (1/2)g_{KK}(K - \bar{K})^2 + (1/2)g_{II}(I - \bar{I})^2 \\ & + (1/2)(g_{KI} + g_{IK})(K - \bar{K})(I - \bar{I}). \end{aligned} \tag{2.29}$$

For latter purposes, we want to write the rhs of this equation by using matrix notation.[9] To take care of the constant and the linear terms we define the vector $(1, K, I)^T$ and the 3×3 matrix $Q = (q_{ij})$ and equate the rhs of (2.29) to the product

$$[1, K, I]Q \begin{bmatrix} 1 \\ K \\ I \end{bmatrix}.$$

Comparing terms on both sides of the resulting expression and using the symmetry of the second order mixed partial derivatives $(g_{KI} = g_{IK})$ yields the elements of Q:

$$\begin{aligned} q_{11} &= g - g_K\bar{K} - g_I\bar{I} + (1/2)g_{KK}\bar{K}^2 + g_{KI}\bar{K}\bar{I} + (1/2)g_{II}\bar{I}^2, \\ q_{12} &= q_{21} = (1/2)(g_K - g_{KK}\bar{K} - g_{KI}\bar{I}), \\ q_{13} &= q_{31} = (1/2)(g_I - g_{II}\bar{I} - g_{KI}\bar{K}), \\ q_{23} &= q_{32} = (1/2)g_{KI}, \\ q_{22} &= (1/2)g_{KK}, \\ q_{33} &= (1/2)g_{II}. \end{aligned}$$

In the next step we use Q and the even larger vector $\mathbf{w} = [1, K, I, 1, K']$ (where K' denotes the next-period stock of capital) to write the rhs of the Bellman equation, $g(K, I) + \beta v(K')$, in matrix notation. This gives:

[9] To prevent confusion, we depart from our usual notation temporarily and let the superscript T denote the transpose operator. As usual in dynamic programming, the prime $'$ denotes next-period variables.

$$[1, K, I, 1, K'] \underbrace{\begin{bmatrix} Q & 0_{3\times 2} \\ 0_{2\times 3} & \beta V^0_{2\times 2} \end{bmatrix}}_{R_{5\times 5}} \begin{bmatrix} 1 \\ K \\ I \\ 1 \\ K' \end{bmatrix}, \; V^0 := \begin{bmatrix} v^0_{11} & v^0_{12} \\ v^0_{21} & v^0_{22} \end{bmatrix}. \quad (2.30)$$

We initialize V^0 with a negative definite matrix, e.g., $V^0 = -0.001 I_2$, where I_2 denotes the two-dimensional identity matrix. Our aim is to eliminate all future variables (here it is just K') using the linear law of motion. Then, we perform the maximization step that allows us to eliminate the controls (here it is just I). After that step we have a new guess for the value function, say V^1. We use this guess as input in a new round of iterations until V^0 and V^1 are sufficiently close together.

Reduction Step. We begin to eliminate K' and the constant from (2.30) so that the resulting quadratic form is reduced to a function of the current state K and the current control I. Note that $K' = I$ can be written as dot product:

$$K' = [0, 0, 1, 0] \begin{bmatrix} 1 \\ K \\ I \\ 1 \end{bmatrix},$$

and observe that

$$\begin{bmatrix} 1 \\ K \\ I \\ 1 \\ K' \end{bmatrix} = \begin{bmatrix} & I_4 & & \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ K \\ I \\ 1 \end{bmatrix}.$$

Thus, we may express (2.30) equivalently as:

$$[1, K, I, 1, K'] R_{5\times 5} \begin{bmatrix} 1 \\ K \\ I \\ 1 \\ K' \end{bmatrix} = [1, K, I, 1] R_{4\times 4} \begin{bmatrix} 1 \\ K \\ I \\ 1 \end{bmatrix},$$

where

$$R_{4\times 4} = \begin{bmatrix} & I_4 & \\ 0 & 0 & 1 & 0 \end{bmatrix}^T R_{5\times 5} \underbrace{\begin{bmatrix} & I_4 & \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{S_{5\times 4}}.$$

So what was the trick? In words: use the rightmost variable in $\mathbf{w}^T = [1, K, I, 1, K']$ and write it as linear function of the remaining variables. This gives a row vector with 4 elements. Append this vector to the identity matrix of dimension 4 to get the transformation matrix $S_{5\times 4}$. The matrix of the Bellman equation with K' eliminated is $R_{4\times 4} = S_{5\times 4}^T R_{5\times 5} S_{5\times 4}$.

In the same way we can eliminate the second constant. The constant in terms of the remaining variables $[1, K, I]$ is determined by the dot product:

$$1 = [1, 0, 0] \begin{bmatrix} 1 \\ K \\ I \end{bmatrix}.$$

Thus, the matrix $S_{4\times 3}$ is now

$$S_{4\times 3} = \begin{bmatrix} & I_3 & \\ 1 & 0 & 0 \end{bmatrix},$$

and the rhs of the Bellman equation in terms of $[1, K, I]$ is

$$g(K, I) + \beta v(I) = [1, K, I] R_{3\times 3} \begin{bmatrix} 1 \\ K \\ I \end{bmatrix}, \quad R_{3\times 3} = S_{4\times 3}^T R_{4\times 4} S_{4\times 3}.$$

Maximization Step. In this last step we eliminate I from the rhs of the Bellman equation to find

$$[1, K] R_{2\times 2} \begin{bmatrix} 1 \\ K \end{bmatrix}.$$

The matrix $R_{2\times 2}$ will be our new guess of the value function. After the last reduction step, the quadratic form is:

$$[1, K, I] \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} 1 \\ K \\ I \end{bmatrix}$$
$$= r_{11} + (r_{12} + r_{21})K + (r_{13} + r_{31})I + (r_{23} + r_{32})KI$$
$$+ r_{22}K^2 + r_{33}I^2.$$

Setting the derivative of this expression with respect to I equal to zero and solving for I gives:

$$I = -\underbrace{\frac{r_{13} + r_{31}}{2r_{33}}}_{i_1} - \underbrace{\frac{r_{23} + r_{32}}{2r_{33}}}_{i_2} K = -\frac{r_{13}}{r_{33}} - \frac{r_{23}}{r_{33}}K,$$

where the last equality follows from the symmetry of R. Thus, we can use

$$S = \begin{bmatrix} & I_2 & \\ -i_1 & & -i_2 \end{bmatrix}$$

to reduce $R_{3\times 3}$ to the new guess of the value function:

$$V^1 = S^T R_{3\times 3} S.$$

We stop iterations, if the maximal element in $|V^1 - V^0|$ is smaller than $\epsilon(1-\beta)$ for some small positive ϵ (see (11.84) in Section 11.4 on this choice).

2.3.3 The General Method

Notation. Consider the following framework: There is a n vector of state variables $\mathbf{x}$, a m vector of control variables $\mathbf{u}$, a current period return function $g(\mathbf{x}, \mathbf{u})$, and a discount factor $\beta \in (0, 1)$. As you will see in a moment, it will be helpful to put $x_1 = 1$. All non-linear relations of the model are part of the specification of g, and the remaining linear relations define the following law of motion:

$$\mathbf{x}' = A\mathbf{x} + B\mathbf{u}. \tag{2.31}$$

Furthermore, there is a point $[\mathbf{x}^{*T}, \mathbf{u}^{*T}]^T$. Usually, this will be the stationary solution of the deterministic counterpart of the model under consideration.

Approximation Step. Let $Q \in \mathbb{R}^{l \times l}$, $l = n + m$, denote the matrix of the linear quadratic approximation of the current period return function $g(\cdot)$, and define the $n + m$ column vector $\mathbf{y} = [\mathbf{x}^T, \mathbf{u}^T]^T$. From a Taylor series expansion of g at $\mathbf{y}^*$, we get:

$$\mathbf{y}^T Q \mathbf{y} = g(\mathbf{y}^*) + \sum_{i=1}^{n+m} g_i (y_i - y_i^*) + \frac{1}{2} \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} g_{ij} (y_i - y_i^*)(y_j - y_j^*),$$

where g_i and g_{ij} are first and second partial derivatives of g at $\mathbf{y}^*$, respectively.[10] Comparing terms on both sides of this expression delivers the elements of $Q = (q_{ij})$:

$$\begin{aligned} q_{11} &= g(\mathbf{y}^*) + \textstyle\sum_{i=1}^{n+m} g_i y_i^* + \frac{1}{2} \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} g_{ij} y_i^* y_j^*, \\ q_{1i} = q_{i1} &= \tfrac{1}{2} g_i - \tfrac{1}{2} \textstyle\sum_{j=1}^{n+m} g_{ij} y_j^*, \; i = 2, 3, \dots, n + m, \\ q_{ij} = q_{ji} &= \tfrac{1}{2} g_{ij}, \; i, j = 2, 3, \dots, n + m. \end{aligned}$$

Except in very rare cases, where g_i and g_{ij} are given by simple analytic expressions, one will use numeric differentiation (see Section 11.3.1). For instance, to use our program `SolveLQA`, the user must supply a procedure `gproc` that returns the value of g at an arbitrary point $[\mathbf{x}^T, \mathbf{u}^T]^T$. Note that you must pass $(1, x_2, \dots, x_n, u_1, \dots, u_n)^T$ to that procedure, even if the 1 is not used in `gproc`. This ensures that any procedure that computes the gradient of g returns a vector with l elements and that any procedure that returns the Hesse matrix returns a $l \times l$ matrix. Given this procedure, our Gauss programs `CDJac` and `CDHesse` compute the gradient vector $\nabla g = [0, g_2, g_3, \dots, g_{n+m}]$ and the Hesse matrix $H := (h_{ij}) \equiv (g_{ij})$, $i, j = 1, 2, \dots, n + m$ from which `SolveLQA` builds Q using the above formulas. All of this is done without any further intervention of the user. If higher accuracy in the computation of the Hesse matrix is desired, the user can supply a routine `MyGrad` that returns the gradient vector of g. He must then set the flag `_MyGrad=1` to let the program know that an analytic gradient is available. `SolveLQA` will then use `MyGrad` to compute the Hesse matrix by using the forward difference Jacobian programmed in `CDJac`.

[10] Note, since $x_1 = 1$, we have $g_1 = 0$ and $g_{1i} = g_{i1} = 0$ for $i = 1, 2, \dots, l$.

Reduction Steps. Let R^s denote the matrix that represents the quadratic form on the rhs of the Bellman equation at reduction step s, where

$$R^1 := \begin{bmatrix} Q_{n+m\times(n+m)} & 0_{(n+m)\times n} \\ 0_{n\times(n+m)} & \beta V^0_{n\times n} \end{bmatrix}.$$

In addition, let $\mathbf{c}_s^T$ denote the $n+1-s$-th row of the matrix

$$C_s = \begin{bmatrix} A & B & 0_{n\times(n-s)} \end{bmatrix}.$$

Then, for $s = 1, 2, \dots, n$ iterate on

$$R^{s+1} = \begin{bmatrix} I_{2n+m-s} \\ \mathbf{c}_s^T \end{bmatrix}^T R^s \begin{bmatrix} I_{2n+m-s} \\ \mathbf{c}_s^T \end{bmatrix}.$$

Maximization Steps. After the last reduction step the matrix R is reduced to a square matrix of size $n+m$. There are m maximization steps to be taken until R is reduced further to a square matrix of size n, which is our new guess of the value function. At step $s = 1, 2, \dots, m$ the optimal choice of the control variable u_{m+1-s} as a linear function of the variables $[x_1, \dots, x_n, u_1, \dots, u_{m-s}]$ is given by the row vector

$$\mathbf{d}_s^T = \left[-\frac{r_{1k}}{r_{kk}}, -\frac{r_{2k}}{r_{kk}}, \dots, -\frac{r_{k-1,k}}{r_{kk}} \right],\ k = n+m-s.$$

Therefore, we iterate on

$$R^{s+1} = \begin{bmatrix} I_{n+m-s} \\ \mathbf{d}_s^T \end{bmatrix}^T R^s \begin{bmatrix} I_{n+m-s} \\ \mathbf{d}_s^T \end{bmatrix},\ s = 1, 2, \dots, m.$$

If R is reduced to size n, we have found a new guess of the value function $V^1 = R^{m+1}$, and we compare its elements to those of V^0. If they are close together,

$$\max_{ij} |v_{ij}^0 - v_{ij}^1| < \epsilon(1-\beta),$$

we stop iterations. Otherwise we replace V^0 with V^1 and restart.

Computation of the Policy Function. It is a good idea to store the vectors $\mathbf{d}_s$ in a $m \times (n+m-1)$ matrix D. After convergence, we can use $D = (d_{ij})$ to derive the policy matrix $F \in \mathbb{R}^{m\times n} = (f_{ij})$ that defines the controls as functions of the states. This works as follows: The policy vector $\mathbf{d}_m$ (i.e., the last row of D) holds the coefficients that determine the first control variable u_1 as function of the n state variables:

$$u_1 = \sum_{i=1}^{n} d_{mi} x_i \;\Rightarrow\; f_{1i} = d_{mi}.$$

The second control is given by

$$\begin{aligned} u_2 &= \sum_{i=1}^{n} d_{m-1,i} x_i + d_{m-1,n+1} u_1 \\ &\Rightarrow\; f_{2i} = d_{m-1,i} + d_{m-1,n+1} f_{1i}. \end{aligned}$$

Therefore, we may compute the coefficients of F recursively from:

$$\begin{aligned} f_{ji} &= d_{m+1-j,i} + \sum_{k=1}^{j-1} d_{m+1-j,n+k} f_{ki}, \\ j &= 1, \dots, m,\; i = 1, \dots, n. \end{aligned}$$

As a final check of the solution, we can use

$$|\mathbf{u}^* - F\mathbf{x}^*|.$$

i.e. the discrepancy between the stationary solution of the controls from the original model and those computed using the linear policy function.

2.4 Linear Approximation

In this section we return to the system of stochastic difference equations (2.23). Remember, this system is one way to characterize the solution of the linear quadratic problem. However, we are by no means restricted to this interpretation. More generally, we may consider this system as an approximation of an arbitrary non-linear model. In the next subsection we explain this approximation by means of the stochastic growth model. Our discussion closely parallels the presentation in Section 2.1. First, we demonstrate that both, the solution to a linearized system of stochastic difference equations and the application of the implicit function theorem provide the same set of equations for the coefficients of the policy function. Second, we obtain these coefficients from the solution of a linear system of stochastic difference equations. Section 2.4.2 presents the solution method for the general case of equations (2.23) and explains the use of our program `SolveLA` that implements this method.

2.4.1 An Illustrative Example

There are two equations that determine the time path of the stochastic Ramsey model from Section 1.3 with strictly positive consumption. They are obtained from equations (1.23):

$$0 = K_{t+1} - (1-\delta)K_t - Z_t f(K_t) + C_t, \tag{2.32a}$$
$$0 = u'(C_t) - \beta E_t u'(C_{t+1})(1-\delta + Z_{t+1} f'(K_{t+1})). \tag{2.32b}$$

We assume that the productivity shock Z_t follows the process

$$\ln Z_t = \varrho \ln Z_{t-1} + \sigma\epsilon_t, \quad \epsilon_t \sim N(0,1). \tag{2.32c}$$

Since $\ln Z_t \simeq \bar{Z}_t, \bar{Z}_t = Z_t - Z^*, Z^* \equiv 1$ this equation may be approximated by

$$\bar{Z}_t = \varrho \bar{Z}_{t-1} + \sigma\epsilon_t. \tag{2.32d}$$

Note, that for $\sigma = 0$ and $Z^* = 1$ this model reduces to the deterministic growth model with the stationary equilibrium determined from

$$C^* = f(K^*) - \delta K^*, \tag{2.33a}$$
$$1 = \beta(1-\delta + f'(K^*)). \tag{2.33b}$$

More generally, equations (2.32) may be written as $E_t \mathbf{g}(\mathbf{x}_t, \mathbf{x}_{t+1}) = \mathbf{0}_{2\times 1}$, $\mathbf{x}_t := [K_t, C_t, Z_t]'$.

Linear Stochastic Difference Equations. At (K^*, C^*, Z^*) the linearized version of this system of equations is given by:

$$\begin{aligned}\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}g_1^1 & g_2^1\\ g_1^2 & g_2^2\end{bmatrix}\begin{bmatrix}\bar{K}_t\\ \bar{C}_t\end{bmatrix} + \begin{bmatrix}g_4^1 & g_5^1\\ g_4^2 & g_5^2\end{bmatrix} E_t \begin{bmatrix}\bar{K}_{t+1}\\ \bar{C}_{t+1}\end{bmatrix}\\ + \begin{bmatrix}g_3^1\\ g_3^2\end{bmatrix}\bar{Z}_t + \begin{bmatrix}g_6^1\\ g_6^2\end{bmatrix} E_t\bar{Z}_{t+1},\end{aligned} \tag{2.34}$$

where $\bar{x}_t$ denotes $x_t - x^*$. Since equation (2.32d) implies $E_t\bar{Z}_{t+1} = \varrho\bar{Z}_t$ the last term in equation (2.34) may also be written as $\varrho[g_6^1, g_6^2]'\bar{Z}_t$. We assume that the linear policy functions for $\bar{K}_{t+1}$ and $\bar{C}_t$ are of the form

$$\bar{K}_{t+1} = h_K^K\bar{K}_t + h_Z^K\bar{Z}_t, \tag{2.35a}$$

$$\bar{C}_t = h_K^C\bar{K}_t + h_Z^C\bar{Z}_t, \tag{2.35b}$$

where h_j^i, $i, j \in \{K, C\}$ denotes the derivative of the policy function of variable i with respect to its jth argument. Substituting this guess in equation (2.34) yields

$$\begin{bmatrix}a_1\\a_2\end{bmatrix}\bar{K}_t + \begin{bmatrix}b_1\\b_2\end{bmatrix}\bar{Z}_t = \begin{bmatrix}0\\0\end{bmatrix},$$

where a_i and b_i, $i = 1, 2$ are collections of coefficients to be given in a moment. Obviously, if (2.35) is a solution to (2.34), this requires $a_i = b_i = 0$, $i = 1, 2$ and, thus, provides four (non-linear) equations in the unknown coefficients h_K^K, h_Z^K, h_K^C, h_Z^C. A modest amount of algebra reveals these relations:

$$a_1 = g_1^1 + g_2^1 h_K^C + (g_4^1 + g_5^1 h_K^C)h_K^K = 0, \tag{2.36a}$$

$$a_2 = g_1^2 + g_2^2 h_K^C + (g_4^2 + g_5^2 h_K^C)h_K^K = 0, \tag{2.36b}$$

$$b_1 = (g_3^1 + g_6^1\varrho) + (g_2^1 + g_5^1\varrho)h_Z^C + (g_4^1 + g_5^1 h_K^C)h_Z^K = 0, \tag{2.36c}$$

$$b_2 = (g_3^2 + g_6^2\varrho) + (g_2^2 + g_5^2\varrho)h_Z^C + (g_4^2 + g_5^2 h_K^C)h_Z^K = 0. \tag{2.36d}$$

Application of the Implicit Function Theorem. We will now demonstrate that the same set of conditions emerges, if we apply the implicit function theorem to the system $E_t\mathbf{g}(\mathbf{x}_t, \mathbf{x}_{t+1}) = \mathbf{0}_{2\times 1}$.

This allows us also to show that the linear policy functions are indeed independent of the parameter σ. We assume non-linear policy functions $K_{t+1} = h^K(K_t, Z_t, \sigma)$ and $C_t = h^C(K_t, Z_t, \sigma)$ with the property $K^* = h^K(K^*, Z^*, 0)$, $C^* = h^C(K^*, Z^*, 0)$ so that a solution of $\mathbf{g}(\cdot) = \mathbf{0}_{2\times 1}$ at $(K^*, Z^*, 0)$ exists. It is not difficult to see that differentiating $\mathbf{g}$ with respect to K_t and Z_t provides the same conditions on the derivatives of h^C and h^K at the stationary solution as presented in equations (2.36). Just note, that g^i, $i = 1, 2$ can be written as

$$\begin{aligned} g^i\Big(&K_t, h^C(K_t, Z_t, \sigma), Z_t, h^K(K_t, Z_t, \sigma), \\ &h^C\big(h^K(K_t, Z_t, \sigma), e^{\varrho \ln Z_t + \sigma \epsilon_{t+1}}, \sigma\big), e^{\varrho \ln Z_t + \sigma \epsilon_{t+1}}\Big), \end{aligned}$$

so that, for instance,

$$\frac{\partial g^1(\cdot)}{\partial K_t} = g_1^1 + g_2^1 h_K^C + g_4^1 h_K^K + g_5^1 h_K^C h_K^K \equiv a_1.$$

Consider the derivatives with respect to σ. They imply:[11]

$$\begin{bmatrix} (g_4^1 + g_5^1 h_K^C) & (g_2^1 + g_5^1) \\ (g_4^2 + g_5^2 h_K^C) & (g_2^2 + g_5^2) \end{bmatrix} \begin{bmatrix} h_\sigma^K \\ h_\sigma^C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This is a system of homogenous equations in h_σ^K and h_σ^C. Since its matrix of coefficients is regular, the only possible solution is $h_\sigma^K = h_\sigma^C = 0$.

We have, thus, seen by means of an example that the application of perturbation methods to a stochastic DGE model allows us to derive linear approximations of the policy functions via the solution of the linearized system of stochastic difference equations.[12]

---

[11] The derivative of the term $Z_{t+1} = e^{\varrho \ln Z_t + \sigma \epsilon_{t+1}}$ with respect to σ evaluated at $Z^* = 1$ and $\sigma = 0$ is ϵ_{t+1}. The expectation of this term as of time t, $E_t \epsilon_{t+1}$, equals zero, the mean of $N(0,1)$.

[12] The generalization of this result is obvious but involves either intricate formulas or the use of tensor notation so that we have decided not to pursue it here. See Schmitt-Grohé and Uribe (2004) for a proof.

Derivation of the Solution. Rather than solving (2.36), we determine the coefficients of the policy functions via the same procedure that we used in Section 2.1. From (2.32) and (2.34) we obtain the following system of linear, stochastic difference equations:

$$E_t \begin{bmatrix} \bar{K}_{t+1} \\ \bar{C}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{\beta} & -1 \\ -\frac{u'f''}{u''} & 1 + \frac{\beta u' f''}{u''} \end{bmatrix}}_{=:W} \begin{bmatrix} \bar{K}_t \\ \bar{C}_t \end{bmatrix} + \underbrace{\begin{bmatrix} f \\ -\frac{\beta u' f f'' + \varrho u'}{u''} \end{bmatrix}}_{=:R} \bar{Z}_t. \tag{2.37}$$

The matrix W equals the Jacobian matrix of the deterministic system (2.3), and, thus, has eigenvalues $\lambda_1 < 1$ and $\lambda_2 > 1$. In the new variables[13]

$$\begin{bmatrix} \tilde{K}_t \\ \tilde{C}_t \end{bmatrix} := T^{-1} \begin{bmatrix} \bar{K}_t \\ \bar{C}_t \end{bmatrix} \Leftrightarrow T \begin{bmatrix} \tilde{K}_t \\ \tilde{C}_t \end{bmatrix} := \begin{bmatrix} \bar{K}_t \\ \bar{C}_t \end{bmatrix} \tag{2.38}$$

the system of difference equations may be written as:[14]

$$E_t \begin{bmatrix} \tilde{K}_{t+1} \\ \tilde{C}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda_1 & s_{12} \\ 0 & \lambda_2 \end{bmatrix}}_{S} \begin{bmatrix} \tilde{K}_t \\ \tilde{C}_t \end{bmatrix} + \underbrace{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}}_{Q=T^{-1}R} \hat{Z}_t. \tag{2.39}$$

Consider the second equation of this system, which is a relation in the new variable $\tilde{C}_t$ and the exogenous shock:

$$E_t \tilde{C}_{t+1} = \lambda_2 \tilde{C}_t + q_2 \bar{Z}_t. \tag{2.40}$$

We can solve this equation for $\tilde{C}_t$ via repeated substitution: from (2.40) we get

$$\tilde{C}_t = \frac{1}{\lambda_2} E_t \tilde{C}_{t+1} - \frac{q_2}{\lambda_2} \bar{Z}_t. \tag{2.41}$$

Shifting the time index one period into the future yields:

---

[13] Remember, T is the matrix that puts W into Schur form $W = TST^{-1}$.

[14] Pre-multiply (2.37) by T^{-1} and use the definitions in (2.38) to arrive at this representation.

$$\tilde{C}_{t+1} = \frac{1}{\lambda_2} E_{t+1}\tilde{C}_{t+2} - \frac{q_2}{\lambda_2}\bar{Z}_{t+1}.$$

Taking expectations as of period t on both sides and noting that (via the law of iterated expectations) $E_t(E_{t+1}\tilde{C}_{t+2}) = E_t\tilde{C}_{t+2}$ yields:

$$E_t\tilde{C}_{t+1} = \frac{1}{\lambda_2} E_t\tilde{C}_{t+2} - \frac{q_2}{\lambda_2} E_t\bar{Z}_{t+1} = \frac{1}{\lambda_2} E_t\tilde{C}_{t+2} - \frac{q_2}{\lambda_2}\varrho\bar{Z}_t, \quad (2.42)$$

due to (2.32d). Substitution of this solution for $E_t\tilde{C}_{t+1}$ into (2.41) results in:

$$\tilde{C}_t = \frac{1}{\lambda_2^2} E_t\tilde{C}_{t+2} - \left[\frac{q_2}{\lambda_2} + \frac{q_2}{\lambda_2}\frac{\varrho}{\lambda_2}\right]\bar{Z}_t.$$

We can use (2.42) to get an expression for $\tilde{C}_{t+3}$ and so on up to period $t+\tau$:

$$\tilde{C}_t = \left[\frac{1}{\lambda_2}\right]^{\tau} E_t\tilde{C}_{t+\tau} - \frac{q_2}{\lambda_2}\sum_{i=0}^{\tau-1}\left[\frac{\varrho}{\lambda_2}\right]^{i}\bar{Z}_t. \quad (2.43)$$

Suppose that the sequence

$$\left\{\frac{1}{\lambda_2^{\tau}} E_t\tilde{C}_{t+\tau}\right\}_{\tau=0}^{\infty}$$

converges towards zero for $\tau \to \infty$. This is not very restrictive: since $1/\lambda_2 < 1$, it is sufficient to assume that $E_t\tilde{C}_{t+\tau}$ is bounded. Intuitively, this assumption rules out speculative bubbles along explosive paths and renders the solution unique. In addition, it guarantees that the transversality condition (1.25) is met. In this case we can compute the limit of (2.43) for $\tau \to \infty$:

$$\tilde{C}_t = -\frac{q_2/\lambda_2}{1-(\varrho/\lambda_2)}\bar{Z}_t. \quad (2.44)$$

We substitute this solution into the second equation of (2.38),[15]

[15] We denote the elements of T^{-1} by (t^{ij}).

$$\tilde{C}_t = t^{21}\bar{K}_t + t^{22}\bar{C}_t,$$

to get the solution for $\bar{C}_t$ in terms of $\bar{K}_t$ and $\bar{Z}_t$:

$$\bar{C}_t = -\underbrace{\frac{t^{21}}{t^{22}}}_{=:h_K^C}\bar{K}_t - \underbrace{\frac{q_2/\lambda_2}{t^{22}(1-(\varrho/\lambda_2))}}_{=:h_Z^C}\bar{Z}_t. \tag{2.45}$$

From the first equation of (2.37),

$$\bar{K}_{t+1} = \frac{1}{\beta}\bar{K}_t - \bar{C}_t + f\bar{Z}_t,$$

we can derive the solution for $\bar{K}_{t+1}$:

$$\begin{aligned}\bar{K}_{t+1} &= \frac{1}{\beta}\bar{K}_t - \underbrace{(h_K^C\bar{K}_t + h_Z^C\bar{Z}_t)}_{=\bar{C}_t} + f\bar{Z}_t\\ \bar{K}_{t+1} &= \underbrace{\left(\frac{1}{\beta} - h_K^C\right)}_{=:h_K^K}\bar{K}_t + \underbrace{(f - h_Z^C)}_{=:h_Z^K}\bar{Z}_t.\end{aligned}$$

Thus, given a sequence of shocks $\{\epsilon_t\}_{t=0}^T$ and an initial $\bar{K}_0$ we may compute the entire time path of consumption and the stock of capital by iteration over

$$\bar{C}_t = h_K^C\bar{K}_t + h_Z^C\bar{Z}_t, \tag{2.46a}$$
$$\bar{K}_{t+1} = h_K^K\bar{K}_t + h_Z^K\bar{Z}_t, \tag{2.46b}$$
$$\bar{Z}_{t+1} = \varrho\bar{Z}_t + \epsilon_{t+1}. \tag{2.46c}$$

The Gauss program `Ramsey3a.g` computes the linear approximations of the policy function of the stochastic growth model from Section 1.3 along the lines described above. The utility function is parameterized as $u(C) = [C^{1-\eta} - 1]/(1-\eta)$ and the production function as $f(K) = K^\alpha$. The program shows how to derive the coefficients of the matrices in equation (2.34) by using numeric differentiation. In the case with logarithmic preferences, complete depreciation δ=1, $\alpha = 0.27$, $\beta = 0.994$ $\varrho = 0.90$, and $\sigma = 0.0072$ the program delivers the following policy functions:

$$\bar{C}_t = 0.736\bar{K}_t + 0.450\bar{Z}_t,$$
$$\bar{K}_{t+1} = 0.270\bar{K}_t + 0.165\bar{Z}_t.$$

In this case, the exact analytic solution is

$$C_t = 0.268 Z_t K_t^{0.27},$$
$$K_{t+1} = 0.732 Z_t K_t^{0.27}$$

Figure 2.3 shows the histograms of the distribution for the capital stock that result from the simulation of both solutions. The simulations use the same sequence of shocks to prevent random differences in the results. By and large, the linear model implies the same stationary distribution of the capital stock as does the true, non-linear model.

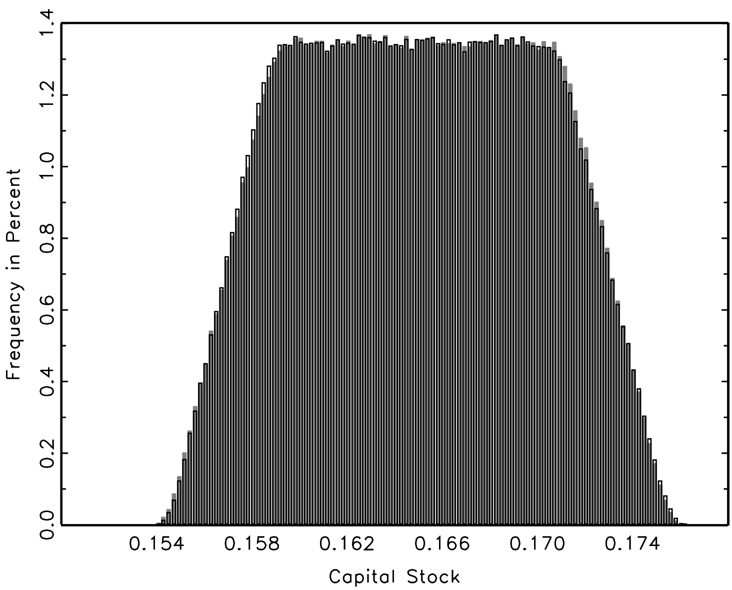

Figure 2.3: Stationary Distribution of the Capital Stock from the Analytic and the Linear Approximate Solution of the Stochastic Infinite-Horizon Ramsey Model

In most applications we want a unit free measure of deviations around the deterministic steady state. Given the linear approximations from above, this is easy to obtain: Just divide both sides of the policy function by the stationary value of the respective lhs variable and rearrange. For instance, using (2.46a), we may write:

$$\hat{C}_t := \frac{C_t - C^*}{C^*} = h_K^C \frac{K^*}{C^*} \underbrace{\frac{K_t - K^*}{K^*}}_{=:\hat{K}_t} + h_Z^C \frac{Z^*}{C^*} \underbrace{\frac{Z_t - Z^*}{Z^*}}_{=:\hat{Z}_t}.$$

Since $\ln(X_t/X^*) \simeq (X_t - X^*)/X^*$, this is a log-linear approximation of the policy function for consumption that relates the percentage deviation of consumption to the percentage deviations of the stock of capital and the productivity shock, respectively.

In the next subsection we basically use the same steps to derive the policy functions for the general system (2.23). If you dislike linear algebra, you may skip this section and note that the program `SolveLA` performs the above explained computations for the general case. The program requires the matrices from (2.23) as input and returns matrices L_j^i that relate the vectors $\mathbf{u}_t$, $\boldsymbol{\lambda}_t$ and $\mathbf{x}_{t+1}$ to the model's state variables in the vectors $\mathbf{x}_t$ and $\mathbf{z}_t$.

2.4.2 The General Method

In this subsection we consider the solution of a system of linear stochastic difference equations given in the form of (2.23), which derives from the LQ problem. There are related ways to state and solve such systems. The list of references includes the classical paper by Blanchard and Kahn (1980), Chapter 3 of the book by Farmer (1993), the papers of King and Watson (1998), (2002), Klein (2000) and the approach proposed by Uhlig (1999). Our statement of the problem is the one proposed by Burnside (1999), but we solve it along the lines of King and Watson (2002).

The Problem. Consider the system of stochastic difference equations (2.47):

$$C_u \mathbf{u}_t = C_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + C_z \mathbf{z}_t, \tag{2.47a}$$

$$\begin{aligned} D_{x\lambda} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} + F_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} &= D_u E_t \mathbf{u}_{t+1} + F_u \mathbf{u}_t \\ &\quad + D_z E_t \mathbf{z}_{t+1} + F_z \mathbf{z}_t. \end{aligned} \tag{2.47b}$$

To ease notation we use $n(x)$ to denote the dimension (i.e., the number of elements) of the vector $\mathbf{x}$. We think of the $n(u)$ vector $\mathbf{u}_t$ as the collection of variables that are determined within period t as linear functions of the model's state variables. We distinguish between three kinds of state variables: those with given initial conditions build the $n(x)$ vector $\mathbf{x}_t$; the $n(\lambda)$ vector $\boldsymbol{\lambda}_t$ collects those variables, whose initial values may be chosen freely. In the LQ problem these are the costate variables. In the stochastic growth model it is just the Lagrange multiplier of the budget constraint. Purely exogenous stochastic shocks are stacked in the $n(z)$ vector $\mathbf{z}_t$. We assume that $\mathbf{z}_t$ is governed by a stable vector autoregressive process of first-order with normally distributed innovations $\boldsymbol{\epsilon}_t$:

$$\mathbf{z}_t = \Pi \mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \Sigma). \tag{2.48}$$

Stability requires that the eigenvalues of the matrix Π lie within the unit circle.

System Reduction. We assume that the first equation can be solved for the vector $\mathbf{u}_t$:

$$\mathbf{u}_t = C_u^{-1} C_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + C_u^{-1} C_z \mathbf{z}_t. \tag{2.49}$$

Shifting the time index one period into the future and taking expectations conditional on information as of period t yields:

$$E_t \mathbf{u}_{t+1} = C_u^{-1} C_{x\lambda} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} + C_u^{-1} C_z E_t \mathbf{z}_{t+1}. \tag{2.50}$$

The solutions (2.49) and (2.50) allow us to eliminate $\mathbf{u}_t$ and $E_t \mathbf{u}_{t+1}$ from (2.47b):

$$\left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right) E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} = -\left(F_{x\lambda} - F_u C_u^{-1} C_{x\lambda}\right) \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + \left(D_z + D_u C_u^{-1} C_z\right) E_t \mathbf{z}_{t+1} + \left(F_z + F_u C_u^{-1} C_z\right) \mathbf{z}_t.$$

Assume that this system can be solved for $E_t(\mathbf{x}_{t+1}, \boldsymbol{\lambda}_{t+1})'$. In other words, the matrix $D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}$ must be invertible. Using $E_t \mathbf{z}_{t+1} = \Pi \mathbf{z}_t$, which is implied by (2.48), we get the following reduced dynamic system:

$$\begin{aligned} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} &= W \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + R \mathbf{z}_t, \\ W &= -\left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right)^{-1} \left(F_{x\lambda} - F_u C_u^{-1} C_{x\lambda}\right), \\ R &= \left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right)^{-1} \\ &\quad \times \left[\left(D_z + D_u C_u^{-1} C_z\right) \Pi + \left(F_z + F_u C_u^{-1} C_z\right)\right]. \end{aligned} \tag{2.51}$$

Change of Variables. Consider the Schur factorization of the matrix W:

$$S = T^{-1} W T,$$

which gives raise to the following partitioned matrices:

$$\begin{aligned} S &= \begin{bmatrix} S_{xx} & S_{x\lambda} \\ 0 & S_{\lambda\lambda} \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix}}_{T^{-1}} \underbrace{\begin{bmatrix} W_{xx} & W_{x\lambda} \\ W_{\lambda x} & W_{\lambda\lambda} \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} T_{xx} & T_{x\lambda} \\ T_{\lambda x} & T_{\lambda\lambda} \end{bmatrix}}_{T}. \end{aligned} \tag{2.52}$$

We assume that the eigenvalues of W appear in ascending order on the main diagonal of S (see 11.1). To find a unique solution, $n(x)$ eigenvalues must lie inside the unit circle and $n(\lambda)$ eigenvalues must have modulus greater than one. In the new variables

$$\begin{bmatrix} \tilde{\mathbf{x}}_t \\ \tilde{\boldsymbol{\lambda}}_t \end{bmatrix} := \begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} \tag{2.53}$$

the dynamic system (2.51) can be rewritten as

$$E_t \begin{bmatrix} \tilde{\mathbf{x}}_{t+1} \\ \tilde{\boldsymbol{\lambda}}_{t+1} \end{bmatrix} = \begin{bmatrix} S_{xx} & S_{x\lambda} \\ 0 & S_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_t \\ \tilde{\boldsymbol{\lambda}}_t \end{bmatrix} + \begin{bmatrix} Q_x \\ Q_\lambda \end{bmatrix} \mathbf{z}_t, \tag{2.54}$$
$$Q = [Q_x, Q_\lambda]' = T^{-1} R.$$

Policy Function for $\boldsymbol{\lambda}_t$. Consider the second line of (2.54), which is a linear system in $\tilde{\boldsymbol{\lambda}}$ alone:

$$E_t \tilde{\boldsymbol{\lambda}}_{t+1} = S_{\lambda\lambda} \tilde{\boldsymbol{\lambda}}_t + Q_\lambda \mathbf{z}_t. \tag{2.55}$$

Its solution is given by:

$$\tilde{\boldsymbol{\lambda}}_t = \Phi \mathbf{z}_t. \tag{2.56}$$

There is a quick and a more illuminating way to compute the matrix Φ. Here is the quick one: Substitute (2.56) into equation (2.55) to obtain

$$E_t \Phi \mathbf{z}_{t+1} = \Phi \Pi \mathbf{z}_t = S_{\lambda\lambda} \Phi \mathbf{z}_t + Q_\lambda \mathbf{z}_t.$$

Thus, Φ must solve the matrix equation

$$\Phi \Pi = S_{\lambda\lambda} \Phi + Q_\lambda.$$

Applying the vec operator to this equations yields (see the rule (11.10b))

$$\operatorname{vec} \Phi = \left[\Pi' \otimes I_{n(\lambda)} - I_{n(z)} \otimes S_{\lambda\lambda} \right]^{-1} \operatorname{vec} Q_\lambda.$$

One may also compute the rows of the matrix Φ in the following steps: The matrix $S_{\lambda\lambda}$ is upper triangular with all of its eigenvalues μ_i on the main diagonal being larger than one in absolute value:

$$S_{\lambda\lambda} = \begin{bmatrix} \mu_1 & s_{12} & \dots & s_{1n(\lambda)} \\ 0 & \mu_2 & \dots & s_{2n(\lambda)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_{n(\lambda)} \end{bmatrix}.$$

Therefore, the last line of (2.55) is a stochastic difference equation in the single variable $\tilde{\lambda}_{n(\lambda)}$, just like equation (2.40):

$$E_t\tilde{\lambda}_{n(\lambda)\,t+1} = \mu_{n(\lambda)}\tilde{\lambda}_{n(\lambda)\,t} + \mathbf{q}'_{n(\lambda)}\mathbf{z}_t, \tag{2.57}$$

where $\mathbf{q}'_{n(\lambda)}$ denotes the last row of the matrix Q_λ. Note, that $\tilde{\lambda}_{n(\lambda)\,t}$ – as every other component of $\tilde{\boldsymbol{\lambda}}_t$ – may be a complex variable. Yet, since the modulus (i.e., the absolute value) of the complex number $\mu_{n(\lambda)}$ is larger than one, the sequence

$$\left\{\frac{1}{\mu^\tau_{n(\lambda)}}E_t\tilde{\lambda}_{n(\lambda)\,t+\tau}\right\}_{\tau=0}^{\infty}$$

will converge to zero if the sequence

$$\left\{E_t\tilde{\lambda}_{n(\lambda)\,t+\tau}\right\}_{\tau=0}^{\infty}$$

is bounded (see Section 12.1). Given this assumption, we know from equation (2.44) that the solution to (2.57) is a linear function of $\mathbf{z}_t$:

$$\tilde{\lambda}_{n(\lambda)\,t} = \underbrace{(\phi_{n(\lambda)\,1}, \phi_{n(\lambda)\,2}, \dots, \phi_{n(\lambda),n(z)})'}_{\boldsymbol{\phi}'_{n(\lambda)}}\,\mathbf{z}_t.$$

To determine the yet unknown coefficients of this function, i.e., the elements of the row vector $\boldsymbol{\phi}'_{n(\lambda)}$, we proceed as follows: we substitute this solution into equation (2.57). This yields:

$$\begin{aligned}
\boldsymbol{\phi}'_{n(\lambda)}E_t\mathbf{z}_{t+1} &= \mu_{n(\lambda)}\boldsymbol{\phi}'_{n(\lambda)}\mathbf{z}_t + \mathbf{q}'_{n(\lambda)}\mathbf{z}_t,\\
\left(\boldsymbol{\phi}'_{n(\lambda)}\Pi - \boldsymbol{\phi}'_{n(\lambda)}\mu_{n(\lambda)}\right)\mathbf{z}_t &= \mathbf{q}'_{n(\lambda)}\mathbf{z}_t,\\
\boldsymbol{\phi}'_{n(\lambda)}\left(\Pi - \mu_{n(\lambda)}I_{n(z)}\right)\mathbf{z}_t &= \mathbf{q}'_{n(\lambda)}\mathbf{z}_t,
\end{aligned}$$

where the second line follows from (2.48). Equating the coefficients on both sides of the last line of the preceding expression gives the solution for the unknown vector $\boldsymbol{\phi}_{n(\lambda)}$:

$$\boldsymbol{\phi}'_{n(\lambda)} = \mathbf{q}'_{n(\lambda)}\left(\Pi - \mu_{n(\lambda)}I_{n(z)}\right)^{-1}. \tag{2.58}$$

Since the eigenvalues of Π are inside the unit circle, this solution exists.

Now, consider the next to last line of (2.55):

$$\begin{aligned} E_t\tilde{\lambda}_{n(\lambda)-1\,t+1} &= \mu_{n(\lambda)-1}\tilde{\lambda}_{n(\lambda)-1\,t} + s_{n(\lambda)-1,n(\lambda)}\tilde{\lambda}_{n(\lambda)\,t} + \mathbf{q}'_{n(\lambda)-1}\mathbf{z}_t, \\ E_t\tilde{\lambda}_{n(\lambda)-1\,t+1} &= \mu_{n(\lambda)-1}\tilde{\lambda}_{n(\lambda)-1\,t} + s_{n(\lambda)-1,n(\lambda)}\boldsymbol{\phi}'_{n(\lambda)}\mathbf{z}_t + \mathbf{q}'_{n(\lambda)-1}\mathbf{z}_t. \end{aligned}$$

The solution to this equation is given by the row vector $\boldsymbol{\phi}'_{n(\lambda)-1}$. Repeating the steps from above, we find:

$$\boldsymbol{\phi}'_{n(\lambda)-1} = \left(\mathbf{q}'_{n(\lambda)-1} + s_{n(\lambda)-1\,n(\lambda)}\boldsymbol{\phi}'_{n(\lambda)}\right)\left(\Pi - \mu_{n(\lambda)-1}I_{n(z)}\right)^{-1}. \tag{2.59}$$

Proceeding from line $n(\lambda)-1$ to line $n(\lambda)-2$ and so forth until the first line of (2.55) we are able to compute all of the rows $\boldsymbol{\phi}'_i$ of the matrix Φ. The respective formula is:

$$\begin{aligned} \boldsymbol{\phi}'_i &= \left[\mathbf{q}'_i + \sum_{j=i+1}^{n(\lambda)} s_{i,j}\boldsymbol{\phi}'_j\right]\left(\Pi - \mu_i I_{n(z)}\right)^{-1}, \\ &\quad i = n(\lambda), n(\lambda)-1, \dots, 1. \end{aligned} \tag{2.60}$$

Given the solution for $\tilde{\boldsymbol{\lambda}}_t$ we can use (2.53) to find the solution for $\boldsymbol{\lambda}_t$ in terms of $\mathbf{x}_t$ and $\mathbf{z}_t$. The second part of (2.53) is:

$$\tilde{\boldsymbol{\lambda}}_t = T^{\lambda x}\mathbf{x}_t + T^{\lambda\lambda}\boldsymbol{\lambda}_t.$$

Together with (2.56) this gives:

$$\boldsymbol{\lambda}_t = \underbrace{-\left(T^{\lambda\lambda}\right)^{-1}T^{\lambda x}}_{L_x^\lambda}\mathbf{x}_t + \underbrace{\left(T^{\lambda\lambda}\right)^{-1}\Phi}_{L_z^\lambda}\mathbf{z}_t. \tag{2.61}$$

Policy Function for $\mathbf{x}_{t+1}$. In obvious notation the first part of (2.51) may be written as:

$$\mathbf{x}_{t+1} = W_{xx}\mathbf{x}_t + W_{x\lambda}\boldsymbol{\lambda}_t + R_x\mathbf{z}_t.$$

Substitution for $\boldsymbol{\lambda}_t$ from (2.61) gives:

$$
\begin{aligned}
\mathbf{x}_{t+1} = & \underbrace{\left(W_{xx} - W_{x\lambda}\left(T^{\lambda\lambda}\right)^{-1} T^{\lambda x}\right)}_{L_x^x} \mathbf{x}_t \\
& + \underbrace{\left(W_{x\lambda}\left(T^{\lambda\lambda}\right)^{-1} \Phi + R_x\right)}_{L_z^x} \mathbf{z}_t .
\end{aligned}
\tag{2.62}
$$

The expression for L_x^x may be considerably simplified. In terms of partitioned matrices the expression $W = TST^{-1}$ may be written as:

$$
\begin{bmatrix} W_{xx} & W_{x\lambda} \\ W_{\lambda x} & W_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} T_{xx} & T_{x\lambda} \\ T_{\lambda x} & T_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} S_{xx} & S_{x\lambda} \\ 0 & S_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix},
$$

which implies:

$$
\begin{aligned}
W_{xx} &= T_{xx}S_{xx}T^{xx} + T_{xx}S_{x\lambda}T^{\lambda x} + T_{x\lambda}S_{\lambda\lambda}T^{\lambda x}, \\
W_{x\lambda} &= T_{xx}S_{xx}T^{x\lambda} + T_{xx}S_{x\lambda}T^{\lambda\lambda} + T_{x\lambda}S_{\lambda\lambda}T^{\lambda\lambda}.
\end{aligned}
$$

Substituting the rhs of these equations into the expression for L_{xx} from (2.62) gives:

$$
L_x^x = T_{xx}S_{xx}\left(T^{xx} - T^{x\lambda}\left(T^{\lambda\lambda}\right)^{-1} T^{\lambda x}\right).
$$

Since

$$
\begin{bmatrix} T_{xx} & T_{x\lambda} \\ T_{\lambda x} & T_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix}^{-1}
$$

the formula for the inverse of a partitioned matrix (11.15a) implies:

$$
\left(T_{xx}\right)^{-1} = T^{xx} - T^{x\lambda}\left(T^{\lambda\lambda}\right)^{-1} T^{\lambda x}.
$$

Putting all pieces together, we find:

$$
L_x^x = T_{xx}S_{xx}T_{xx}^{-1}.
$$

Policy Function for $\mathbf{u}_t$. Using equation (2.49) the solutions for $\mathbf{x}_t$ and $\boldsymbol{\lambda}_t$ imply the following policy function for the vector $\mathbf{u}_t$:

$$\begin{aligned}\mathbf{u}_t &= \underbrace{C_u^{-1}C_{x\lambda}\begin{bmatrix} I_{n(x)} \\ L_x^\lambda \end{bmatrix}}_{L_x^u}\mathbf{x}_t \\ &+ \underbrace{\left(C_u^{-1}C_{x\lambda}\begin{bmatrix} 0_{n(x)\times n(z)} \\ L_z^\lambda \end{bmatrix} + C_u^{-1}C_z\right)}_{L_z^u}\mathbf{z}_t.\end{aligned} \tag{2.63}$$

Implementation. Our Gauss program `SolveLA` performs the computation of the policy matrices according to the formulas given by equations (2.61), (2.62), and (2.63). It uses the Gauss intrinsic command `Schtoc` to get the matrices S and T. However, the eigenvalues on the main diagonal of S are not ordered. We use the Givens rotation described in Section 11.1 to sort the eigenvalues in ascending order. The program's input are the matrices from (2.47), the matrix Π from (2.48), and the number of elements $n(x)$ of the vector $\mathbf{x}_t$. The program checks whether $n(x)$ of the eigenvalues of W are inside the unit circle. If not, it stops with an error message. Otherwise it returns the matrices L_x^x, L_z^x, L_x^λ, L_z^λ, L_x^u, and L_z^u. A second version of this program, `SolveLA2`, uses the Gauss foreign language interface and calls a routine (written in Fortran) that returns S and T so that the eigenvalues of the complex matrix S with modulus less than one appear in the upper left block of S. This routine in turn calls the program `ZGEES` from the Fortran LAPACK library. Our Fortran version of `SolveLA` also uses `ZGGES` to get the Schur decomposition with sorted eigenvalues. The Gauss version of `SolveLA` (and `SolveLA2`) also solves purely deterministic models. Just set the matrices C_z, F_z, D_z and Π equal to the Gauss missing value code.

The matrices that are an input to both programs can be obtained in two ways. The first and probably more cumbersome approach is to use paper and pencil to derive the coefficients of the matrices analytically. If the differentiation is done with respect to the (natural) logs of the variables, `SolveLA` returns the coefficients of the log-linear policy functions. Otherwise the coefficients refer to the linear approximation. One may, however, also use numeric differentiation to obtain the matrices from (2.47). We provide an example in the Gauss program `Ramsey3a.g` where we show how to solve the stochastic growth model by using `SolveLA`.

2.5 Quadratic Approximation

In this section we consider quadratic approximations of the policy functions of DGE models. We introduce you to this topic in the next subsection. Then, we consider two examples before we provide the general algorithm in Subsection 2.5.4.

2.5.1 Introduction

We begin with the quadratic approximation of the solution of a system of static equilibrium conditions. Consider the equilibrium condition $g(x, y) = 0$ and suppose that a solution exists at (x^*, y^*). Let $y = h(x)$ be the solution in an ϵ neighborhood of x^*. A second-order Taylor series approximation of h at x^* is given by

$$h(x^* + \epsilon) \simeq y^* + h'(x^*)\epsilon + \frac{1}{2}(h'')^2(x^*)\epsilon^2.$$

Differentiating $g(x, h(x))$ once provides

$$g_1(x, h(x)) + g_2(x, h(x))h'(x). \tag{2.64}$$

At (x^*, y^*) this expression must equal zero, from which we obtain the solution

$$h'(x^*) = -\frac{g_1(x^*, y^*)}{g_2(x^*, y^*)}.$$

Differentiating (2.64) again and setting the result equal to zero yields:

$$h''(x^*) = -\frac{g_{11} + (g_{12} + g_{21})h' + g_{22}(h')^2}{g_2}.$$

This formula still looks pretty simple. Though straight forward, the generalization to the case of n exogenous and m endogenous variables $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{0}_{m\times 1}$ produces formulas with lots of indices. First note that in this context the quadratic approximation of the solution $h^j(\mathbf{x})$, $j = 1, 2, \dots, m$ is given by

$$\hat{h}^j(\mathbf{x}) = h^j(\mathbf{x}^*) + \mathbf{h}^j_{\mathbf{x}}\bar{\mathbf{x}} + \frac{1}{2}\bar{\mathbf{x}}' H^j \bar{\mathbf{x}}, \tag{2.65}$$

where $\mathbf{h}^j_{\mathbf{x}} = [h^j_{x_1}, h^j_{x_2}, \dots, h^j_{x_n}]'$ is the vector of linear coefficients and $H^j = (h^j_{il})$ is the n-by-n matrix of quadratic coefficients. The vectors $\mathbf{h}^j_{\mathbf{x}}$ are determined from the matrix equation

$$\mathbf{h}_x = -D_{\mathbf{y}}^{-1} D_{\mathbf{x}}$$

where $D_{\mathbf{y}}$ ($D_{\mathbf{x}}$) is the matrix of partial derivatives of $\mathbf{g}(\mathbf{x}, \mathbf{y})$ with respect to the variables in the vector $\mathbf{y}$ ($\mathbf{x}$) (see equation (11.38)). Note, that a single element in this matrix equation is given by

$$\begin{aligned} &0 = g^j_{x_k}(\mathbf{x}, \mathbf{h}(\mathbf{x})) + \sum_{l=1}^{m} g^j_{y_l}(\mathbf{x}, \mathbf{h}(\mathbf{x})) h^l_{x_k}, \\ &j = 1, 2, \dots, m, \quad k = 1, 2, \dots, n. \end{aligned}$$

Differentiating this expression with respect to variable x_i provides

$$\begin{aligned} 0 = g^j_{x_k x_i} &+ \sum_{l=1}^{m} g^j_{x_k y_l} h^l_{x_i} + \sum_{l=1}^{m} g^j_{y_l} h^l_{x_k x_i} + \sum_{l=1}^{m} g^j_{y_l x_i} h^l_{x_k} \\ &+ \sum_{s=1}^{m}\sum_{l=1}^{m} g^j_{y_l y_s} h^s_{x_i} h^l_{x_k}, \quad j = 1, \dots, m;\ i, k = 1, \dots, n. \end{aligned}$$

These mn^2 equations can be arranged to n^2 matrix equations in the coefficients $h^j_{x_k x_j}$, $j = 1, 2, \dots, m$. Due to the symmetry of the Hesse matrices $n(n+1)/2$ of these equations are redundant. As you will see in the next examples, since the structure of the equilibrium conditions of DGE models is not as simple as $\mathbf{g}(\mathbf{x}, \mathbf{h}(\mathbf{x})) = 0_{m\times 1}$, the respective formulas to compute the Hesse matrices H^j are more involved.

2.5.2 The Deterministic Growth Model

We return to the deterministic growth model considered in Sections 1.2 and 2.1. We let $K_{t+1} = h^K(K_t)$ and $C_t = h^C(K_t)$ denote the policy functions for the next-period capital stock and consumption, respectively. For both functions we seek a second order approximation at the stationary solution K^* of the form

$$\begin{aligned}\bar{K}_{t+1} &= h_K^K \bar{K}_t + \frac{1}{2} h_{KK}^K \bar{K}_t^2,\\ \bar{C}_t &= h_K^C \bar{K}_t + \frac{1}{2} h_{KK}^C \bar{K}_t^2,\end{aligned}$$

where h_K^i and h_{KK}^i, $i \in \{K, C\}$ denote the first and second derivative of the policy function of variable i with respect to the stock of capital K. Of course, all derivatives are evaluated at the stationary solution K^*. To obtain the coefficients h_K^i and h_{KK}^i, we use a more general exposition. Observe that the resource constraint $g^1(\cdot)$ and the Euler equation for the optimal next-period capital stock $g^2(\cdot)$, equations (2.1a) and (2.1b), have the following structure:

$$\begin{aligned}&g^i(K, C, K', C')\\ &\equiv g^i(K, h^C(K), h^K(K), h^C(h^K(K))) = 0, \quad i = 1, 2,\end{aligned}$$

where we have omitted the time indices. To distinguish between current period variables and next-period variables we used a prime to denote the latter. Differentiating with respect to K yields (we suppress the arguments of g^i but not of h^i)

$$\begin{aligned}&g_K^i + g_C^i h_K^C(K) + g_{K'}^i h_K^K(K) + g_{C'}^i h_K^C(K') h_K^K(K) = 0,\\ &i = 1, 2.\end{aligned} \tag{2.66}$$

We have already solved these two equations in Section 2.1, so let us assume here that we know h_K^K and h_K^C. To obtain equations in h_{KK}^K and h_{KK}^C, we must differentiate (2.66) with respect to K. This yields:

$$\begin{bmatrix} g_{K'}^1 + g_{C'}^1 h_K^C & g_C^1 + g_{C'}^1 \left(h_K^K\right)^2 \\ g_{K'}^2 + g_{C'}^2 h_K^C & g_C^2 + g_{C'}^2 \left(h_K^K\right)^2 \end{bmatrix} \begin{bmatrix} h_{KK}^K \\ h_{KK}^C \end{bmatrix} = \begin{bmatrix} \mathbf{h}_K^T H(g^1) \mathbf{h}_K \\ \mathbf{h}_K^T H(g^2) \mathbf{h}_K \end{bmatrix}, \tag{2.67}$$

where

$$\mathbf{h}_K^T = \left[1, h_K^C, h_K^K, h_K^C h_K^K\right], \quad H(g^i) := \begin{bmatrix} g_{KK}^i & \cdots & g_{KC'}^i \\ \vdots & \ddots & \vdots \\ g_{C'K}^i & \cdots & g_{C'C'}^i \end{bmatrix}.$$

Since (2.67) is a system of two linear equations it is easily solved for h_{KK}^K and h_{KK}^C. Usually, we will use numeric differentiation to obtain the coefficients of equation (2.67). If $u(C) := (C^{1-\eta} - 1)/(1-\eta)$ and $f(K) = (1-\delta)K + K^\alpha$, the matrix on the lhs of (2.67) is given by

$$\begin{bmatrix} 1 & 1 \\ \left[\eta h_K^C + \alpha\beta(1-\alpha)C^*(K^*)^{\alpha-2}\right] & \eta\left[\left(h_K^K\right)^2 - 1\right] \end{bmatrix}$$

and the vector on the rhs, say $\mathbf{b}$, has elements

$$b_1 := \alpha(\alpha-1)(K^*)^{\alpha-2}$$

and

$$\begin{aligned} b_2 := {} & \eta(1+\eta)\frac{1}{C^*}\left[\left(h_{KK}^K\right)^2 - 1\right]\left(h_K^C\right)^2 \\ & + \alpha\beta(1-\alpha)(K^*)^{\alpha-2}\left(h_K^K\right)^2\left(2\eta h_K^C + (2-\alpha)\frac{C^*}{K^*}\right). \end{aligned}$$

In the Gauss program `Ramsey2b.g` we compute the coefficients of the quadratic policy functions using both analytic and numeric derivatives. Figure 2.4 displays the policy function for consumption from the linear, the quadratic solution and the analytic solution ($\alpha = 0.27$ and $\beta = 0.994$).

To compare the accuracy of the linear with the accuracy of the quadratic approximation this program also computes the residuals of the Euler equation (2.1b) over a grid of 200 points in the interval $[0.9K^*, 1.1K^*]$. For the parameter values $\alpha = 0.27$, $\beta = 0.994$, $\eta = 2$, and $\delta = 0.011$ we find that the maximum absolute Euler equation residual from the linear solution is about 13 times larger than that obtained from the quadratic policy function which is 2.4×10^{-6}, and, thus, very small. We also find that there is no noteworthy difference in accuracy, if we use analytic instead of numeric derivatives.

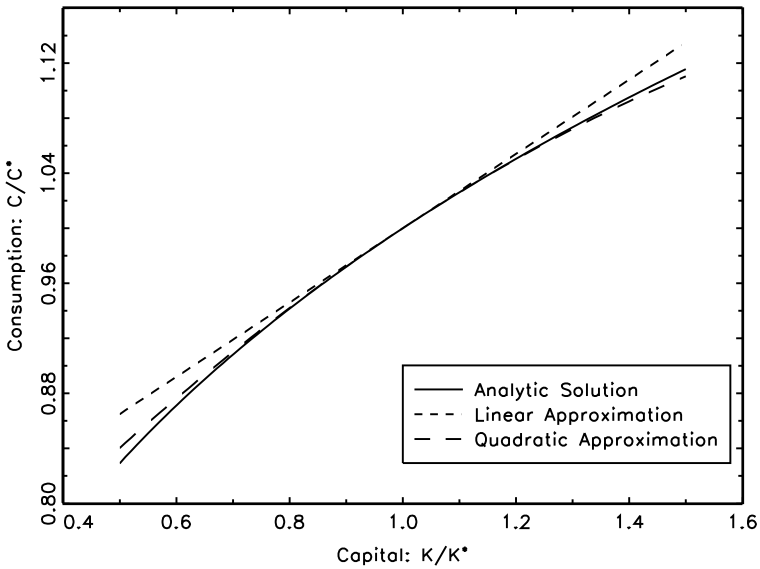

Figure 2.4: Policy Functions of Consumption of the Deterministic Growth Model

2.5.3 The Stochastic Growth Model

The Framework. We return to the stochastic growth model considered in Sections 1.3 and 2.4.1 assuming $u(C) = [C^{1-\eta} - 1]/(1 - \eta)$ and $f(K) = K^{\alpha}$. As in the previous subsection, we drop the time indices from all variables and use a prime to designate variables that pertain to the next period. This allows us to the write the equilibrium conditions as[16]

$$
\begin{aligned}
0 &= Eg^i(K, C, z, K', C', z'), \quad i = 1, 2, && (2.68a)\\
C &= h^C(K, z, \sigma),\\
C' &= h^C(h^K(K, z, \sigma), z', \sigma),
\end{aligned}
$$

[16] You may probably wonder why we use $z = \ln Z$ as a state variable and not Z itself. In the present context, in which we know what the equilibrium conditions look like, we could indeed have used Z. Yet, when writing a general purpose routine, we have no information about the structure of the equilibrium conditions. In this case, we are bound to assume that the shocks evolve according to a linear first-order autoregressive process.

$$K' = h^K(K, z, \sigma),$$
$$z' = \varrho z + \sigma\epsilon', \quad \epsilon' \sim N(0, 1),$$

where

$$g^1(\cdot) = K' - (1-\delta)K - e^z K^\alpha + C, \tag{2.68b}$$
$$g^2(\cdot) = C^{-\eta} - \beta(C')^{-\eta}\left(1 - \delta + \alpha e^{z'}(K')^{\alpha-1}\right). \tag{2.68c}$$

The operator E denotes expectations with respect to information available at the current period.

As in Section 2.4.1 we consider the model in a neighborhood of $\sigma = 0$, where it reduces to the deterministic growth model with stationary solution $(K^*, C^*, z^* = 0)$ determined by equations (2.33). For $i \in \{C, K\}$ we look for quadratic approximations of the policy function h^i given by

$$\begin{aligned} h^i(K, z, \sigma) = {} & h^i(K^*, z^*, \sigma = 0) \\ & + h^i_K \bar{K} + h^i_z \bar{z} + h^i_\sigma \sigma \\ & + \frac{1}{2}\left[\bar{K}, \bar{z}, \sigma\right] \begin{bmatrix} h^i_{KK} & h^i_{Kz} & h^i_{K\sigma} \\ h^i_{zK} & h^i_{zz} & h^i_{z\sigma} \\ h^i_{\sigma K} & h^i_{\sigma z} & h^i_{\sigma\sigma} \end{bmatrix} \begin{bmatrix} \bar{K} \\ \bar{z} \\ \sigma \end{bmatrix}, \end{aligned} \tag{2.69}$$

where the bar denotes deviations from the stationary solution. Note that the Hesse matrix in (2.69) is a symmetric matrix, i.e., $h^i_{jk} = h^i_{kj}$, $j, k \in \{K, z, \sigma\}$. To determine the coefficients of these functions we closely follow SCHMITT-GROHÉ and URIBE (2004).[17]

As in Section 2.4.1 we differentiate (2.68a) with respect to K, z, and σ. To represent the respective formulas we define the vector function

$$\mathbf{h} := \begin{bmatrix} h^C(K, z, \sigma) \\ h^K(K, z, \sigma) \\ h^C(h^K(K, z, \sigma), \varrho z + \sigma\epsilon', \sigma) \end{bmatrix}$$

with the vector of derivatives denoted by $\mathbf{h}_K$, $\mathbf{h}_z$, and $\mathbf{h}_\sigma$, respectivley. In addition, we use $g^i_{[\mathbf{i}]}$ for the (column) vector of first

[17] In a recent paper LOMBARDO and SUTHERLAND (2007) outline an algorithm that also provides second-order accurate solutions. Their procedure relies on methods developed for the solution of linear models.

derivatives of g^i with respect to the indices in the vector $\mathbf{i}$ and $g^i_{[\mathbf{i}_1][\mathbf{i}_2]}$ for the matrix of second partial derivatives with respect to the indices in $\mathbf{i}_1$ (for the rows of the matrix) and $\mathbf{i}_2$ (for the columns). To avoid confusion, we denote the transpose of a vector by the superscript T.

Consider the derivatives of conditions (2.68a) with respect to K, z, and σ:

$$0 = E\left\{\left[1, \mathbf{h}_K^T\right] g^i_{[K,C,K',C']}\right\}, \tag{2.70a}$$

$$0 = E\left\{\left[\mathbf{h}_z^T, 1, \varrho\right] g^i_{[C,K',C',z,z']}\right\}, \tag{2.70b}$$

$$0 = E\left\{\left[\mathbf{h}_\sigma^T, \epsilon'\right] g^i_{[C,K',C',z']}\right\}. \tag{2.70c}$$

Since we have already seen how we can compute the coefficients of the linear part of (2.69) in Section 2.4.1, we proceed to the coefficients of the quadratic part. For the following derivations we will keep in mind that we found $h^i_\sigma = 0$.

Coefficients of the Hesse Matrices. Differentiating equation (2.70a) with respect to K provides two linear equations in the coefficients h^i_{KK}:

$$0 = \mathbf{h}_{KK}^T g^i_{[C,K',C']} + \left[1, \mathbf{h}_K^T\right] g^i_{[K,C,K',C'][K,C,K',C']} \begin{bmatrix} 1 \\ \mathbf{h}_K \end{bmatrix}, \tag{2.71a}$$

where $\mathbf{h}_{KK}$ is the vector of second derivatives of $\mathbf{h}$ with respect to K. This equation corresponds to equation (2.67) in the deterministic case.

To determine h^i_{Kz}, we differentiate (2.70a) with respect to z, yielding

$$0 = \mathbf{h}_{Kz}^T g^i_{[C,K',C']} + \left[1, \mathbf{h}_K^T\right] g^i_{[K,C,K',C'][C,K',C',z]} \begin{bmatrix} \mathbf{h}_Z \\ 1 \end{bmatrix} + \varrho\left[1, \mathbf{h}_K^T\right] g^i_{[K,C,K',C'][z']}. \tag{2.71b}$$

The first term in this equation equals

$$\left(g^i_{K'} + g^i_{C'} h^C_K\right) h^K_{KZ} + \left(g^i_C + \varrho g^i_{C'} h^K_K\right) h^C_{KZ} + g^i_{C'} h^K_K h^C_{KK} h^K_Z.$$

Thus, (2.71b) provide two linear equations in h^K_{KZ} and h^C_{KZ}.

Differentiating conditions (2.70a) with respect to σ provides conditions on $h^i_{K\sigma}$:

$$0 = E\left\{ \mathbf{h}^T_{K\sigma} g^i_{[C,K'C']} + \left[1, \mathbf{h}^T_K\right] g^i_{[K,C,K',C'][C,K',C']} \mathbf{h}_\sigma \right. \tag{2.71c}$$
$$\left. + \left[1, \mathbf{h}^T_K\right] g^i_{[K,C,K',C'][z']} \epsilon' \right\}.$$

The expectation of the first term in curly brackets is

$$E\left\{\mathbf{h}^T_{K\sigma} g^i_{[C,K',C']}\right\} = \left(g^i_{K'} + g^i_{C'} h^C_K\right) h^K_{K\sigma} + \left(g^i_C + g^i_{C'} h^K_K\right) h^C_{K\sigma},$$

since $h^K_\sigma = 0$ and $E(h^K_K h^C_{KZ} \epsilon') = 0$. At the stationary solution the second term in (2.71c) is obviously zero, since $\mathbf{h}_\sigma$ is a vector with zeros. The expectation of the third term is also zero since $E(\epsilon') = 0$. Thus, system (2.71c) is a linear homogeneous system with solution $h^i_{K\sigma} = 0$.

To determine the coefficients h^i_{zz}, we differentiate (2.70b) with respect to z. The result is:

$$0 = \mathbf{h}^T_{zz} g^i_{[C,K',C']}$$
$$+ \left[\mathbf{h}^T_z, 1, \varrho\right] g^i_{[C,K',C',z,z'][C,K',C',z,z']} \begin{bmatrix} \mathbf{h}_z \\ 1 \\ \varrho \end{bmatrix}. \tag{2.71d}$$

The first term on the rhs of this equation equals

$$\mathbf{h}^T_{zz} g^i_{[C,K',C']} = \left(g^i_{K'} + g^i_{C'} h^C_K\right) h^K_{ZZ} + \left(g^i_C + g^i_{C'} \varrho^2\right) h^C_{ZZ}$$
$$+ g^i_{C'} h^K_Z \left(h^C_{KK} h^K_Z + 2\varrho h^C_{KZ}\right).$$

Differentiating (2.70b) with respect to σ provides

$$0 = E\left\{ \mathbf{h}^T_{z\sigma} g^i_{[C,K',C']} \right.$$
$$\left. + \left[\mathbf{h}^T_z, 1, \varrho\right] g^i_{[C,K',C',z,z'][C,K',C',z']} \begin{bmatrix} \mathbf{h}_\sigma \\ \epsilon' \end{bmatrix} \right\}. \tag{2.71e}$$

As in equation (2.71c) all terms except the coefficients of $h^K_{Z\sigma}$ and $h^C_{Z\sigma}$ are equal to zero. Therefore, $h^K_{Z\sigma} = h^C_{Z\sigma} = 0$.

Finally, we turn to the coefficients $h^i_{\sigma\sigma}$. They are obtained from differentiating equations (2.70c) with respect to σ. This delivers:

$$\begin{aligned}
0 &= E\Big\{\mathbf{h}^T_{\sigma\sigma} g^i_{[C,K',C']} \\
&\qquad + \left[\mathbf{h}^T_\sigma, \epsilon'\right] g^i_{[C,K',C',z'][C,K',C',z']} \begin{bmatrix} \mathbf{h}_\sigma \\ \epsilon' \end{bmatrix} \Big\}, \qquad (2.71\text{f})\\
\mathbf{h}^T_{\sigma\sigma} &= \left[h^C_{\sigma\sigma}, h^K_{\sigma\sigma}, h^C_{\sigma\sigma} + h^C_K h^K_{\sigma\sigma} + \Delta\right], \\
\Delta &:= h^K_\sigma \left(h^C_{KK} h^K_\sigma + h^C_{KZ}\epsilon' + h^C_{K\sigma}\right) \\
&\quad + \epsilon' \left(h^C_{ZK} h^K_\sigma + h^C_{ZZ}\epsilon' + h^C_{z\sigma}\right) + h^C_{\sigma K} h^K_\sigma + h^C_{\sigma z}\epsilon'.
\end{aligned}$$

To evaluate this expression, observe that

1. at $\sigma = 0$ the vector of derivatives $\mathbf{h}^T_\sigma$ equals $[0, 0, h^C_Z \epsilon]$, since $h^i_\sigma = 0$,
2. $h^i_{\sigma j} = h^i_{j\sigma} = 0$ for $i \in \{K, C\}$ and $j \in \{K, z\}$,
3. $E(\epsilon')^2 = 1$ and $E(\epsilon') = 0$.

Thus, equations (2.71f) reduce to

$$\begin{aligned}
0 &= \left(g^i_{K'} + g^i_{C'} h^C_K\right) h^K_{\sigma\sigma} + \left(g^i_C + g^i_{C'}\right) h^C_{\sigma\sigma} \\
&\quad + g^i_{C'C'}(h^C_Z)^2 + 2g^i_{C'z'} h^C_Z + g^i_{z'z'} + g^i_{C'} h^C_{ZZ}.
\end{aligned}$$

Our Gauss program `Ramsey3b.g` computes the quadratic approximation of the policy function from these formulas. It employs numeric differentiation to compute $g^i_{[\cdot]}$ as well as the Hesse matrices that appear in (2.71).

Table 2.1 presents the coefficients from this exercise for the parameter values $\alpha = 0.27$, $\beta = 0.994$, $\eta = 1$, and $\delta = 1$. The second column shows solutions obtained from using the Gauss commands `gradp` and `hessp` that provide forward difference approximations of the first and second partial derivatives, respectively.[18] Our own procedures `CDJac` and `CDHesse` imple-

[18] See Section 11.3.1 on numeric differentiation, where we explain forward difference as well as central difference formulas for the numeric computation of derivatives.

Table 2.1

Coefficient	Forward Differences	Central Differences	Analytic solution
h^K_K	0.270000	0.270000	0.270000
h^K_Z	0.164993	0.164993	0.164993
h^K_{KK}	−1.194628	−1.194595	−1.194595
h^K_{KZ}	0.269781	0.270000	0.270000
h^K_{ZZ}	0.156787	0.164995	0.164993
$h^K_{\sigma\sigma}$	−0.023160	0.000001	0.000000
h^C_K	0.736036	0.736036	0.736036
h^C_Z	0.449782	0.449781	0.449781
h^C_{KK}	−3.256642	−3.256537	−3.256538
h^C_{KZ}	0.735831	0.736036	0.736036
h^C_{ZZ}	0.479034	0.449782	0.449781
$h^C_{\sigma\sigma}$	0.023160	−0.000001	0.000000

ment central difference formulas that involve a smaller approximation error. The fourth column presents the coefficients computed from the quadratic approximation of the analytic solutions $h^K = \alpha\beta e^z K^\alpha$ and $h^C = (1-\alpha\beta)e^z K^\alpha$, respectively. There is no noteworthy difference in the linear coefficients as well as in h^i_{KK}. There is a small difference between the solutions for h^K_{KZ}, but the numeric value of $h^i_{\sigma\sigma}$ is far from its true value of zero when we use forward difference formulas. This imprecision can also be seen from the residuals of the Euler equation

$$C_t^{-\eta} = E_t \beta C_{t+1}^{-\eta} \left(1 - \delta + \alpha(e^{\varrho z_t + \sigma\epsilon_{t+1}})K_{t+1}^{\alpha-1}\right). \tag{2.72}$$

We compute the residuals on a grid of 400 equally spaced points on the square $[0.9K^*, 1.2K^*] \times [\ln(0.95), \ln(1.05)]$. With respect to the maximum absolute value of these residuals the solution displayed in the second column of Table 2.1 is about 2.5 times worse than the solution based on the numbers in column four. The Euler equation residual from the linear solution is almost 37 times larger than the Euler equation residual from the quadratic solution displayed in column four. When we use the parameter values from Table 1.1 for α, β, η, δ, ϱ, and σ, the linear solution is about 13 times less accurate than the quadratic solution, whose maximum absolute Euler equation residual is 4.6×10^{-6}.

Computation of the Euler Equation Residual. Here we briefly explain our computation of the residual in the stochastic growth model. Given the approximate policy functions $\hat{h}^K$ and $\hat{h}^C$ the term to the right of the expectations operator E_t in equation (2.72) can be written as

$$\begin{aligned}\phi(K, Z, \sigma, \epsilon) := \beta &\left(\hat{h}^C(\hat{h}^K(K, Z, \sigma), e^{\varrho z_t + \sigma\epsilon}, \sigma)\right)^{-\eta} \\ &\times \left(1 - \delta + \alpha e^{\varrho z_t + \sigma\epsilon}\left(\hat{h}^K(K, Z, \sigma)\right)^{\alpha - 1}\right).\end{aligned}$$

For given values of K, z, and σ this is a function of the stochastic variable ϵ that has a standard normal distribution. Therefore, the rhs of equation (2.72) is given by

$$\Delta := \int_{-\infty}^{\infty} \phi(K, Z, \sigma, \epsilon) \frac{e^{-\frac{\epsilon^2}{2}}}{\sqrt{2\pi}} d\epsilon.$$

We use the Gauss-Hermite four point integration formula given in equation (11.77) to compute this expectation. Given Δ, the Euler equation residual at (K, Z) is defined as

$$\tilde{R} = \frac{\Delta^{-1/\eta}}{\hat{h}^C(K, Z, \sigma)} - 1.$$

2.5.4 Generalization

Framework. Equations (2.68a) are readily generalized. Just replace K by an $n(x)$ vector $\mathbf{x}$ of state variables, C by a $n(y)$ vector $\mathbf{y}$ of control and costate variables, Z by an $n(z)$ vector of shocks $\mathbf{z}$, and ϵ by a $n(z)$ vector $\boldsymbol{\epsilon}$ of $N(\mathbf{0}_{n(z)}, I_{n(z)})$ distributed innovations so that $\mathbf{z}_t = \Pi\mathbf{z}_{t-1} + \sigma\Omega\boldsymbol{\epsilon}$. The $n(z)$ by $n(z)$ matrix Ω allows for possible correlations between the elements of $\mathbf{z}$. To see this, note that the conditional variance of $\mathbf{z}_t$ given $\mathbf{z}_{t-1}$ is given by $E(\sigma\Omega\boldsymbol{\epsilon})(\sigma\Omega\boldsymbol{\epsilon})^T = \sigma^2\Omega\Omega^T$, where the superscript T denotes the transposition of a matrix or a vector.

The $n(x)+n(y)$ equilibrium conditions are

$$0 = Eg^i(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{x}',\mathbf{y}',\mathbf{z}'), \quad i = 1,2,\dots,n(x)+n(y), \tag{2.73a}$$

where

$$\mathbf{y} = \mathbf{h}^y(\mathbf{x},\mathbf{z},\sigma), \tag{2.73b}$$
$$\mathbf{x}' = \mathbf{h}^x(\mathbf{x},\mathbf{z},\sigma), \tag{2.73c}$$
$$\mathbf{y}' = \mathbf{h}^y(\mathbf{x}',\mathbf{z}',\sigma), \tag{2.73d}$$
$$\mathbf{z}' = \Pi\mathbf{z} + \sigma\Omega\boldsymbol{\epsilon}' \tag{2.73e}$$

The quadratic approximation of the policy function h^i, $i \in \{x_1,\dots,x_{n(x)},y_1,\dots,y_{n(y)}\}$ is an expression of the form

$$\begin{aligned} h^i &= h^i(\mathbf{x}^*,\mathbf{z}^*,\sigma=0) + (\mathbf{l}^i)^T \begin{bmatrix}\bar{\mathbf{x}}\\ \bar{\mathbf{z}}\end{bmatrix} \\ &\quad + \frac{1}{2}\left[\bar{\mathbf{x}}^T,\bar{\mathbf{z}}^T,\sigma\right] \underbrace{\begin{bmatrix} H^i_{\mathbf{xx}} & H^i_{\mathbf{xz}} & \mathbf{0} \\ H^i_{\mathbf{zx}} & H^i_{\mathbf{zz}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & H^i_{\sigma\sigma}\end{bmatrix}}_{H^i} \begin{bmatrix}\bar{\mathbf{x}}\\ \bar{\mathbf{z}}\\ \sigma\end{bmatrix}. \end{aligned} \tag{2.74}$$

The row vector $\mathbf{l}^i$ holds the coefficients of the linear part and the matrices $H^i_{\mathbf{xx}}$, $H^i_{\mathbf{xz}}$, and $H^i_{\mathbf{zz}}$ contain the coefficients of the quadratic part with respect to the state variables $\mathbf{x}$ and $\mathbf{z}$. As before, the bar denotes deviations from the equilibrium $\mathbf{x}^*$ and $\mathbf{z}^*$, respectively. The scalar $H^i_{\sigma\sigma}$ is the coefficient of σ^2. Note that in the general model both the linear coefficients of σ are zero and the matrices $H^i_{\mathbf{x}\sigma}$ and $H^i_{\mathbf{z}\sigma}$ are zero matrices as in the example of the previous subsection.[19]

Computation of the Quadratic Part. To obtain these matrices we proceed as in our example. Given the vector

[19] See Schmitt-Grohé and Uribe (2004) for a proof.

$$\mathbf{h} := \begin{bmatrix} h^{y_1}(\mathbf{x}, \mathbf{z}, \sigma) \\ \vdots \\ h^{y_{n(y)}}(\mathbf{x}, \mathbf{z}, \sigma) \\ h^{x_1}(\mathbf{x}, \mathbf{z}, \sigma) \\ \vdots \\ h^{x_{n(x)}}(\mathbf{x}, \mathbf{z}, \sigma) \\ h^{y_1}(\mathbf{x}', \mathbf{z}', \sigma) \\ \vdots \\ h^{y_{n(y)}}(\mathbf{x}', \mathbf{z}', \sigma) \end{bmatrix}, \tag{2.75}$$

we use $\mathbf{h}_i$ to denote the vector whose elements are the derivatives of the elements of $\mathbf{h}$ with respect to variable i.

We begin with the coefficients of the matrices $H_{\mathbf{xx}}$. We differentiate equations (2.73a) with respect to x_j and evaluate the result at the point $(\mathbf{x}^*, \mathbf{z}^*, \sigma = 0)$:

$$0 = \left[1, \quad \mathbf{h}_{x_j}^T\right] g^i_{[x_j, \mathbf{y}, \mathbf{x}', \mathbf{y}']}, \quad i = 1, 2, \ldots, n(x) + n(y). \tag{2.76}$$

Differentiating this expression with respect to x_k provides a set of $(n(x) + n(y))n(x)^2$ conditions in the unknown coefficients of the matrices $H^i_{\mathbf{xx}}$:

$$\mathbf{h}^T_{x_j x_k} g^i_{[x_j, \mathbf{y}, \mathbf{x}', \mathbf{y}']} = -\left[1, \mathbf{h}^T_{x_j}\right] g^i_{[x_j, \mathbf{y}, \mathbf{x}', \mathbf{y}'][x_k, \mathbf{y}, \mathbf{x}', \mathbf{y}']} \begin{bmatrix} 1 \\ \mathbf{h}_{x_k} \end{bmatrix}, \tag{2.77a}$$

$$i = 1, \ldots, n(x) + n(y),\ j = 1, \ldots, n(x),\ k = 1, \ldots, n(x),$$

where

$$\mathbf{h}^T_{x_j} = \left[h^{y_1}_{x_j}, \ldots, h^{y_{n(y)}}_{x_j}, h^{x_1}_{x_j}, \ldots, h^{x_{n(x)}}_{x_j}, \Delta^1_1, \ldots, \Delta^1_{n(y)}\right], \tag{2.77b}$$

$$\Delta^1_i = \sum_{l=1}^{n(x)} h^{y_i}_{x_l} h^{x_l}_{x_j}.$$

and

$$\mathbf{h}^T_{x_j x_k} = \left[h^{y_1}_{x_j x_k}, \ldots, h^{y_{n(y)}}_{x_j x_k}, h^{x_1}_{x_j x_k}, \ldots, h^{x_{n(x)}}_{x_j x_k}, \Delta^2_i, \ldots, \Delta^2_{n(y)}\right], \tag{2.77c}$$

$$\Delta^2_i = \sum_{l=1}^{n(x)} h^{y_i}_{x_l} h^{x_l}_{x_j x_k} + \sum_{l=1}^{n(x)} h^{x_l}_{x_j} \sum_{r=1}^{n(x)} h^{y_i}_{x_l x_r} h^{x_r}_{x_k}.$$

Different from our example in the previous subsection the system of equations (2.77a) cannot be factored into smaller systems in the pairs of coefficients (x_j, x_k), since all the unknown coefficients $h^{y_i}_{x_l x_r}$ appear in each equation. The huge linear system (2.77a) may be written as $A\mathbf{w} = \mathbf{q}$, where

$$\mathbf{w} := \operatorname{vec}\left[H^{x_1}_{\mathbf{xx}}, \quad \dots, \quad H^{x_{n(x)}}_{\mathbf{xx}}, \quad H^{y_1}_{\mathbf{xx}}, \quad \dots, \quad H^{y_{n(y)}}_{\mathbf{xx}}\right].$$

The element $h^{x_i}_{x_j x_k}$ in this vector has the index $ix(i,j,k) = (i-1)n(x)^2 + (j-1)n(x) + k$. The index of $h^{y_i}_{x_j x_k}$ is $iy(i,j,k) = n(x)^3 + ix(i,j,k)$. Using the functions ix and iy it is easy to loop over $j = 1, \dots, n(x)$, $k = 1, \dots, n(x)$, and $i = 1, \dots, n(x) + n(y)$ to set up the matrix A and the vector $\mathbf{q}$ from (2.77a).

The elements of the matrices $H^i_{\mathbf{xz}}$ solve

$$\mathbf{h}^T_{x_j z_k} g^i_{[\mathbf{y},\mathbf{x}',\mathbf{y}']} = -\left[1, \mathbf{h}^T_{x_j}\right] g^i_{[x_j,\mathbf{y},\mathbf{x}',\mathbf{y}'][\mathbf{y},\mathbf{x}',\mathbf{y}',z_k,\mathbf{z}']} \begin{bmatrix} \mathbf{h}_{z_k} \\ 1 \\ \pi_{1k} \\ \vdots \\ \pi_{n(z)k} \end{bmatrix}, \tag{2.78}$$
$$i = 1, \dots, n(x) + n(y),\ j = 1, \dots, n(x),\ k = 1, \dots, n(z),$$

where π_{lk} is the element in the lth row and kth column of the matrix Π from equation (2.73e). This system is derived from differentiating (2.76) with respect to z_k. The elements of the vector $\mathbf{h}_{z_k}$ are the derivatives of (2.75) with respect to z_k:

$$\begin{aligned} \mathbf{h}^T_{z_k} &:= \left[h^{y_1}_{z_k}, \dots, h^{y_{n(y)}}_{z_k}, h^{x_1}_{z_k}, \dots, h^{x_{n(x)}}_{z_k}, \Delta^3_1, \dots, \Delta^3_{n(y)}\right], \\ \Delta^3_i &= \sum_{l=1}^{n(x)} h^{y_i}_{x_l} h^{x_l}_{z_k} + \sum_{l=1}^{n(z)} h^{y_i}_{z_l} \pi_{lk}. \end{aligned} \tag{2.79}$$

Differentiating the elements of (2.77b) with respect to z_k provides the vector $\mathbf{h}_{x_j z_k}$:

$$\mathbf{h}^T_{x_j z_k} := \left[h^{y_1}_{x_j z_k}, \dots, h^{y_{n(y)}}_{x_j z_k}, h^{x_1}_{x_j z_k}, \dots, h^{x_{n(y)}}_{x_j z_k}, \Delta^4_1, \dots, \Delta^4_{n(y)}\right],$$
$$\Delta^4_i := \sum_{l=1}^{n(x)} h^{y_i}_{x_l} h^{x_l}_{x_j z_k} + \sum_{l=1}^{n(x)} h^{x_l}_{x_j} \left[\sum_{r=1}^{n(x)} h^{y_i}_{x_l x_r} h^{x_r}_{z_k} + \sum_{r=1}^{n(z)} h^{y_i}_{x_l z_r} \pi_{rk}\right].$$

The system of equations (2.78) may also be written as $A\mathbf{w} = \mathbf{q}$. But note that different from (2.77a) the lhs of (2.78) not only contains the elements of $H^i_{\mathbf{xz}}$ but also terms that belong to the vector $\mathbf{q}$.

To obtain the matrices $H^i_{\mathbf{zz}}$ we first differentiate (2.73a) with respect to z_j and then with respect to z_k. The result is:

$$\begin{aligned}
\mathbf{h}^T_{z_j z_k} g^i_{[\mathbf{y},\mathbf{x}',\mathbf{y}']} &= -\left[\mathbf{h}^T_{z_j}, 1, \pi_{1j}, \dots, \pi_{n(z),j}\right] \\
&\quad \times g^i_{[\mathbf{y},\mathbf{x}',\mathbf{y}',z_j,\mathbf{z}'][\mathbf{y},\mathbf{x}',\mathbf{y}',z_k,\mathbf{z}']} \begin{bmatrix} \mathbf{h}_{z_k} \\ 1 \\ \pi_{1k} \\ \vdots \\ \pi_{n(z),k} \end{bmatrix},
\end{aligned} \tag{2.80}$$

$$\begin{aligned}
\mathbf{h}^T_{z_j z_k} &= \left[h^{y_1}_{z_j z_k}, \dots, h^{y_{n(y)}}_{z_j z_k}, h^{x_1}_{z_j z_k}, \dots, h^{x_{n(x)}}_{z_j z_j}, \Delta^5_1, \dots, \Delta^5_{n(y)}\right], \\
\Delta^5_i &= \sum_{l=1}^{n(x)} h^{y_i}_{x_l} h^{x_l}_{z_j z_k} \\
&\quad + \sum_{l=1}^{n(x)} h^{x_l}_{z_j} \left[\sum_{r=1}^{n(x)} h^{y_i}_{x_l x_r} h^{x_r}_{z_k} + \sum_{r=1}^{n(z)} h^{y_i}_{x_l z_r} \pi_{rk}\right] \\
&\quad + \sum_{l=1}^{n(z)} \pi_{lj} \left[\sum_{r=1}^{n(x)} h^{y_i}_{z_l x_r} h^{x_r}_{z_k} + \sum_{r=1}^{n(z)} h^{y_i}_{z_l z_r} \pi_{rk}\right].
\end{aligned}$$

In the last step, we determine $H^i_{\sigma\sigma}$. Differentiating (2.73a) twice with respect to σ yields

$$\begin{aligned}
0 &= E\left\{\mathbf{h}^T_{\sigma\sigma} g^i_{[\mathbf{y},\mathbf{x}',\mathbf{y}']}\right\} \\
&\quad + E\left\{ \left[\mathbf{h}^T_{\sigma}, \Delta^6_1, \dots, \Delta^6_{n(z)}\right] g^i_{[\mathbf{y},\mathbf{x}',\mathbf{y}',\mathbf{z}'][\mathbf{y},\mathbf{x}',\mathbf{y}',\mathbf{z}']} \begin{bmatrix} \mathbf{h}_{\sigma} \\ \Delta^6_1 \\ \vdots \\ \Delta^6_{n(z)} \end{bmatrix} \right\}, \\
\Delta^6_i &:= \sum_{s=1}^{n(z)} \omega_{is} \epsilon'_s,
\end{aligned} \tag{2.81}$$

where ω_{is} is the element in the ith row and sth column of the matrix Ω from equation (2.73e). At the stationary equilibrium, the vector $\mathbf{h}_\sigma$ is given by

$$\mathbf{h}_\sigma^T = \left[\underbrace{0,\dots,0,}_{n(y)\text{ elements}}\ \underbrace{0,\dots,0,}_{n(x)\text{ elements}}\ \Delta_1^7,\dots,\Delta_{n(y)}^7\right],$$

$$\Delta_i^7 = \sum_{s=1}^{n(z)} h_{z_s}^{y_i} \sum_{r=1}^{n(z)} \omega_{sr}\epsilon'_r,$$

since in the general model as well as in our example $h_\sigma^{y_i} = h_\sigma^{x_j} = 0$. The vector $\mathbf{h}_{\sigma\sigma}$ is given by

$$\begin{aligned}
\mathbf{h}_{\sigma\sigma}^T &= \left[h_{\sigma\sigma}^{y_1},\dots,h_{\sigma\sigma}^{y_{n(y)}}, h_{\sigma\sigma}^{x_1},\dots,h_{\sigma\sigma}^{x_{n(x)}}, \Delta_1^8,\dots,\Delta_{n(y)}^8\right],\\
\Delta_i^8 &= \sum_{s=1}^{n(x)} h_{x_s}^{y_i} h_{\sigma\sigma}^{x_s}\\
&\quad + \sum_{s=1}^{n(x)} h_\sigma^{x_s}\left(\sum_{r=1}^{n(x)} h_{x_s x_r}^{y_i} h_\sigma^{x_r} + \sum_{r=1}^{n(z)} h_{x_s z_r}^{y_i} \sum_{t=1}^{n(z)} \omega_{rt}\epsilon'_t + h_{x_s\sigma}^{y_i}\right)\\
&\quad + \sum_{s=1}^{n(z)} \left(\sum_{r=1}^{n(z)} \omega_{sr}\epsilon'_r\right)\left[\sum_{t=1}^{n(x)} h_{z_s x_t}^{y_i} h_\sigma^{x_t} + \sum_{t=1}^{n(z)} h_{z_s z_t}^{y_i} \sum_{u=1}^{n(z)} \omega_{tu}\epsilon'_u\right]\\
&\quad + \sum_{s=1}^{n(x)} h_{\sigma x_s}^{y_i} h_\sigma^{x_s} + \sum_{s=1}^{n(z)} h_{\sigma z_s}^{y_i} \sum_{r=1}^{n(z)} \omega_{sr}\epsilon'_r + h_{\sigma\sigma}^{y_i}.
\end{aligned}$$

Consider the expectation of the first term on the rhs of (2.81) and note that

1. the vector $g^i_{[\mathbf{y},\mathbf{x}',\mathbf{y}']}$ does not contain any stochastic variables,
2. in addition to $h_\sigma^{y_i} = 0$ and $h_\sigma^{x_i} = 0$ also $h_{\sigma z_s}^{y_i} = h_{\sigma x_s}^{y_i} = 0$,
3. $E(\epsilon'_i\epsilon'_j) = 0$ for all $i \neq j$ and $E(\epsilon'_i\epsilon'_i) = 1$.

Therefore, we get

$$
\begin{aligned}
&E\left\{\mathbf{h}_{\sigma\sigma}^{T} g^{i}_{[\mathbf{y},\mathbf{x}',\mathbf{y}']}\right\}\\
&= \left[h^{y_1}_{\sigma\sigma},\dots,h^{y_{n(y)}}_{\sigma\sigma},h^{x_1}_{\sigma\sigma},\dots,h^{x_{n(x)}}_{\sigma\sigma},\Delta^9_1,\dots,\Delta^9_{n(y)}\right] g^{i}_{[\mathbf{y},\mathbf{x}',\mathbf{y}']},\\
&\Delta^9_i = \sum_{s=1}^{n(x)} h^{y_i}_{x_s} h^{x_s}_{\sigma\sigma} + h^{y_i}_{\sigma\sigma} + \sum_{s=1}^{n(z)}\sum_{r=1}^{n(z)} h^{y_i}_{z_s z_r} \sum_{t=1}^{n(z)} \omega_{st}\omega_{rt}.
\end{aligned}
$$

By using a well known property of the trace operator, the expectation of the second term on the rhs of (2.81) equals[20]

$$
\operatorname{tr}\left\{ g^{i}_{[\mathbf{y}',\mathbf{z}'][\mathbf{y}',\mathbf{z}']} E\left[\Delta^7_1,\dots,\Delta^7_{n(y)},\Delta^6_1,\dots,\Delta^6_{n(z)}\right] \begin{bmatrix} \Delta^7_1 \\ \vdots \\ \Delta^7_{n(y)} \\ \Delta^6_1 \\ \vdots \\ \Delta^6_{n(z)} \end{bmatrix} \right\}.
$$

The expectation of the cross-products involved in this expression are readily evaluated to be

$$
\begin{aligned}
E\left[\Delta^7_i \Delta^7_j\right] &= \sum_{q=1}^{n(z)}\sum_{s=1}^{n(z)} h^{y_i}_{z_q} h^{y_j}_{z_s} \sum_{r=1}^{n(z)} \omega_{qr}\omega_{sr},\\
E\left[\Delta^7_i \Delta^6_j\right] &= \sum_{s=1}^{n(z)} h^{y_i}_{z_s} \sum_{r=1}^{n(z)} \omega_{sr}\omega_{jr},\\
E\left[\Delta^6_i \Delta^6_j\right] &= \sum_{r=1}^{n(z)} \omega_{ir}\omega_{jr}.
\end{aligned}
$$

Implementation. Our Gauss program `SolveQA` implements the computation of the Hesse matrices H^i in (2.74). It requires the coefficients of the linear part, the matrices Π, Ω, the Jacobian matrix of $\mathbf{g}$ stored in a matrix `gmat`, say, and the $n(x)+n(y)$ Hesse matrices of g^i as input. The latter must be gathered in a three-dimensional array `hcube`, say. The program returns two

[20] The second term on the rhs of (2.81), say a, is a scalar so that $a = \operatorname{tr}(a)$. Yet, for any two conformable matrices A and B, it holds that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

three-dimensional arrays: `xcube` contains the $n(x)$ Hesse matrices H^{x_i} and `ycube` stores the $n(y)$ Hesse matrices H^{y_i}.

Of course, there is other software available on the world wide web. SCHMITT-GROHÉ and URIBE (2004) provide Matlab programs that compute the matrices of the quadratic part in our equation (2.74). An advantage of their program is its ability to handle symbolic differentiation if you own the respective Matlab toolbox. Other programs that can handle quadratic approximations are Dynare[21] mainly developed by JUILLARD and Gensys written by SIMS.[22]

2.6 Applications

In this section we consider three applications of the methods presented in the previous sections. First, we solve the benchmark model introduced in Chapter 1, second, we consider a simplified version of the time-to-build model of KYDLAND and PRESCOTT (1982), and third, we develop a monetary model with nominal rigidities that give raise to what has been called the New Keynesian Phillips curve.

2.6.1 The Benchmark Model

In Example 1.5.1, we present the benchmark model, in which a representative agent chooses feed-back rules for consumption and labor supply that maximize his expected live time utility over an infinite time horizon. This section shows how we can obtain linear and quadratic approximations of these feed-back rules by using the methods introduced in Sections 2.2 through 2.5.

Linear and Quadratic Policy Functions. Our starting point is the system of stochastic difference equations which we obtained

[21] See the user's guide written by GRIFFOLI (2007).
[22] See KIM, KIM, SCHAUMBURG, and SIMS (2005) on this program.

in Section 1.5. We repeat these equations for your convenience:[23]

$$0 = c_t^{-\eta}(1-N_t)^{\theta(1-\eta)} - \lambda_t, \tag{2.82a}$$
$$0 = \theta c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)-1} - (1-\alpha)\lambda_t Z_t N_t^{-\alpha} k_t^{\alpha}, \tag{2.82b}$$
$$0 = a k_{t+1} - (1-\delta)k_t + c_t - Z_t N_t^{1-\alpha} k_t^{\alpha}, \tag{2.82c}$$
$$0 = \lambda_t - \beta a^{-\eta} E_t \lambda_{t+1}\left(1-\delta+\alpha Z_{t+1} N_{t+1}^{1-\alpha} k_{t+1}^{\alpha-1}\right). \tag{2.82d}$$

Equation (2.82a) states that the shadow price of an additional unit of capital, λ_t, must equal the agent's marginal utility of consumption. Condition (2.82b) equates the marginal rate of substitution between consumption and leisure with the marginal product of labor. Equation (2.82c) is the economy's resource constraint. According to equation (2.82d) the marginal utility of consumption must equal the discounted expected utility value of the return on investment in the future stock of capital. We complete the model by specifying the law of motion for the natural log of the productivity shock $z_t := \ln Z_t$:

$$z_t = \varrho z_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2). \tag{2.82e}$$

In Section 1.5 we explain the choice of the model's parameters α, β, δ, η, and θ. With these values at hand, we can compute the stationary solution (k, λ, c, N) from equations (1.46). The vectors $\mathbf{x}_t$, $\mathbf{u}_t$, and $\boldsymbol{\lambda}_t$ from equations (2.47) are then given by $\mathbf{x}_t \equiv k_t - k$, $\boldsymbol{\lambda}_t \equiv \lambda_t - \lambda$, $\mathbf{u}_t := [c_t - c, N_t - N]'$, and $\mathbf{z}_t \equiv \ln Z_t$. In our Fortran program `Benchmark.for` we use numeric differentiation of (2.82) at (k, λ, c, N) to obtain the Jacobian matrix `gmat`. From this matrix we derive the coefficients of the matrices C_u, $C_{x\lambda}$, C_z, $D_{x\lambda}$, $F_{x\lambda}$, D_u, F_u, D_z, and F_z, that appear in (2.47). A call to `SolveLA` returns the coefficients of the linear approximate policy functions. To obtain the coefficients of the quadratic part, we differentiate each equation of (2.82) twice using `CDHesse`. This

[23] The symbols have the following meaning: C_t is consumption, N_t are working hours, K_t is the stock of capital, Λ_t is the shadow price of an additional unit of capital and Z_t is the level of total factor productivity. Except for $\lambda_t := A_t^{\eta}\Lambda_t$, the other lower case variables are scaled by the level of labor-augmenting technical progress A_t , that is, $c_t := C_t/A_t$ and $k_t := K_t/A_t$.

provides the three dimensional array `hcube` that is an input to `SolveQA`. Thus, it requires four steps to compute the solutions:

Step 1: solve for (k, λ, c, n),
Step 2: write a procedure that receives the vector of 10 elements $(k, \lambda, c, n, z, k', \lambda', c', n', z')$ and that returns the lhs of (2.82),
Step 3: compute `gmat` and `hcube` by using `CDJac` and `CDHesse`, respectively,
Step 4: set up the matrices required by `SolveLA` and `SolveQA`.

Linear Quadratic Algorithm. At first sight, it seems that the law of motion of the productivity shock z_t in equation (2.82e) is the only linear equation of the benchmark model. Yet, if we use investment expenditures

$$i_t = Z_t N_t^{1-\alpha} k_t^{\alpha} - c_t \tag{2.83}$$

instead of consumption c_t, equation (2.82c) can be written as:

$$k_{t+1} = \frac{1}{a} i_t + \frac{1-\delta}{a} k_t, \tag{2.84}$$

which is linear in k_{t+1}, k_t, and i_t. Let $\mathbf{x}_t := [1, k_t, z_t]'$ denote the vector of states and $\mathbf{u}_t := [i_t, N_t]'$ the vector of controls. Then, for our model, the transition equation (2.31) is given by:

$$\mathbf{x}_{t+1} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & (1-\delta)/a & 0 \\ 0 & 0 & \varrho \end{bmatrix}}_{A} \mathbf{x}_t + \underbrace{\begin{bmatrix} 0 & 0 \\ 1/a & 0 \\ 0 & 0 \end{bmatrix}}_{B} \mathbf{u}_t + \begin{bmatrix} 0 \\ 0 \\ \epsilon_t \end{bmatrix}. \tag{2.85}$$

The remaining non-linearities are handled by the algorithm. The current period return function in the scaled variables is given by:

$$g(\mathbf{x}, \mathbf{u}) := \frac{1}{1-\eta} \left(e^{z_t} N_t^{1-\alpha} k_t^{\alpha} - i_t \right)^{1-\eta} (1 - N_t)^{\theta(1-\eta)}.$$

You must write a subroutine, say `GProc`, that takes the vector `ybar`$=[1, k, z, i, N]'$ as input and returns the value of g at this point.

There is a final issue that concerns the appropriate discount factor. The value function v that solves the Bellman equation

$$\begin{aligned} v(k,z) = \max_{k',N} \quad & u\left(e^z N^{1-\alpha} k^\alpha + (1-\delta)k - ak', 1-N\right) \\ & + \tilde{\beta} E\left[v(k',z')|z\right] \end{aligned}$$

is a function in the scaled variables. It is, thus, inappropriate to use β which pertains to the original variables. $\tilde{\beta}$ is found by observing that equations (2.82) solve the following scaled problem:

$$\begin{aligned} \max_{c_0,N_0} \quad & \sum_{t=0}^{\infty} \tilde{\beta}^t \Bigg\{ \Bigg[\frac{c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ & + \lambda_t \left(Z_t N_t^{1-\alpha} k_t^\alpha + (1-\delta)k_t - c_t - ak_{t+1}\right) \Bigg] \Bigg\}, \\ & \tilde{\beta} := \beta a^{1-\eta}. \end{aligned} \tag{2.86}$$

Other Variables of Interest. Both the program `SolveLA` and `SolveQA` provide approximations of the policy functions for k_{t+1}, c_t, and N_t. From these we obtain the solution for output y_t, investment i_t, and the real wage w_t, respectively, via

$$y_t = Z_t N_t^{1-\alpha} k_t^\alpha, \tag{2.87a}$$

$$i_t = y_t - c_t, \tag{2.87b}$$

$$w_t = (1-\alpha) Z_t N_t^{-\alpha} k_t^\alpha. \tag{2.87c}$$

The program `SolveLQA` provides linear approximate solutions for i_t and N_t from which we derive c_t via equation (2.87b). Given c_t the resource constraint (2.82c) yields the solution for k_{t+1}.

Time series for output y_t, consumption c_t, investment i_t, hours N_t, and the real wage w_t are derived by iterations that start at the stationary solution $k_1 = k$. We use a random number generator to obtain a sequence of innovations $\{\epsilon_t\}_{t=1}^T$. The sequence of capital stocks and the sequence of productivity shocks follow from

$$\left.\begin{aligned} k_{t+1} &= \hat{h}^k(k_t, z_t), \\ z_{t+1} &= \varrho z_t + \epsilon_{t+1}, \end{aligned}\right\} t = 1, 2, \ldots, T-1,$$

where $\hat{h}^k(\cdot)$ denotes the approximate policy function for the next-period stock of capital. Once we have computed the sequences $\{k_t\}_{t=1}^T$ and $\{z_t\}_{t=1}^T$, the sequences for the other variables of the model are obtained from the respective approximate policy functions and from (2.87).

Results. Table 2.2 summarizes the results of our simulations carried out with the Fortran program `Benchmark.for`. We used the parameter values from Table 1.1. The length T of our artificial time series for output, investment, consumption, working hours, and the real wage is 60 quarters.[24] The second moments displayed in Table 2.2 refer to HP-filtered percentage deviations from a variable's stationary solution.[25] They are averages over 500 simulations. We use the same sequence of shocks for all three solution methods to prevent random differences in the results.

The first message from Table 2.2 is that except for the small difference in the standard deviation of investment of 0.01 between the linear and the linear quadratic solution there are virtually no differences in the second moments across our three different methods. There are, however, differences in accuracy. As explained in Section 1.6.2, we use two measures of accuracy: the residuals of the Euler equation (2.82d) and the DM-statistic.

The Euler equation residuals are computed over a grid of 400 equally spaced points over the square $[\underline{k}; \overline{k}] \times [\underline{z}; \overline{z}]$. We choose $\underline{z} = \ln 0.95$ and $\overline{z} = \ln 1.05$ because in more than ninety percent of our simulations z_t remains in this interval. The largest interval for the stock of capital that we consider is $\mathscr{K} = [0.8; 1.2]k$, where k is the stationary solution. Yet, even the much smaller interval $[0.9; 1.1]k$ covers all simulated sequences of the capital stock. We compute the Euler equation residual as the rate by which consumption had to be raised over $\hat{h}^c(k, z)$ so that the lhs of equation (2.82d) matches its rhs. The numbers displayed in Table 2.2 are the maximum absolute values over the square indicated in the left-most column of the table.

[24] See Section 1.5 on the issues of parameter choice and model evaluation.
[25] See Section 12.4 on the HP-Filter.

Table 2.2

	Linear			Linear Quadratic			Quadratic		
				Second Moments					
Variable	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	1.44	1.00	0.64	1.44	1.00	0.64	1.44	1.00	0.64
Investment	6.11	1.00	0.64	6.12	1.00	0.64	6.11	1.00	0.64
Consumption	0.56	0.99	0.66	0.56	0.99	0.66	0.56	0.99	0.66
Hours	0.77	1.00	0.64	0.77	1.00	0.64	0.77	1.00	0.64
Real Wage	0.67	0.99	0.65	0.67	0.99	0.65	0.67	0.99	0.65
				Euler Equation Residuals					
$[0.90; 1.10]k$	1.835E-4			7.656E-4			1.456E-5		
$[0.85; 1.15]k$	3.478E-4			9.322E-4			4.085E-5		
$[0.80; 1.20]k$	5.670E-4			1.100E-3			8.845E-5		
				DM-Statistic					
<3.816	2.0			1.3			2.7		
>21.920	3.4			8.9			3.0		

Notes: s_x:=standard deviation of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x. All second moments refer to HP-filtered percentage deviations from a variable's stationary solution. Euler equation residuals are computed as maximum absolute value over a grid of 400 equally spaced points on the square $\mathscr{K} \times [\ln 0.95; \ln 1.05]$, where $\mathscr{K}$ is defined in the respective row of the left-most column. The 2.5 and the 97.5 percent critical values of the $\chi^2(11)$-distribution are displayed in the last two lines of the first column. The table entries refer to the percentage fraction out of 1,000 simulations where the DM-statistic is below (above) its respective critical value.

First, note that all residuals are quite small. Even in the worst case, the required change of consumption is merely 0.11 percent. Second, and as expected from a local method, accuracy diminishes with the distance from the stationary solution. For instance, consider the linear policy function. The Euler equation residual over $[0.85; 1.15]k$ ($[0.8; 1.2]k$) is almost two times (three times) larger than the maximum residual over $[0.9; 1.1]k$. Third, the Euler equation residuals of the linear quadratic approach are worse than those of the linear approach. For the former, the maximum absolute Euler equation residual over $[0.9; 1.1]k$ is more than four times larger than the Euler equation residual of the linear solution method. Fourth, although the quadratic policy function delivers a more accurate solution than the linear policy function, the difference between the respective Euler equation residuals becomes smaller as one moves further away from the stationary solution: Over $[0.9; 1.1]k$ the Euler equation residual of the linear solution is more than twelve times larger than the Euler equation residual of the quadratic solution. Yet over $[0.8; 1.2]k$ it is only six times larger. Fifth, there are several possible ways to compute the Euler equation residuals. For instance, since both the linear and the quadratic perturbation method deliver a policy function for λ, we could use this function in the computation. We, however, used the policy functions for consumption and hours and inferred λ from equation (2.82a), since the linear quadratic approach delivers only policy functions for investment and hours. The difference is considerable: When we use the linear approximate policy function for λ we find a maximum Euler equation residual over $[0.9; 1.1]k$ that is 26 times larger than that displayed in Table 2.2.

As explained in Section 1.6.2 (and more formally in Section 12.3), the DM-statistic aims to detect systematic forecast errors with respect to the rhs of the Euler equation (2.82d). For this purpose, we simulate the model and compute the ex-post forecast error

$$e_t := \beta a^{-\eta}\lambda_{t+1}\left(1 - \delta + \alpha Z_{t+1} N_{t+1}^{1-\alpha} k_{t+1}^{\alpha-1}\right) - \lambda_t,$$

where λ_t is given by equation (2.82a). We use simulated time series with many periods so that the asymptotic properties of the test statistic will apply. The simulations always start at the stationary solution. To prevent the influence of the model's transitional dynamics on our results, we discard a small fraction of the initial observations. In effect, we use 3,000 points. We regress e_t on a constant, five lags of consumption, and five lags of the productivity shock and compute the Wald-statistic (which is the DM-statistic in this context) of the null that all coefficients from this regression are equal to zero. We use White's (1980) heteroscedasticity robust covariance estimator. Under the null the Wald-statistic has a χ^2-distribution with 11 degrees of freedom. We run 1,000 tests and computed the fraction of the DM-statistic below (above) the 2.5 (97.5) percent critical value (displayed in the first column of Table 2.2). If systematic errors are not present, about 2.5 percent of our simulations should yield test statistics below (above) the respective critical values. Both, the linear and the quadratic policy functions provide satisfactory results. Yet, the linear policy functions obtained from the linear-quadratic approach are less good. The null is far more often rejected than can be expected, namely in almost 9 percent of our simulations.

Finally, note that the second moments as well as the DM-statistic depend on the random numbers used for the productivity shock z_t. Thus, when you repeat our calculations, you will find at least small differences to our results.

2.6.2 Time to Build

Gestation Period. In the benchmark model investment projects require one quarter to complete. In their classic article Kydland and Prescott (1982) use a more realistic gestation period. Based on published studies of investment projects they assume that it takes four quarters for an investment project to be finished. The investment costs are spread out evenly over this period. Yet, the business cycle in this extended model is similar to the business cycle in their benchmark model with a one quarter lag. We introduce the time-to-build assumption into the benchmark model of the previous section. Our results confirm their findings. Nevertheless, we think this venture is worth the while, since it nicely demonstrates the ease of applying the linear quadratic solution algorithm to a rather tricky dynamic model.

The model that we consider uses the same specification of the household's preferences and production technology as the model in the previous section. The timing of investment expenditures differs from this model in the following way. In each quarter t the representative household launches a new investment project. After four quarters this project is finished and adds to the capital stock. The investment costs are spread out over the entire gestation period. More formally, let S_{it}, $i = 1, 2, 3, 4$ denote an investment project that is finished after i periods and that requires the household to pay the fraction ω_i of its total costs. At any period, there are four unfinished projects so that total investment expenditures I_t amount to

$$I_t = \sum_{i=1}^{4} \omega_i S_{it}, \quad \sum_{i=1}^{4} \omega_i = 1. \tag{2.88}$$

Obviously, the S_{it} are related to each other in the following way:

$$\begin{aligned} S_{1t+1} &= S_{2t}, \\ S_{2t+1} &= S_{3t}, \\ S_{3t+1} &= S_{4t}, \end{aligned} \tag{2.89}$$

and the capital stock evolves according to

$$K_{t+1} = (1-\delta)K_t + S_{1t}. \tag{2.90}$$

First-Order Conditions. Since the model exhibits growth, we transform it to a stationary problem. As in Section 2.6.1 we put $c_t := C_t/A_t$, $i_t := I_t/A_t$, $k_t := K_t/A_t$, $\lambda_t := \Lambda_t A_t^{\eta}$, $s_{it} = S_{it}/A_t$, and $\tilde{\beta} := \beta a^{1-\eta}$. In this model, the vector of states is $\mathbf{x}_t = [1, k_t, s_{1t}, s_{2t}, s_{3t}, \ln Z_t]'$ and the vector of controls is $\mathbf{u} = [s_{4t}, N_t]'$. From (2.89) and (2.90) we derive the following law of motion of the stationary variables:

$$\mathbf{x}_{t+1} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1-\delta}{a} & \frac{1}{a} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{a} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{a} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho \end{bmatrix}}_{:=A} \mathbf{x}_t + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{1}{a} & 0 \\ 0 & 0 \end{bmatrix}}_{=:B} \mathbf{u}_t + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \epsilon_t \end{bmatrix}. \tag{2.91}$$

The remaining task is to compute the stationary equilibrium. Consider the Lagrangean of the stationary problem:

$$\begin{aligned}\mathscr{L} = E_0 \sum_{t=0}^{\infty} \tilde{\beta}^t \Bigg\{ & \frac{c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ & + \lambda_t \big(Z_t N_t^{1-\alpha} k_t^{\alpha} - \sum_{i=1}^{4} \omega_i s_{it} - c_t \big) \\ & + \gamma_t ((1-\delta)k_t + s_{1t} - a k_{t+1}) \Bigg\},\end{aligned}$$

where γ_t is the Lagrange multiplier of the transformed constraint (2.90). Differentiating this expression with respect to c_t, N_t, s_{4t} and k_{t+4} provides the following conditions:[26]

$$\lambda_t = c_t^{-\eta}(1-N_t)^{\theta(1-\eta)}, \tag{2.92a}$$

$$\frac{\theta c_t}{1-N_t} = (1-\alpha) Z_t N_t^{-\alpha} k_t^{\alpha}, \tag{2.92b}$$

$$\begin{aligned}0 = E_t \big[& -\omega_4 \lambda_t - (\tilde{\beta}/a)\omega_3 \lambda_{t+1} - (\tilde{\beta}/a)^2 \omega_2 \lambda_{t+2} \\ & - (\tilde{\beta}/a)^3 \omega_1 \lambda_{t+3} + (\tilde{\beta}/a)^3 \gamma_{t+3} \big],\end{aligned} \tag{2.92c}$$

$$\begin{aligned}0 = E_t \big[& -(\tilde{\beta}/a)^3 \gamma_{t+3} + (\tilde{\beta}/a)^4 (1-\delta)\gamma_{t+4} \\ & + (\tilde{\beta}/a)^4 \lambda_{t+4} \alpha Z_{t+4} N_{t+4}^{1-\alpha} k_{t+4}^{\alpha-1} \big].\end{aligned} \tag{2.92d}$$

The first and the second condition are standard and need no comment. The third and the fourth condition imply the following Euler equation in the shadow price of capital:

$$\begin{aligned}0 = E_t \Big\{ & \omega_4 [(\tilde{\beta}/a)(1-\delta)\lambda_{t+1} - \lambda_t] \\ & + \omega_3 (\tilde{\beta}/a)[(\tilde{\beta}/a)(1-\delta)\lambda_{t+2} - \lambda_{t+1}] \\ & + \omega_2 (\tilde{\beta}/a)^2 [(\tilde{\beta}/a)(1-\delta)\lambda_{t+3} - \lambda_{t+2}] \\ & + \omega_1 (\tilde{\beta}/a)^3 [(\tilde{\beta}/a)(1-\delta)\lambda_{t+4} - \lambda_{t+3}] \\ & + (\tilde{\beta}/a)^4 \alpha \lambda_{t+4} Z_{t+4} N_{t+4}^{1-\alpha} k_{t+4}^{\alpha-1} \Big\}.\end{aligned}$$

[26] To keep track of the various terms that involve s_{4t} and k_{t+4}, it is helpful to write out the sum for $t = 0, 1, 2, 3, 4$.

Stationary Equilibrium. On a balanced growth path, where $Z_t = 1$ and $\lambda_t = \lambda_{t+1}$ for all t, this expression reduces to

$$\frac{y}{k} = \frac{a - \tilde{\beta}(1-\delta)}{\alpha\tilde{\beta}} \left[\omega_1 + (a/\tilde{\beta})\omega_2 + (a/\tilde{\beta})^2\omega_3 + (a/\tilde{\beta})^3\omega_4\right]. \tag{2.93}$$

Given a, β, δ, and η, we can solve this equation for the output-capital ratio y/k. From $(1-\delta)k + s_1 = ak$ we find $s_1 = (a+\delta-1)k$, the stationary level of new investment projects started in each period. Total investment per unit of capital is then given by

$$\frac{i}{k} = \frac{1}{k}\sum_{i=1}^{4}\omega_i s_i = (a+\delta-1)\sum_{i=1}^{4} a^{i-1}\omega_i.$$

Using this, we can solve for

$$\frac{c}{k} = \frac{y}{k} - \frac{i}{k}.$$

Since $y/c = (y/k)/(c/k)$, we can finally solve the stationary version of (2.92b) for N. This solution in turn provides $k = N(y/k)^{1/(\alpha-1)}$, which allows us to solve for i and c. The final step is to write a procedure that returns the current period utility as a function of $\mathbf{x}$ and $\mathbf{u}$. The latter is given by:

$$g(\mathbf{x}, \mathbf{u}) := \frac{1}{1-\eta}\left(e^{\ln Z_t} N_t^{1-\alpha} k_t^{\alpha} - \sum_{i=1}^{4} s_{it}\right)^{1-\eta} (1-N_t)^{\theta(1-\eta)}.$$

Results. The Gauss program `ToB.g` implements the solution. We use the parameter values from Table 1.1 and assume $\omega_i = 0.25, i = 1, \dots, 4$. Table 2.3 displays the averages of 500 time series moments computed from the simulated model. We used the same random numbers in both the simulations of the benchmark model and the simulations of the time-to-build model. Thus, the differences revealed in Table 2.3 are systematic and not random.

In the time-to-build economy output, investment, and hours are a little less volatile than in the benchmark economy. The intuition behind this result is straightforward. When a positive technological shock hits the benchmark economy the household takes

Table 2.3

	Benchmark			Time to Build		
Variable	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	1.45	1.00	0.63	1.37	1.00	0.63
Investment	6.31	0.99	0.63	5.85	0.99	0.65
Consumption	0.57	0.99	0.65	0.58	0.97	0.56
Hours	0.78	1.00	0.63	0.71	0.98	0.65
Real Wage	0.68	0.99	0.64	0.68	0.98	0.58

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x.

the chance, works more at the higher real wage and transfers part of the increased income via capital accumulation into future periods. Since the shock is highly autocorrelated, the household can profit from the still above average marginal product of capital in the next quarter. Yet in the time-to-build economy intertemporal substitution is not that easy. Income spent on additional investment projects will not pay out in terms of more capital income until the fourth quarter after the shock. However, at this time a substantial part of the shock has faded. This reduces the incentive to invest and, therefore, the incentive to work more.

Lawrence Christiano and Richard Todd (1996) embed the time-to-build structure in a model where labor augmenting technical progress follows a random walk. They use a different parameterization of the weights ω_i. Their argument is that investment projects typically begin with a lengthy planning phase. The overwhelming part of the project's costs are spent in the construction phase. As a consequence, they set $\omega_1 = 0.01$ and $\omega_2 = \omega_3 = \omega_4 = 0.33$. This model is able to account for the positive autocorrelation in output growth, whereas the Kydland and Prescott (1982) parameterization of the same model – $\omega_i = 0.25$, $i = 1, \dots, 4$ – is not able to replicate this empirical finding. However, the random walk assumption does not lent itself to the linear quadratic approach, and, therefore we will not pursue this matter any further.

2.6.3 New Keynesian Phillips Curve

Money in General Equilibrium. So far we have restricted our attention to non-monetary economies. In this subsection we focus on the interaction of real and monetary shocks to explain the business cycle.

Introducing money into a dynamic general equilibrium model is not an easy task. As a store of value money is dominated by other interest bearing assets like corporate and government bonds or stocks, and in the basically one-good Ramsey model there is no true need for a means of exchange. So how do we guarantee a positive value of pure fiat outside money in equilibrium?

Monetary theory has pursed three approaches (see, e.g., WALSH (2003)). The first device is to assume that money yields direct utility, the second strand of the literature imposes transaction costs, and the third way is to guarantee an exclusive role for money as a store of value. We will pursue the second approach in what follows and assume transaction costs to be proportional to the volume of trade. Moreover, a larger stock of real money balances relative to the volume of trade reduces transaction costs (see LEEPER and SIMS (1994)). Different from other approaches, as, e.g., the cash-in-advance assumption, our particular specification implies the neutrality of monetary shocks in the log-linear model solution. This allows us to focus on other deviations from the standard model that are required to explain why money has short-run real effects.

The most prominent explanation for the real effects of money that has been pursued in the recent literature are nominal rigidities that arise from sticky wages and/or prices.[27] Among the var-

[27] A non-exhaustive list of models of nominal rigidities includes BERGIN and FEENSTRA (2000), CHARI, KEHOE, and MCGRATTAN (2000), CHO and COOLEY (1995), COOLEY and HANSEN (1995, 1998), CHRISTIANO, EICHENBAUM, and EVANS (1997), HAIRAULT and PORTIER (1995).

ious models probably the CALVO (1983) model has gained the most widespread attention. For this reason we use the discrete time version of his assumption on price setting to introduce nominal frictions into the monetary economy that we consider in the following paragraphs.

The CALVO (1983) hypothesis provides a first-order condition for the optimal relative price of a monopolistically competitive firm that is able to adjust its price optimally whereas a fraction of other firms is not permitted to do so. The log-linear version of this condition (see equation (A.4.11e) in Appendix 4) relates the current inflation rate to the expected inflation rate and a measure of labor market tension. It thus provides solid microfoundations for the well-known Phillips curve that appears in many textbooks. This curve plays the role of a short-run aggregate supply function and relates inflation to expected inflation and cyclical unemployment.[28] In the CALVO (1983) model the deviation of marginal costs from their average level measures labor market tension. Since this equation resembles the traditional Phillips curve it is sometimes referred to as the New Keynesian Phillips curve.

The Household Sector. The representative household has the usual instantaneous utility function u defined over consumption C_t and leisure $1 - N_t$, where N_t are working hours:

$$u(C_t, 1 - N_t) := \frac{C_t^{1-\eta}(1 - N_t)^{\theta(1-\eta)}}{1 - \eta}. \tag{2.94}$$

The parameters of this function are non-negative and satisfy $\eta > \theta/(1 + \theta)$. The household receives wages, rental income from capital services, dividends D_t and a lump-sum transfer from the government T_t. We use P_t to denote the aggregate price level. The wage rate in terms of money is W_t and the rental rate in terms of consumption goods is r_t. The household allocates its income net of transaction costs TC_t to consumption, additional holdings of physical capital K_t and real money balances M_t/P_t, where M_t is the beginning-of-period stock of money. This produces the following budget constraint:

[28] See, e.g., MANKIW (2000), pp. 364.

$$\frac{M_{t+1}-M_t}{P_t} + K_{t+1} - (1-\delta)K_t \le \frac{W_t}{P_t}N_t + r_t K_t + D_t \\ + T_t - TC_t - C_t. \tag{2.95}$$

Transactions costs are given by the following function

$$TC_t = \gamma \left(\frac{C_t}{M_{t+1}/P_t} \right)^\kappa C_t, \quad \gamma,\ \kappa > 0. \tag{2.96}$$

Importantly, the assumption that the costs TC_t depend upon the ratio of consumption to real end-of-period money holdings M_{t+1}/P_t is responsible for the neutrality of money in our model. The household maximizes the expected discounted stream of future utility

$$E_0 \sum_{t=0}^{\infty} \beta^t u(C_t, 1-N_t)$$

subject to (2.95) and (2.96).

Money Supply. The government sector finances the transfers to the household sector from money creation. Thus,

$$T_t = \frac{M_{t+1}-M_t}{P_t}. \tag{2.97}$$

We assume that the monetary authority is not able to monitor the growth rate of money supply perfectly. In particular, we posit the following stochastic process for the growth factor of money supply $\mu_t := M_{t+1}/M_t$:

$$\hat{\mu}_{t+1} = \rho^\mu \hat{\mu}_t + \epsilon_t^\mu, \quad \hat{\mu}_t := \ln \mu_t - \ln \mu,\ \epsilon_t^\mu \sim N(0, \sigma^\mu). \tag{2.98}$$

In the stationary equilibrium money grows at the rate $\mu - 1$.

Price Setting. To motivate price setting by individual firms we assume that there is a final goods sector that assembles the output of a large number J_t of intermediary producers to the single good Y_t according to

$$Y_t = \left[J_t^{-1/\epsilon} \sum_{j=1}^{J_t} Y_{jt}^{(\epsilon-1)/\epsilon} \right]^{\epsilon/(\epsilon-1)}, \quad \epsilon > 1. \tag{2.99}$$

The money price of intermediary product j is P_{jt} and final output sells at the price P_t. The representative firm in the final sector takes all prices as given. Maximizing its profits $P_t Y_t - \sum_{j=1}^{J_t} P_{jt} Y_j$ subject to (2.99) produces the following demand for good j:

$$Y_{jt} = \left(\frac{P_{jt}}{P_t}\right)^{-\epsilon} \frac{Y_t}{J_t}. \tag{2.100}$$

Accordingly, ϵ is the price elasticity of demand for good j. It is easy to demonstrate that the final goods producers earn no profits if the aggregate price index P_t is given by the following function:

$$P_t = \left[\frac{1}{J_t}\sum_{j=1}^{J_t} P_{jt}^{1-\epsilon}\right]^{1/(1-\epsilon)}. \tag{2.101}$$

An intermediary producer j combines labor N_{jt} and capital services K_{jt} according to the following production function:

$$Y_{jt} = Z_t(A_t N_{jt})^{1-\alpha} K_{jt}^{\alpha} - F, \quad \alpha \in (0,1), F > 0. \tag{2.102}$$

F is a fixed cost in terms of forgone output. We will use F to determine the number of firms on a balanced growth path from the zero profit condition. As in all our other models A_t is an exogenous, deterministic process for labor augmenting technical progress,

$$A_{t+1} = aA_t, \quad a \geq 1,$$

and Z_t is a stationary, stochastic process for total factor productivity that follows

$$\hat{Z}_{t+1} = \rho^Z \hat{Z}_t + \epsilon_t^Z, \quad \hat{Z}_t = \ln Z_t,\ \epsilon_t^Z \sim N(0, \sigma^Z).$$

Note that α, F, A_t, and Z_t are common to all intermediary producers, who also face the same price elasticity ϵ.

From now on we must distinguish between two types of firms, which we label A and N, respectively. Type A firms are allowed to set their price P_{At} optimally, whereas type N firms are not. To prevent their relative price P_{Nt}/P_t from falling short of the aggregate price level, type N firms are permitted to increase their price according to the average inflation factor π. This is the inflation factor on a balanced growth path without any uncertainty. Thus

$$P_{Nt} = \pi P_{Nt-1}. \tag{2.103}$$

To which type an individual firm j belongs is random. At each period $(1-\varphi)J_t$ of firms receive the signal to choose their optimal relative price P_{At}/P_t. The fraction $\varphi \in [0,1]$ must apply the rule (2.103). Those firms that are free to adjust their price solve the following problem:

$$\begin{aligned} \max_{P_{At}} \quad & E_t \sum_{\tau=t}^{\infty} \varphi^{\tau-t} \varrho_\tau \left[\left(\frac{\pi^{\tau-t} P_{At}}{P_\tau} \right) Y_{A\tau} - g_\tau (Y_{A\tau} + F) \right] \\ \text{s.t.} \quad & Y_{A\tau} = \left(\frac{\pi^{\tau-t} P_{At}}{P_\tau} \right)^{-\epsilon} \frac{Y_\tau}{J_\tau}. \end{aligned} \tag{2.104}$$

The sum to the right of the expectations operator E_t is the discounted flow of real profits earned until the firm will be able to reset its price optimally again. Real profits are given by the value of sales in units of the final good $[(\pi^{\tau-t} P_{At})/P_\tau]Y_\tau$ minus production cost $g_\tau(Y_{A\tau}+F)$, where g_τ are the firm's variable unit costs.[29] The term $\varphi^{\tau-t}$ captures the probability that in period τ the firm is still a type N producer. ϱ_τ is the discount factor for time τ profits. We show in Section 6.3.4 that this factor is related to the household's discount factor β and marginal utility of wealth Λ_τ by the following formula:

$$\varrho_\tau = \beta^{\tau-t} \frac{\Lambda_\tau}{\Lambda_t}. \tag{2.105}$$

Intermediary producers distribute their profits to the household sector. Thus,

[29] We show in Appendix 4 that g_τ also equals the firm's marginal costs. Note further that equation (2.102) implies that the firm must produce the amount $Y_{jt} + F$ in order to sell Y_{jt}.

$$D_t := \sum_{j=1}^{J_t} \frac{P_{jt}}{P_t} Y_{jt} - \frac{W_t}{P_t} N_{jt} - r_t K_{jt}. \tag{2.106}$$

This equation closes the model. To streamline the presentation we restrict ourselves to the properties of the stationary equilibrium and the simulation results. Appendix 4 provides the mathematical details of the analysis and the loglinear model used for the simulation.

Stationary Equilibrium. The model of this section depicts a growing economy. For this reason we must scale the variables so that they are stationary on a balanced growth path. As previously, we use the following definitions: $c_t := C_t/A_t$, $y_t := Y_t/A_t$, $k_t := K_t/A_t$, $\lambda_t := \Lambda_t A_t^{\eta}$. In addition, we define the inflation factor $\pi_t := P_t/P_{t-1}$ and real end-of-period money balances $m_t := M_{t+1}/(A_t P_t)$. The stationary equilibrium of the deterministic model has the following properties:

1. The productivity shock and the money supply shock equal their respective means $Z_t = Z \equiv 1$ and $\mu_t = \mu$ for all t.
2. Inflation is constant: $\pi = \frac{P_t}{P_{t-1}}$ for all t.
3. All (scaled) variables are constant.
4. All firms in the intermediary sector earn zero profits.

There are two immediate consequences of these assumptions. First, inflation is directly proportional to the growth rate of money supply $\mu - 1$:[30]

$$\pi = \frac{\mu}{a}.$$

Second, the optimal relative price of type A firms satisfies

$$\frac{P_A}{P} = \frac{\epsilon}{\epsilon - 1} g,$$

i.e., it is determined as a markup on the firm's marginal costs g. Furthermore, the formula for the price index given in equation (A.4.5) implies $P_A = P$ so that $g = (\epsilon - 1)/\epsilon$ and $P_N = P$. Since

[30] See equation (A.4.2c) for $m_t = m_{t+1}$.

all firms charge the same price, the market share of each producer is Y/J. Therefore, working hours and capital services are equal across firms, $N_j = N/J$, and $K_j = K/J$, and profits amount to

$$D_j = \frac{Y}{J} - g\left(\frac{Y}{J} + F\right).$$

Imposing $D_j = 0$ for all j and using $Y/J = (AN/J)^{1-\alpha}(K/J)^{\alpha} - (F/J)$ yields

$$j := \frac{J_t}{A_t} = \frac{N^{1-\alpha}k^{\alpha}}{\epsilon F}.$$

Thus, to keep profits at zero, the number of firms must increase at the rate $a - 1$ on the balanced growth path.[31] The production function (2.102) thus implies

$$y = \frac{\epsilon - 1}{\epsilon} N^{1-\alpha}k^{\alpha}.$$

Using this in the first-order condition for cost minimization with respect to capital services (see equation (A.4.3b)) implies

$$r = \alpha(y/k).$$

Eliminating r from the Euler equation for capital delivers the well known relation between the output-capital ratio and the household's discount factor β:

$$\frac{y}{k} = \frac{a^{\eta} - \beta(1-\delta)}{\alpha\beta}. \tag{2.107a}$$

This result allows us to solve for the consumption-output ratio via the economy's resource constraint (see (A.4.9)):

$$\frac{c}{y} = \left(1 + \frac{1 - a - \delta}{y/k}\right)\left[1 + \gamma\left(\frac{C}{\mu(M/P)}\right)^{\kappa}\right]^{-1}.$$

[31] Alternatively, we could have assumed that fixed costs are given by $A_t F$ so that the number of firms does not grow without bounds.

The stationary version of the Euler condition for money balances (see equation (A.4.2e)) delivers:

$$\frac{\beta a^{1-\eta}}{\mu} = 1 - \kappa\gamma \left(\frac{C}{\mu(M/P)} \right)^{1+\kappa}. \tag{2.107b}$$

We need a final equation to determine the stationary level of working hours. Using the results obtained so far we derive this relation from the household's first-order condition with respect to labor supply (see equation (A.4.2b)):

$$\frac{N}{1-N} = \frac{1-\alpha}{\theta} \left(1 + \frac{1-a-\delta}{y/k} \right)^{-1} h(c/x), \tag{2.107c}$$

$$h(c/x) := \frac{1+\gamma(c/x)^{\kappa}}{1+\gamma(1+\kappa)(c/x)^{\kappa}}, \; \frac{c}{x} := \frac{PC}{\mu M}.$$

It is obvious from equation (2.107a) that the output-capital ratio and therefore also the capital-labor ratio k/N and labor productivity y/N are independent of the money growth rate. As can be seen from (2.107b), the velocity of end-of-period money balances $c/x \equiv C/(\mu(M/P))$ is an increasing function of the money growth rate. In the benchmark model of Section 2.6.1 working hours are determined by the first two terms on the rhs of (2.107c). The presence of money adds the factor $h(c/x)$. It is easy to show that $h(c/x)$ is an decreasing function of the velocity of money (c/x). Since $N/(1-N)$ increases with N, steady-state working hours are a decreasing function of the money growth rate.

Calibration. We do not need to assign new values to the standard parameters of the model. The steady state relations presented in the previous paragraph show that the usual procedure to calibrate β, α, a, and δ is still valid. We will also use the empirical value of N to infer θ from (2.107c). This implies a slightly smaller value of θ as compared to the value of this parameter in the benchmark model. Nothing is really affected from this choice.

Unfortunately, there is no easy way to determine the parameters of the productivity shock, since there is no simple aggregate production function that we could use to identify Z_t. The problem becomes apparent from the following equation, which we derive in Appendix 4:

$$\hat{y}_t = \vartheta(1-\alpha)\hat{N}_t + \vartheta\alpha\hat{k}_t + \vartheta\hat{Z}_t(1-\vartheta)\hat{j}_t, \quad \vartheta = \frac{\epsilon}{\epsilon - 1}. \tag{2.108}$$

This equation is the model's analog to the log-linear aggregate production function in the benchmark model given by

$$\hat{y}_t = (1-\alpha)\hat{N}_t + \alpha\hat{k}_t + \hat{Z}_t.$$

Since $\vartheta > 1$ we overstate the size of $\hat{Z}_t$, when we use this latter equation to estimate the size of the technology shock from data on output, hours, and the capital stock. Furthermore, in as much as the entry of new firms measured by $\hat{j}_t$ depends upon the state of the business cycle, the usual measure of $\hat{Z}_t$ is further spoiled. We do not consider this book to be the right place to develop this matter further. Possible remedies have been suggested for instance by Rotemberg and Woodford (1995) and Hairault and Portier (1995). Instead, we continue to use the parameters from the benchmark model so that we are able to compare our results to those obtained in the Section 2.6.1 and Section 2.6.2.

What we further need are the parameters of the money supply process, of the transaction costs function, and of the structure of the monopolistic intermediary goods sector.

Our measure of money supply is the West-German monetary aggregate M1 per capita. As in Section 1.5 we focus on the period 1975.i through 1989.iv. The average quarterly growth rate of this aggregate was 1.67 percent. We fitted an AR(1) process to the deviations of μ_t from this value. The autocorrelation parameter from this estimation is not significantly different from zero and the estimated standard deviation of the innovations is $\sigma^\mu = 0.0173$. We use the average velocity of M1 with respect to consumption of 0.84 to determine γ from (2.107b). Finally, we can use the following observation to find an appropriate value of κ: The lhs of equation (2.107b) is equal to

$$\frac{1}{\pi(1-\delta+r)}.$$

The term in the denominator is the nominal interest rate factor, i.e., one plus the nominal interest rate q, say. This implies the following long run interest rate elasticity of the demand for real money balances:

$$\frac{d(M/P)/(M/P)}{dq/q} = \frac{-1}{(1+\kappa)\pi(1-\delta+r)}.$$

The estimate of this elasticity provided by HOFFMAN, RASCHE, and TIESLAU (1995) is about -0.2. Since $1/R \approx 1$ we use $\kappa = 4$.

Table 2.4

Preferences	Production	
β=0.994	a=1.005	α=0.27
η=2.0	δ=0.011	ρ^Z=0.90
N=0.13	σ^Z=0.0072	

Money Supply	Transactions Costs	Market Structure
μ=1.0167	$C/(M/P)$=0.84	φ=0.25
ρ^μ=0.0	κ=4.0	ϵ=6.0
σ^μ=0.0173		

The degree of nominal rigidity in our model is determined by the parameter φ. According to the estimates found in ROTEMBERG (1987) it takes about four quarters to achieve full price adjustment. Therefore, we use $\varphi = 0.25$. LINNEMANN (1999) presents estimates of markups for Germany, which imply a price elasticity of $\epsilon = 6$. Table 2.4 summarizes this choice of parameters.

Results. The Gauss program `NKPK.g` implements the solution. To understand the mechanics of the model, we consider the case without nominal frictions first. Figure 2.5 displays the time paths of several variables after a one-time shock to the money supply process (2.98) in period $t = 3$ of size σ^μ. Before this shock the economy was on its balanced growth path, after this shock the growth factor of money follows (2.98) with $\epsilon_t^\mu = 0$.

The case $\rho^\mu = 0$ highlights the unanticipated effect of the shock, since after period 3 the money growth rate is back on its

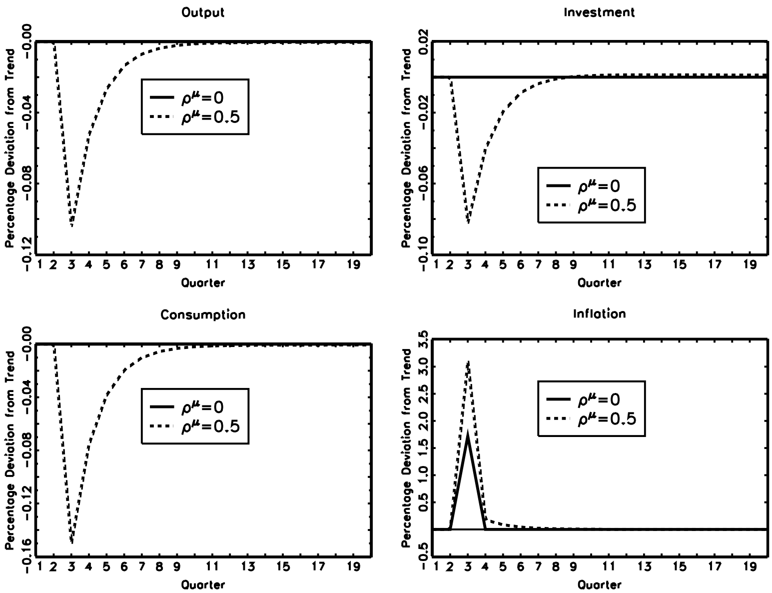

Figure 2.5: Real Effects of a Monetary Shock in the Model Without Nominal Rigidities

stationary path. The money transfer in period 3 raises the household's income unexpectedly. Since both consumption and leisure are normal goods the household's demand for consumption increases and its labor supply decreases. The latter raises the real wage so that marginal costs increase. Higher costs and excess demand raise inflation. This increase just offsets the extra amount of money so that the real stock of money does not change. Therefore, none of the real variables really changes. Money is neutral. This can be seen in Figure 2.5 since the impulse responses of output, consumption, and investment coincide with the zero line.

Things are different when the shock is autocorrelated. In this case there is also an anticipated effect. Households know that money growth will remain above average for several periods and expect above average inflation. This in turn increases the expected costs of money holdings and households reduce their cash holdings. As a consequence, the velocity of money with respect to consumption increases. To offset this negative effect on transaction costs the households reduce consumption. Their desire to smooth consumption finally entails less investment. Note however that these effects are very small. For instance, consumption in period 3 is 0.16 percent below its stationary value, and investment is 0.08 percent below its steady state level.

We find very different impulse responses, if nominal rigidities are present. This can be seen in Figure 2.6. Since inflation cannot adjust fully, households expect above average inflation even in the case of $\rho^\mu = 0$. This creates a desire to shift consumption to the current period so that there is excess demand. Monopolistically competitive firms are willing to satisfy this demand since their price exceeds their marginal costs. Thus output increases. The household's desire to spread the extra income over several periods spurs investment into physical capital.

There is another noteworthy property of the model: The spike-like shape of the impulse responses. Consumption, hours, output, and investment are almost back on their respective growth paths

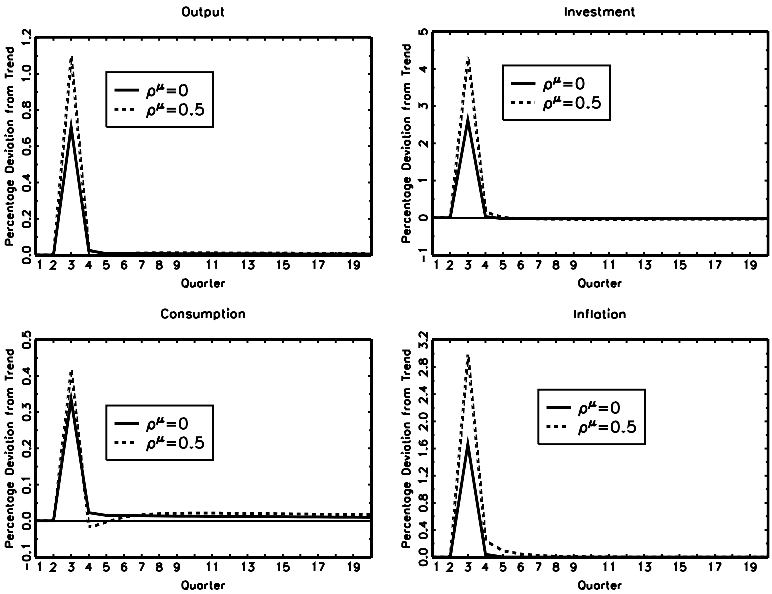

Figure 2.6: Impulse Responses to a Monetary Shock in the New Keynesian Phillips Curve Model

after period 3, irrespective of whether or not the monetary shock is autocorrelated. This is in stark contrast to the findings of empirical studies. For instance, according to the impulse responses estimated by Cochrane (1998) and, more recently, by Christiano, Eichenbaum, and Evans (2005) the response of output is hump shaped and peaks after eight quarters. The apparent failure of the model to explain the persistence of a monetary shock has let many researches to question the usefulness of the New Keynesian Phillips curve. In a recent paper Eichenbaum and Fisher (2004) argue that the Calvo (1983) model is able to explain persistent effects of monetary shocks if one abandons the convenient but arbitrary assumption of a constant price elasticity. Walsh (2005) argues that labor market search, habit persistence in consumption, and monetary policy inertia together can explain the long-lasting effects of monetary shocks. However, as Heer and Maussner (2007) point out, this result may be due to the assumption of prohibitively high costs of capital adjustment. In Christiano, Eichenbaum, and Evans (2005) wage staggering and variable capacity utilization account for the close fit between the estimated and the model-implied impulse responses of output and inflation.

Table 2.5 reveals the contribution of monetary shocks to the business cycle. To fully understand the model we must disentangle several mechanisms that work simultaneously. For this reason, columns 2 to 4 present simulations, where neither monetary shocks, nor nominal rigidities, nor monopolistic elements are present. This requires to set $\vartheta = 1$, $\varphi = 0$, and $\sigma^{\mu} = 0$ in the program `NKPK.g`. Obviously, this model behaves almost like the benchmark model (see Table 2.2).

Next consider columns 5 to 7. In this model, there are no monetary shocks, but there are monopolistic price setters facing nominal rigidities. The most immediate differences are: output is more volatile and hours are less volatile than in the benchmark model. How can this happen? Note that under monopolistic price setting the marginal product of labor is larger than it is under perfect competition. The same is true for the marginal product of capital. Thus, a technology shock that shifts the production function

Table 2.5

	$\vartheta=1,\varphi=0,\sigma^\mu=0$			$\sigma^\mu=0$			$\sigma^\mu=0.0173$		
Variable	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	1.43	1.00	0.63	1.55	1.00	0.68	1.69	1.00	0.56
	(1.14)	(1.00)	(0.80)	(1.14)	(1.00)	(0.80)	(1.14)	(1.00)	(0.80)
Consump-	0.53	0.99	0.65	0.55	0.98	0.72	0.64	0.98	0.52
tion	(1.18)	(0.79)	(0.84)	(1.18)	(0.79)	(0.84)	(1.18)	(0.79)	(0.84)
Invest-	6.16	1.00	0.63	6.87	1.00	0.67	7.31	1.00	0.58
ment	(2.59)	(0.75)	(0.79)	(2.59)	(0.75)	(0.79)	(2.59)	(0.75)	(0.79)
Hours	0.76	1.00	0.63	0.59	0.99	0.75	0.97	0.86	0.23
	(0.78)	(0.40)	(0.31)	(0.78)	(0.40)	(0.31)	(0.78)	(0.40)	(0.31)
Real	0.67	0.99	0.65	0.66	0.99	0.72	0.81	0.97	0.45
Wage	(1.17)	(0.41)	(0.91)	(1.17)	(0.41)	(0.91)	(1.17)	(0.41)	(0.91)
Inflation	0.27	−0.53	−0.07	0.31	−0.48	−0.05	1.62	0.30	−0.06
	(0.28)	(0.04)	(−0.03)	(0.28)	(0.04)	(−0.03)	(0.28)	(0.04)	(−0.03)

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x. Empirical magnitudes in parenthesis.

outward boosts output more than it would do in a competitive environment. Due to the fixed costs of production, the shock also raises profits and thus dividend payments to the household. This in turn increases the household's demand for leisure. Since prices do not fully adjust, these effects are a bit smaller than they are in a purely real model without nominal frictions.[32]

Columns 8 to 10 present the results from simulations where both technology shocks and monetary shocks are present. The most noteworthy effect concerns working hours. The standard deviation of this variable increases by 64 percent. The wealth effect that we identified above now works in the opposite direction: A monetary shock squeezes the profits of firms, since marginal costs rise and prices cannot fully adjust. As a consequence, the house-

[32] A detailed comparison between a real and a monetary model of monopolistic price setting appears in Maussner (1999).

hold's demand for leisure falls. But note, most of the shock is absorbed by inflation, which increases substantially.

Appendix 3: Solution of the Stochastic LQ Problem

In this Appendix we provide the details of the solution of the stochastic linear quadratic (LQ) problem. If you are unfamiliar with matrix algebra, you should consult 11.1 before proceeding.

Using matrix algebra we may write the Bellman equation (2.15) as follows:

$$\begin{aligned}\mathbf{x}'P\mathbf{x} + d = \max_{\mathbf{u}} \Big[&\mathbf{x}'Q\mathbf{x} + \mathbf{u}'R\mathbf{u} + 2\mathbf{u}'S\mathbf{x} \\ &+ \beta E\Big(\mathbf{x}'A'PA\mathbf{x} + \mathbf{u}'B'PA\mathbf{x} + \boldsymbol{\epsilon}'PA\mathbf{x} \\ &+ \mathbf{x}'A'PB\mathbf{u} + \mathbf{u}'B'PB\mathbf{u} + \boldsymbol{\epsilon}'PB\mathbf{u} \\ &+ \mathbf{x}'A'P\boldsymbol{\epsilon} + \mathbf{u}'B'P\boldsymbol{\epsilon} + \boldsymbol{\epsilon}'P\boldsymbol{\epsilon} + d\Big)\Big].\end{aligned} \tag{A.3.1}$$

Since $E(\boldsymbol{\epsilon}) = \mathbf{0}$ the expectation of all linear forms involving the vector of shocks $\boldsymbol{\epsilon}$ evaluate to zero. The expectation of the quadratic form $\boldsymbol{\epsilon}'P\boldsymbol{\epsilon}$ is:

$$E\left(\sum_{i=1}^{n}\sum_{i=1}^{n} p_{ij}\epsilon_i\epsilon_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\sigma_{ij},$$

where σ_{ij} (σ_{ii}) denotes the covariance (variance) between ϵ_i and ϵ_j (of ϵ_i). It is not difficult to see that this expression equals $\operatorname{tr}(P\Sigma)$. Furthermore, since $P = P'$ and

$$z := \mathbf{u}'B'PA\mathbf{x} = z' = (\mathbf{x}'A'PB'\mathbf{u})'$$

we may write the Bellman equation as

$$\begin{aligned}\mathbf{x}'P\mathbf{x} + \mathbf{d} = \max_{\mathbf{u}} \Big[&\mathbf{x}'Q\mathbf{x} + 2\mathbf{u}S\mathbf{x} + \mathbf{u}'R\mathbf{u} + \beta\mathbf{x}'A'PA\mathbf{x} \\ &+ 2\beta\mathbf{x}'A'PB\mathbf{u} + \beta\mathbf{u}'B'PB\mathbf{u} + \beta\operatorname{tr}(P\Sigma) + \beta d\Big].\end{aligned} \tag{A.3.2}$$

This is equation (2.16) in the main text. Differentiation of the rhs of this expression with respect to $\mathbf{u}$ yields

$$2S\mathbf{x} + 2R\mathbf{u} + 2\beta(\mathbf{x}'A'PB)' + 2\beta(B'PB)\mathbf{u}.$$

Setting this equal to the zero vector and solving for $\mathbf{u}$ gives

$$\begin{aligned}
&\underbrace{(R+\beta B'PB)}_{C^{-1}}\mathbf{u} = -\underbrace{(S+\beta B'PA)}_{D}\mathbf{x} \\
\Rightarrow \quad &\mathbf{u} = -CD\mathbf{x}.
\end{aligned} \tag{A.3.3}$$

If we substitute this solution back into (A.3.2), we get:

$$\begin{aligned}
\mathbf{x}'P\mathbf{x} + d &= \mathbf{x}'Q\mathbf{x} - 2(CD\mathbf{x})'S\mathbf{x} + (CD\mathbf{x})'RCD\mathbf{x} + \beta\mathbf{x}'A'PA\mathbf{x} \\
&\quad - 2\beta\mathbf{x}'A'PBCD\mathbf{x} + \beta(CD\mathbf{x})'B'PBCD\mathbf{x} + \beta\,\mathrm{tr}\,(P\Sigma) + \beta d \\
&= \mathbf{x}'Q\mathbf{x} + \beta\mathbf{x}'A'PA\mathbf{x} \\
&\quad - 2\mathbf{x}'D'C'S\mathbf{x} - 2\beta\mathbf{x}'A'PBCD\mathbf{x} \\
&\quad + \mathbf{x}'D'C'RCD\mathbf{x} + \beta\mathbf{x}'D'C'B'PBCD\mathbf{x} \\
&\quad + \beta\,\mathrm{tr}\,(P\Sigma) + \beta d.
\end{aligned}$$

The expression on the fourth line can be simplified to

$$\begin{aligned}
&- 2\mathbf{x}'D'C'S\mathbf{x} - \underbrace{2\beta\mathbf{x}'A'PBCD\mathbf{x}}_{=2\beta\mathbf{x}'D'C'B'PA\mathbf{x}} \\
&= -2\mathbf{x}'D'C'\underbrace{(S+\beta B'PA)}_{D}\mathbf{x} = -2\mathbf{x}'D'C'D\mathbf{x}.
\end{aligned}$$

The terms on the fifth line add to

$$\mathbf{x}'D'C'\underbrace{(R+\beta B'PB)C}_{I}D\mathbf{x} = \mathbf{x}'D'C'D.$$

Therefore,

$$\mathbf{x}'P\mathbf{x} + d = \mathbf{x}'Q\mathbf{x} + \beta\mathbf{x}'A'PA\mathbf{x} - \mathbf{x}'D'C'D\mathbf{x} + \beta\,\mathrm{tr}(P\Sigma) + \beta d. \tag{A.3.4}$$

For this expression to hold, the coefficient matrices of the various quadratic forms on both sides of equation (A.3.4) must satisfy the matrix equation

$$P = Q + \beta A'PA + D'C'D,$$

and the constant d must be given by

$$d = \frac{\beta}{1-\beta}\,\mathrm{tr}(P\Sigma).$$

This finishes the derivation of the solution of LQ the problem.

Appendix 4: Derivation of the Log-Linear Model of the New Keynesian Phillips Curve

In this appendix we provide the details of the solution of the model from Section 2.6.3.

The Household's Problem. The Lagrangean of the household's problem is:

$$\begin{aligned}\mathscr{L} = E_0 \sum_{t=0}^{\infty} \beta^t \Bigg\{ & \frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ & + \Lambda_t \Bigg[\frac{W_t}{P_t} N_t + (r_t - \delta)K_t + D_t + T_t \\ & - \gamma \left(\frac{C_t}{M_{t+1}/P_t} \right)^{\kappa} C_t - C_t - (K_{t+1} - K_t) - \frac{M_{t+1} - M_t}{P_t} \Bigg] \Bigg\}.\end{aligned}$$

Differentiating this expression with respect to C_t, N_t, K_{t+1} and M_{t+1} provides the following first-order conditions:

$$\begin{aligned} 0 &= C_t^{-\eta}(1-N_t)^{\theta(1-\eta)} - E_t \Lambda_t \left[1 + \gamma(1+\kappa) \left(\frac{C_t}{M_{t+1}/P_t} \right)^{\kappa} \right], \\ 0 &= \theta C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)-1} - \Lambda_t \frac{W_t}{P_t}, \\ 0 &= \Lambda_t - \beta E_t \Lambda_{t+1}(1 - \delta + r_{t+1}), \\ 0 &= E_t \left\{ -\frac{\Lambda_t}{P_t} + \kappa\gamma \left(\frac{C_t}{M_{t+1}/P_t} \right)^{\kappa+1} \frac{\Lambda_t}{P_t} + \beta \frac{\Lambda_{t+1}}{P_{t+1}} \right\}. \end{aligned} \tag{A.4.1}$$

As usual, we must define variables that are stationary. We choose $c_t := C_t/A_t$, $k_t := K_t/A_t$, $\lambda_t := \Lambda_t A_t^{\eta}$, $w_t := W_t/(P_t A_t)$, $m_{t+1} := M_{t+1}/(A_t P_t)$, and $j_t := J_t/A_t$. The inflation factor is $\pi_t := P_t/P_{t-1}$. Since the price level is determined in period t, this variable is also a period t variable. The growth factor of money supply, also determined in period t, is given by $\mu_t := M_{t+1}/M_t$, where M_t is the beginning-of-period money stock and M_{t+1} the end-of-period money stock. In these variables, we can rewrite the system (A.4.1) as follows:

$$c_t^{-\eta}(1-N_t)^{\theta(1-\eta)} = \lambda_t \left(1 + \gamma(1+\kappa) \left(\frac{c_t}{m_{t+1}} \right)^{\kappa} \right), \tag{A.4.2a}$$

$$\lambda_t w_t = \theta c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)-1}, \tag{A.4.2b}$$

$$m_{t+1} = \frac{\mu_t}{a\pi_t} m_t, \tag{A.4.2c}$$

$$\lambda_t = \beta a^{-\eta} E_t \lambda_{t+1} \left(1 - \delta + r_{t+1}\right), \tag{A.4.2d}$$

$$\beta a^{-\eta} E_t \frac{\lambda_{t+1}}{\pi_{t+1}} = \lambda_t \left(1 - \kappa\gamma \left(\frac{c_t}{m_{t+1}}\right)^{\kappa+1}\right). \tag{A.4.2e}$$

Price Setting. To study the price setting behavior, it is convenient to first solve the firm's cost minimization problem

$$\min_{N_{jt},K_{jt}} \quad \frac{W_t}{P_t} N_{jt} + r_t K_{jt} \quad \text{s.t. } (2.102).$$

The first-order conditions for this problem are easy to derive. They are:

$$w_t = g_t(1-\alpha) Z_t N_{jt}^{-\alpha} (K_{jt}/A_t)^{\alpha} = g_t(1-\alpha) Z_t (k_t/N_t)^{\alpha}, \tag{A.4.3a}$$

$$r_t = g_t \alpha Z_t N_{jt}^{1-\alpha} (K_{jt}/A_t)^{\alpha-1} = g_t \alpha Z_t (k_t/N_t)^{\alpha-1}, \tag{A.4.3b}$$

where g_t is the Lagrange multiplier of the constraint (2.102), and $w_t := W_t/(P_t A_t)$ is the real wage rate per unit of effective labor.[33] It is well known from elementary production theory that g_t equals the marginal costs of production. Furthermore, the constant scale assumption with respect to $Y_{jt} + F$ also implies that g_t are the variable unit costs of production:

$$g_t = \frac{(W_t/P_t) N_{jt} + r_t K_{jt}}{Y_{jt} + F}.$$

Marginal costs as well as the capital-output ratio are the same in all intermediary firms due to the symmetry that is inherent in the specification of the demand and production function. For later use we note the factor demand functions that are associated with this solution:

[33] Note that g_t is equal for all firms. This can be seen by using

$$\frac{w_t}{r_t} = \frac{1-\alpha}{\alpha} \frac{K_{jt}}{N_{jt}},$$

which implies that all firms choose the same capital-labor ratio $k_t/N_t \equiv K_{jt}/N_{jt}$, since all firms face the same real wages and rental rates. Via equation (A.4.3b) this also implies $g_t = g_{jt}$ for all j.

$$N_{jt} = \frac{Y_{jt}+F}{A_t Z_t}\left(\frac{1-\alpha}{\alpha}\right)^{\alpha}\left(\frac{w_t}{r_t}\right)^{-\alpha}, \tag{A.4.4a}$$

$$K_{jt} = \frac{Y_{jt}+F}{Z_t}\left(\frac{1-\alpha}{\alpha}\right)^{\alpha-1}\left(\frac{w_t}{r_t}\right)^{1-\alpha}. \tag{A.4.4b}$$

In each period $(1-\varphi)J_t$ firms choose their optimal money price P_{At} and φJ_t firms increase their price according to average inflation,

$$P_{Nt} = \pi P_{Nt-1}.$$

Therefore, the aggregate price level given in equation (2.101) is:

$$P_t = \left[(1-\varphi)P_{At}^{1-\epsilon} + \varphi(\pi P_{Nt-1})^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}.$$

Now observe that the pool of firms that are not allowed to choose their price optimally consists itself of firms that were able to set their optimal price in the previous period and those unlucky ones that were not allowed to do so. Thus, P_{Nt-1} is in turn the following index:

$$P_{N\,t-1} = \left[(1-\varphi)P_{A\,t-1}^{1-\epsilon} + \varphi(\pi P_{N\,t-2})^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}.$$

Using this formula recursively establishes:

$$P_t = \left[(1-\varphi)\left\{P_{A\,t}^{1-\epsilon} + \varphi(\pi P_{A\,t-1})^{1-\epsilon} + \varphi^2(\pi^2 P_{A\,t-2})^{1-\epsilon} + \dots\right\}\right]^{\frac{1}{1-\epsilon}},$$

which implies

$$\varphi(\pi P_{t-1})^{1-\epsilon} = \left[(1-\varphi)\{\varphi(\pi P_{A\,t-1})^{1-\epsilon} + \varphi^2(\pi^2 P_{A\,t-2})^{1-\epsilon} + \dots\}\right].$$

Thus, the aggregate price level can equivalently be written as

$$P_t = \left[(1-\varphi)P_{At}^{1-\epsilon} + \varphi(\pi P_{t-1})^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}. \tag{A.4.5}$$

We now turn to the first-order conditions that determine the optimal price of type A firms. Maximizing the expression in (2.104) with respect to P_{At} provides the following condition:

$$\begin{aligned}\underbrace{\frac{\epsilon-1}{\epsilon}}_{=:1/\vartheta} P_{At}E_t\sum_{\tau=t}^{\infty}\varphi^{\tau-t}\varrho_\tau\left(\frac{\pi^{\tau-t}}{P_\tau}\right)^{(1-\epsilon)}\frac{Y_\tau}{J_\tau}\\ =E_t\sum_{\tau=t}^{\infty}\varphi^{\tau-t}\varrho_\tau\left(\frac{\pi^{\tau-t}}{P_\tau}\right)^{-\epsilon}g_\tau\frac{Y_\tau}{J_\tau}.\end{aligned}$$

We multiply both sides by $P_t^{-\epsilon}$ and replace ϱ_τ by the rhs of equation (2.105). The result is:

$$\begin{aligned}
&\frac{1}{\vartheta}\left(\frac{P_{At}}{P_t}\right) E_t \sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}\frac{\lambda_\tau}{\lambda_t}\pi^{(1-\epsilon)(\tau-t)}\left(\frac{P_\tau}{P_t}\right)^{\epsilon-1}\frac{Y_\tau}{J_\tau} \\
&= E_t \sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}\frac{\lambda_\tau}{\lambda_t}\pi^{-\epsilon(\tau-t)} g_\tau \left(\frac{P_\tau}{P_t}\right)^{\epsilon}\frac{Y_\tau}{J_\tau}.
\end{aligned} \tag{A.4.6}$$

Our next task is to determine aggregate output and employment. Note from (2.100) that final goods producers use different amounts of type A and N goods since the prices of these inputs differ. Therefore, aggregate output is:

$$\begin{aligned}
Y_t &= (1-\varphi)J_t\frac{P_{At}}{P_t}Y_{At} + \varphi J_t\frac{\pi}{\pi_t}Y_{Nt} \\
&= (1-\varphi)J_t\left[\frac{P_{At}}{P_t}\left(Z_tA_tN_{At}(K_{At}/A_tN_{At})^{1-\alpha} - F\right)\right] \\
&\quad + \varphi J_t\left[\frac{\pi}{\pi_t}\left(Z_tA_tN_{Nt}(K_{Nt}/A_tN_{Nt})^{1-\alpha} - F\right)\right].
\end{aligned}$$

Using the fact that all producers choose the same capital-labor ratio k_t/N_t provides:

$$\begin{aligned}
Y_t = A_t &\left[\frac{P_{At}}{P_t}Z_t\underbrace{(1-\varphi)J_tN_{At}}_{n_tN_t}(k_t/N_t)^{1-\alpha} + \frac{\pi}{\pi_t}Z_t\underbrace{\varphi J_tN_{Nt}}_{(1-n_t)N_t}(k_t/N_t)^{1-\alpha}\right] \\
&- J_tF\left[(1-\varphi)\frac{P_{At}}{P_t} + \varphi\frac{\pi}{\pi_t}\right],
\end{aligned}$$

where the fraction of workers employed by type A firms n_t is given by:

$$n_t := \frac{(1-\varphi)J_tN_{At}}{N_t}. \tag{A.4.7}$$

From this we derive the following equation in terms of aggregate output per efficiency unit A_t:

$$\begin{aligned}
y_t := \frac{Y_t}{A_t} &= Z_tN_t^{1-\alpha}k_t^{\alpha}\left[n_t\frac{P_{At}}{P_t} + (1-n_t)\frac{\pi}{\pi_t}\right] \\
&- j_tF\left[(1-\varphi)\frac{P_{At}}{P_t} + \varphi\frac{\pi}{\pi_t}\right].
\end{aligned} \tag{A.4.8}$$

In the log-linear version of this equation the variable n_t drops out. Thus, there is no need to derive the equation that determines this variable.

Finally, consider the household's budget constraint (2.95). In equilibrium it holds with equality. Using the government's budget constraint (2.97) and the definition of dividends (2.106), we end up with the following resource constraint:

$$ak_{t+1} = y_t + (1-\delta)k_t - \gamma \left(\frac{c_t}{m_{t+1}}\right)^{\kappa} c_t - c_t. \tag{A.4.9}$$

The Log-Linear Model. The dynamic model consists of equations (A.4.2), (A.4.3), (A.4.5),(A.4.6), (A.4.8), and (A.4.9). The stationary equilibrium of this system is considered in the main text so that we can focus on the derivation of the log-linear equations. First, consider the variables that play the role of the control variables in the system (2.47). These are the deviations of consumption, working hours, output, the inflation factor, the real wage rate, and the rental rate of capital from their respective steady state levels:

$$\mathbf{u}_t := [\hat{c}_t, \hat{N}_t, \hat{y}_t, \hat{\pi}_t, \hat{w}_t, \hat{r}_t]'.$$

The state variables with predetermined initial conditions are the stock of capital and beginning-of-period money real money balances. Thus, in terms of (2.47):

$$\mathbf{x}_t = [\hat{k}_t, \hat{m}_t]'.$$

Purely exogenous are the technological shock $\hat{Z}_t$, the monetary shock $\hat{\mu}_t$, and the entrance rate of firms $\hat{j}_t$ into the intermediary goods sector. For the latter we will assume it is independent of the state of the business cycle so that $\hat{j}_t = 0$ for all t.[34] Thus,

$$\mathbf{z}_t = [\hat{Z}_t, \hat{\mu}_t]'.$$

The remaining variables are the shadow price of capital λ_t, firms' marginal costs g_t, and real end-of-period money balances m_{t+1}. Note, that we cannot determine the latter from equation (A.4.2c), since we

[34] For instance, ROTEMBERG and WOODFORD (1995) link $\hat{j}_t$ to the technological shock.

need this equation to determine π_t. Thus, in addition to λ_t and g_t, this variable is a costate. To keep to the dating convention in (2.47) we define the auxiliary variable $x_t \equiv m_{t+1}$. Hence, our vector of costate variables comprises:

$$\lambda_t = [\hat{\lambda}_t, \hat{g}_t, \hat{x}_t]'.$$

We first present the static equations that relate control variables to state and costate variables. The log-linear versions of equations (A.4.2a) through (A.4.2c) are

$$-(\eta + \xi_1)\hat{c}_t - \xi_2 \hat{N}_t = \hat{\lambda}_t - \xi_1 \hat{x}_t, \tag{A.4.10a}$$

$$(1-\eta)\hat{c}_t - \xi_3 \hat{N}_t - \hat{w}_t = \hat{\lambda}_t, \tag{A.4.10b}$$

$$\hat{\pi}_t = \hat{m}_t - \hat{x}_t + \hat{\mu}_t, \tag{A.4.10c}$$

$$\begin{aligned}
\xi_1 &:= \frac{\kappa\gamma(1+\kappa)(c/x)^\kappa}{1+\gamma(1+\kappa)(c/x)^\kappa}, \quad \frac{c}{x} = \frac{C}{\mu(M/P)},\\
\xi_2 &:= \theta(1-\eta)\frac{N}{1-N},\\
\xi_3 &:= [\theta(1-\eta)-1]\frac{N}{1-N}.
\end{aligned}$$

The log-linear cost-minimizing conditions (A.4.3) deliver two further equations:

$$\alpha\hat{N}_t + \hat{w}_t = \alpha\hat{k}_t + \hat{g}_t + \hat{Z}_t, \tag{A.4.10d}$$

$$(\alpha-1)\hat{N}_t + \hat{r}_t = (\alpha-1)\hat{k}_t + \hat{g}_t + \hat{Z}_t. \tag{A.4.10e}$$

To derive the sixth equation we use the formula for the price level to write

$$\pi_t = \frac{P_t}{P_{t-1}} = \left[(1-\varphi)\left(\frac{P_{At}}{P_t}\underbrace{\frac{P_t}{P_{t-1}}}_{\pi_t}\right)^{1-\epsilon} + \varphi\pi^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}.$$

Log-linearizing at $P_A/P = 1$ provides:

$$\hat{\pi}_t = \frac{1-\varphi}{\varphi}\widehat{P_{At}/P}_t.$$

We use this relation to derive

$$\hat{y}_t - \vartheta(1-\alpha)\hat{N}_t = \vartheta\alpha\hat{k}_t + \vartheta\hat{Z}_t + (1-\vartheta)\hat{j}_t. \tag{A.4.10f}$$

from equation (A.4.8). The six equations (A.4.10a) through (A.4.10f) determine the control variables. We now turn to the dynamic equations that determine the time paths of $\hat{k}_t$, $\hat{m}_t$, $\hat{x}_t \equiv \hat{m}_{t+1}$, $\hat{\lambda}_t$, and $\hat{g}_t$. The log-linear versions of the resource constraint (A.4.9), the Euler equations for capital and money balances (A.4.2d) and (A.4.2e), and the definition $x_t := m_{t+1}$ are:

$$\begin{aligned}
aE_t\hat{k}_{t+1} - (1-\delta)\hat{k}_t - \xi_4\hat{x}_t &= \frac{y}{k}\hat{y}_t - \xi_5\hat{c}_t, && \text{(A.4.11a)}\\
-E_t\hat{\lambda}_{t+1} + \hat{\lambda}_t &= \xi_6 E_t\hat{r}_{t+1}, && \text{(A.4.11b)}\\
E_t\hat{\lambda}_{t+1} - \hat{\lambda}_t - \xi_7\hat{x}_t &= -\xi_7\hat{c}_t + E_t\hat{\pi}_{t+1}, && \text{(A.4.11c)}\\
E_t\hat{m}_{t+1} - \hat{x}_t &= 0, && \text{(A.4.11d)}
\end{aligned}$$

$$\begin{aligned}
\xi_4 &:= \kappa\gamma(c/x)^{\kappa}(c/k),\\
\xi_5 &:= (1+\gamma(1+\kappa)(c/x)^{\kappa})(c/k),\\
\xi_6 &:= 1-\beta a^{-\eta}(1-\delta),\\
\xi_7 &:= \frac{\kappa\gamma(1+\kappa)(c/x)^{1+\kappa}}{1-\gamma\kappa(c/x)^{1+\kappa}}.
\end{aligned}$$

The remaining fifth equation is the log-linear condition for the firms' optimal price:

$$\frac{(1-\varphi)(1-\varphi\beta a^{-\eta})}{\varphi}\hat{g}_t = -\beta a^{-\eta}E_t\hat{\pi}_{t+1} + \hat{\pi}_t. \tag{A.4.11e}$$

This looks nice and resembles a Phillips curve since it relates the current inflation rate to the expected future rate of inflation and a measure of labor market tension, which is here given by the deviation of marginal costs from their steady state level. It requires a substantial amount of algebra to get this relation and it is this task to which we turn next. Considering (A.4.6) we find:

$$\begin{aligned}
&(\widehat{P_{At}/P_t})\frac{1}{\vartheta}\frac{y}{j}\underbrace{\left(1+\varphi\beta a^{-\eta}+(\varphi\beta a^{-\eta})^2+\dots\right)}_{(1-\varphi\beta a^{-\eta})^{-1}}\\
&\quad+\frac{1}{\vartheta}\frac{y}{j}\sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}E_t\left[(\widehat{\lambda_\tau/\lambda_t})+(\epsilon-1)(\widehat{P_\tau/P_t})+(\widehat{y_\tau/j_\tau})\right]\\
&=g\frac{y}{j}\sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}E_t\left[(\widehat{\lambda_\tau/\lambda_t})+\epsilon(\widehat{P_\tau/P_t})+(\widehat{y_\tau/j_\tau})+\hat{g}_\tau\right].
\end{aligned}$$

Since $\vartheta g = 1$ and $\widehat{P_{At}/P_t} = [\varphi/(1-\varphi)]\hat{\pi}_t$ (see above), we can simplify this expression to

$$\frac{\varphi}{(1-\varphi)(1-\varphi\beta a^{-\eta})}\hat{\pi}_t = \sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}E_t\left[(\widehat{P_\tau/P_t}) + \hat{g}_\tau\right]. \tag{A.4.12}$$

Next, we shift the time index one period into the future, multiply through by $\varphi\beta a^{-\eta}$, and compute the conditional expectation of the ensuing expression:[35]

$$\begin{aligned}&\left(\frac{\varphi}{1-\varphi}\right)\left(\frac{\varphi\beta a^{-\eta}}{1-\varphi\beta a^{-\eta}}\right)E_t\hat{\pi}_{t+1}\\&= E_t\Bigg[\left(\varphi\beta a^{-\eta}\right)^2\widehat{\left(\frac{P_{t+2}}{P_{t+1}}\right)} + \left(\varphi\beta a^{-\eta}\right)^3\widehat{\left(\frac{P_{t+3}}{P_{t+1}}\right)} + \cdots + \varphi\beta a^{-\eta}\hat{g}_{t+1}\\&\quad + (\varphi\beta a^{-\eta})^2\hat{g}_{t+2} + \ldots\Bigg].\end{aligned}$$

We subtract this equation from (A.4.12) to arrive at:

$$\begin{aligned}&\frac{\varphi}{(1-\varphi)(1-\varphi\beta a^{-\eta})}\left(\hat{\pi}_t - \varphi\beta a^{-\eta}E_t\hat{\pi}_{t+1}\right)\\&= \hat{g}_t + E_t\Bigg[\varphi\beta a^{-\eta}\widehat{\left(\frac{P_{t+1}}{P_t}\right)} + \left(\varphi\beta a^{-\eta}\right)^2\left\{\widehat{\left(\frac{P_{t+2}}{P_t}\right)} - \widehat{\left(\frac{P_{t+2}}{P_{t+1}}\right)}\right\}\\&\quad + \left(\varphi\beta a^{-\eta}\right)^3\left\{\widehat{\left(\frac{P_{t+3}}{P_t}\right)} - \widehat{\left(\frac{P_{t+3}}{P_{t+1}}\right)}\right\} + \ldots\Bigg].\end{aligned} \tag{A.4.13}$$

Since

$$\widehat{\left(\frac{P_\tau}{P_t}\right)} = \sum_{s=t+1}^{\tau}\hat{\pi}_s,$$

the terms in curly brackets reduce to $\hat{\pi}_{t+1}$ so that the sum in brackets equals

[35] Here we use the law of iterated expectations according to which $E_t x_{t+1} = E_t\left(E_{t+1}x_{t+1}\right)$.

$$\hat{\pi}_{t+1} \underbrace{\left[\varphi\beta a^{-\eta} + (\varphi\beta a^{-\eta})^2 + \dots\right]}_{(\varphi\beta a^{-\eta})/(1-\varphi\beta a^{-\eta})}.$$

Substituting these results back into (A.4.13) delivers equation (A.4.11e).

To determine the time path of investment, we start from

$$i_t = y_t - \left(1 + \gamma \left(\frac{c_t}{x_t}\right)^{\kappa}\right) c_t, \quad x_t \equiv m_{t+1}.$$

The log-linearized version of this equation is:

$$\begin{aligned}
&\hat{i}_t = \iota_1 \hat{y}_t - \iota_2 \hat{c}_t + \iota_3 \hat{x}_t,\\
&\quad \iota_1 := (y/i) = \frac{y/k}{a + \delta - 1}, \quad \iota_2 := \left(1 + (1+\kappa)\gamma \left(\frac{c}{x}\right)^{\kappa}\right) \frac{c}{i},\\
&\quad \iota_3 := \kappa\gamma \left(\frac{C}{\mu(M/P)}\right)^{\kappa} \frac{c}{i}, \quad \frac{c}{i} = \frac{y}{i} - 1.
\end{aligned}$$

Problems

2.1 **Certainty Equivalence**

Consider the deterministic linear quadratic optimal control problem of maximizing

$$\sum_{t=0}^{\infty} \beta^t \left[\mathbf{x}_t' Q \mathbf{x} + \mathbf{u}_t' R \mathbf{u}_t + 2\mathbf{u}_t' S \mathbf{x}_t\right]$$

subject to the linear law of motion

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t.$$

Adapt the steps followed in Section 2.2 and Appendix 3 to this problem and show that the optimal control as well as the matrix P are the solutions to equations (2.17) and (2.18), respectively.

2.2 **Relation Between the LQ Problems** (2.12) **and** (2.19)

Show that the linear quadratic problem with the current period return function

$$\begin{aligned} g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) := {} & \mathbf{x}_t' A_{xx} \mathbf{x}_t + \mathbf{u}_t' A_{uu} \mathbf{u}_t + \mathbf{z}_t A_{zz} \mathbf{z}_t \\ & + 2\mathbf{u}_t' A_{ux} \mathbf{x}_t + 2\mathbf{u}_t' A_{uz} \mathbf{z}_t + 2\mathbf{x}_t A_{xz} \mathbf{z}_t \end{aligned}$$

and the law of motion

$$\mathbf{x}_{t+1} = B_x \mathbf{x}_t + B_u \mathbf{u}_t + B_z \mathbf{z}_t$$

is a special case of the problem stated in equations (2.12) and (2.11). Toward that purpose define

$$\tilde{\mathbf{x}}_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{z}_t \end{bmatrix}, \quad \tilde{\boldsymbol{\epsilon}}_t = \begin{bmatrix} 0_{n\times 1} \\ \boldsymbol{\epsilon}_t \end{bmatrix}$$

and show how the matrices A, B, Q, R, and S must be chosen so that both problems coincide.

2.3 **Convex Costs of Price Adjustment**

Instead of the Calvo (1983) model, consider the following model of price setting introduced in Hairault and Portier (1995). Intermediate producers face convex costs of adjusting their price given by

$$PC_{jt} := (\psi/2)\left(\frac{P_{jt}}{P_{jt-1}} - \pi\right)^2.$$

Thus they solve the following problem:

$$\max \quad E_0 \sum_{t=0}^{\infty} \varrho_t \left[(P_{jt}/P_t) Y_{jt} - (W_t/P_t) N_{jt} - r_t K_{jt} - PC_{jt} \right],$$

s.t.

$$Y_{jt} = (P_{jt}/P_t)^{-\epsilon} (Y_t/J_t),$$
$$Y_{jt} = Z_t (A_t N_{jt})^{\alpha} K_{jt}^{1-\alpha} - F.$$

Calibrate the parameter ψ so that a one percent deviation of the firm's inflation factor P_{jt}/P_{jt-1} from average the average inflation factor entails costs of 0.01 percent of the firm's value added. Do you find more persistence of a money supply shock with this alternative specification of nominal rigidities? What happens, if you increase ψ?

2.4 **Government Spending in a Real Business Cycle Model**
In most OECD countries, wages and labor productivity are acyclic or even negatively correlated with output and working hours, while, in the stochastic Ramsey model, however, these correlations are positive and close to one (please compare table 2.2). One possible remedy for this shortcoming of the stochastic growth model is the introduction of a government spending shock. The following model is adapted from Baxter and King (1993) and Ambler and Paquet (1996).
Consider the stochastic growth model where the number of agents is normalized to one. Assume that utility is also a function of government consumption, where due to our normalization per capita government spending G_t is also equal to total government spending G_t. In particular, government consumption substitutes for private consumption C_t^p:

$$C_t = C_t^p + \vartheta G_t,$$

with $\vartheta < 1$ as some forms of government spending, for example military spending, do not provide utility for private consumption. The household maximizes her intertemporal utility:

$$\max_{C_0^p, N_0} \quad E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta} (1 - N_t)^{\theta(1-\eta)}}{1-\eta} \right],$$
$$\beta \in (0,1), \eta \geq 0, \theta \geq 0, \eta > \theta/(1+\theta),$$

subject to the budget constraint

$$C_t^p + I_t^p = (1-\tau)(w_t N_t + r_t K_t^p) + Tr_t.$$

Both wage income $w_t N_t$ and interest income $r_t K_t$ are taxed at the constant rate τ. The household also receives lump-sum transfers Tr_t from the government. The private capital stock evolves according to:

$$K^p_{t+1} = (1-\delta)K^P_t + I^p_t,$$

where δ denotes the depreciation rate. Production is described by a Cobb-Douglas Production Function, $Y_t = Z_t N_t^\alpha K_t^{1-\alpha}$, where the productivity Z_t follows an AR(1) process, $Z_{t+1} = Z_t^\varrho e^{\epsilon_t}$, with $\epsilon_t \sim N(0, \sigma^2)$ and $\varrho = 0.90$ and $\sigma = 0.007$. Factors are rewarded by their marginal products. Government consumption $G_t = g_t \bar{G}$ follows a stochastic process:

$$\ln g_t = \rho_g \ln g_{t-1} + \epsilon^g_t,$$

with $\epsilon^g_t \sim N(0, \sigma_g^2)$ and $\rho_g = 0.95$ and $\sigma_g = 0.01$. In the steady state, government consumption is constant and equal to 20% of output, $\bar{G} = 0.2\bar{Y}$. In equilibrium, the government budget is balanced:

$$\tau(w_t N_t + r_t K^p_t) = G_t + Tr_t.$$

The model is calibrated as follows: $\beta = 0.99$, $\eta = 2.0$, $\psi = 0.5$, $\alpha = 0.6$, $\delta = 0.02$. θ and τ are chosen so that the steady state labor supply $\bar{N}$ and transfers $\overline{Tr}$ are equal to 0.30 and 0, respectively.

a) Compute the steady state.
b) Compute the log-linear solution. Simulate the model and assume that ϵ_t and ϵ^g_t are uncorrelated. What happens to the correlation of labor productivity and wages with output and employment?
c) Assume that transfers are zero, $Tr_t = 0$, and that the income tax τ_t always adjusts in order to balance the budget. How are your results affected?
d) Assume now that the government expenditures are split evenly on government consumption G_t and government investment I^G_t. Government capital K^G_t evolves accordingly

$$K^G_{t+1} = (1-\delta)K^G_t + I^G_t,$$

and production is now given by

$$Y_t = Z_t = Z_t N_t^\alpha K_t^{1-\gamma} \left(K^G_t\right)^{1-\alpha-\gamma}$$

with $\alpha = 0.6$ and $\gamma = 0.3$. Recompute the model.

2.5 **Government Spending and Nominal Rigidities**
In the previous problem, you have learned about the 'wealth effect' of government demand. An increase in government expenditures results in a reduction of transfers and, hence, wealth of the households is decreased. Consequently, the households increase their labor supply and both employment and output increase. In this problem, you will learn about the traditional Keynesian IS-LM effect. Expansionary fiscal policy increases

aggregate demand and demand-constrained firms increases their output as prices are fixed in the short run. The model follows LINNEMANN and SCHABERT (2003).
Households maximize the expected value of a discounted stream of instantaneous utility:

$$\max_{C_0,N_0} \quad E_0\left[\sum_{t=0}^{\infty}\beta^t \frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta}\right],$$
$$\beta\in(0,1), \eta\geq 0, \theta\geq 0, \eta>\theta/(1+\theta).$$

A role for money is introduced into the model with the help of a cash-in-advance constraint:

$$P_tC_t \leq M_t + P_tTr_t,$$

Nominal consumption purchases P_tC_t are constrained by nominal beginning-of period money balances M_t and nominal government transfers P_tTr_t.[36] The household holds two kinds of assets, nominal money M_t and nominal bonds, B_t. Bonds yield a gross nominal return R_t. In addition, agents receive income from labor, $P_tw_tN_t$, government transfers, P_tTr_t, and from firm profits, $\int_0^1 \Omega_{it}\, di$. The budget constraint is given by:

$$M_{t+1}+B_{t+1}+P_tc_t = P_tw_tN_t + R_tB_t + M_t + P_tTr_t + \int_0^1 \Omega_{it}di.$$

The number of firms i is one, $i\in(0,1)$. Firms are monopolistically competitive and set their prices in a staggered way as in the model of Section 2.6.3. Accordingly, profit maximization of the firms implies the New Keynesian Phillips curve:

$$\hat{\pi}_t = \psi\widehat{mc}_t + \beta E_t\left\{\hat{\pi}_{t+1}\right\}, \quad \psi=(1-\varphi)(1-\beta\varphi)\varphi^{-1},$$

where mc_t denotes marginal costs (compare (A.4.11e)).
Firms produce with labor only:

$$y_{it}=N_{it}.$$

Cost minimization implies that the real wage is equal to marginal costs:

[36] Government transfers are included in this cash-in-advance specification in order to avoid the following: an expansionary monetary policy consisting in a rise of M_{t+1} already increases prices P_t due to the expected inflation effect. Accordingly, real money balances M_t/P_t fall and so does real consumption C_t if government transfers do not enter the cash-in-advance constraint. This, however, contradicts empirical evidence.

$$w_t = mc_t.$$

The government issues money and nominal riskless one-period bonds and spends its revenues on government spending, G_t, and lump-sum transfers:

$$P_t Tr_t + P_t G_t + M_t + R_t B_t = B_{t+1} + M_{t+1}.$$

Real government expenditures follow an AR(1)-process:

$$\ln G_t = \rho \ln G_{t-1} + (1-\rho)\ln G + \epsilon_t$$

with $\epsilon_t \sim N(0, \sigma^2)$ and $\rho = 0.90$ and $\sigma = 0.007$.
Monetary policy is characterized by a forward-looking interest-rate rule:

$$\hat{R}_{t+1} = \rho_\pi E - t\hat{\pi}_{t+1} + \rho_y E_t \hat{y}_{t+1}, \quad \rho_{pi} > 1.$$

The restriction ρ_π is imposed in order to ensure uniqueness of the equilibrium.

a) Compute the first-order conditions of the household.
b) Compute the stationary equilibrium that is characterized by a zero-supply of bonds, $B_t = 0$,[37] and $R > 1$ (in this case, the cash-in-advance constraint is always binding). Furthermore, in equilibrium, the aggregate resource constraint is given by $y_t = c_t + G_t$ and firms are identical, $y_{it} = y_t = N_t = N_{it}$. Define the equilibrium with the help of the stationary variables $\{\pi_t, w_t, m_t \equiv \frac{M_t}{P_{t-1}}, R_t, y_t, G_t\}$.
c) Compute the steady-state.
d) Calibrate the model as in the previous problem. In addition, set $\rho_\pi = 1.5$, $\rho_y \in \{0, 0.1, 0.5\}$, $\pi = 1$, and $\varphi = 0.75$.
e) Log-linearize the model and compute the dynamics. How does consumption react to an expansionary fiscal policy? Does it increase (as IS-LM implies) or decrease (due to the wealth effect)?
f) Assume now that the interest-rate rule is subject to an exogenous autocorrelated shock with autoregressive parameter $\rho_R \in \{0, 0.5\}$. How does a shock affect the economy?
g) Assume that monetary policy is described by a money-growth rule that is subject to an autoregressive shock. Recompute the model for an autoregressive parameter $\rho_\mu \in \{0, 0.5\}$ and compare the impulse responses to those implied by an interest-rate rule.

[37] Why can we set the nominal bonds supply equal to zero?

http://www.springer.com/978-3-642-03148-9

Dynamic General Equilibrium Modeling
Computational Methods and Applications
Heer, B.; Maussner, A.
2009, XXXII, 702 p., Softcover
ISBN: 978-3-642-03148-9