Optimization under Commitment and Discretion, the Recursive Saddlepoint Method, and Targeting Rules and Instrument Rules: Lecture Notes

Lars E.O. Svensson
Sveriges Riksbank and Stockholm University

Version 2.993: May 2023

Contents

1 The model

Let the dynamic equations of the model be

$$\begin{bmatrix} X_{t+1} \\ Hx_{t+1|t} \end{bmatrix} = A \begin{bmatrix} X_t \\ x_t \end{bmatrix} + Bi_t + \begin{bmatrix} C \\ 0 \end{bmatrix} \varepsilon_{t+1} \tag{1.1}$$

for $t \geq 0$, where X_t is an n_X-vector of *predetermined variables* (one element of X_t can be unity, in order to handle constants), X_0 is given, x_t is an n_x-vector of *forward-looking variables*, i_t is an n_i-vector of *instruments* (control variables), and ε_t is an n_ε-vector of exogenous zero-mean iid shocks. The matrices A, B, C, and H are of dimension $(n_X + n_x) \times (n_X + n_x)$, $(n_X + n_x) \times n_i$, $n_X \times n_\varepsilon$, and $n_x \times n_x$, respectively. Without loss of generality, I can normalize the shocks so the covariance matrix of ε_t is I. Then the covariance matrix of the shocks to X_{t+1} is CC'. For any vector z_t, $z_{t+1|t}$ denotes the rational expectation $\mathrm{E}_t z_{t+1}$. A common special case is when the matrix $H \equiv I$, but in general H need not be invertible. Some rows of H may be zero.

A variable is a *predetermined* variable if and only if it has exogenous one-period-ahead forecast errors (cf. Klein (2000)).[1] The one-period-ahead forecast errors of X_t, $X_{t+1} - \mathrm{E}_t X_{t+1} = C\varepsilon_{t+1}$, which by assumption is exogenous. Thus, a predetermined variable is determined by lagged variables and contemporaneous exogenous shocks. Predetermined variables that only depend on lagged values of themselves and contemporaneous exogenous shocks are *exogenous* variables. A variable that is not a predetermined variable is a *non-predetermined* variable. The non-predetermined variables are the forward-looking variables x_t and the instruments i_t. Non-predetermined variables have forecast errors, $x_{t+1} - \mathrm{E}_t x_{t+1}$ and $i_{t+1} - \mathrm{E}_t i_{t+1}$, which are endogenous, that is, endogenous functions of the exogenous shocks.

The two blocks of (1.1) can be written

$$X_{t+1} = A_{11}X_t + A_{12}x_t + B_1 i_t + C\varepsilon_{t+1} \tag{1.2}$$

$$Hx_{t+1|t} = A_{21}X_t + A_{22}x_t + B_2 i_t, \tag{1.3}$$

where A and B are partitioned conformably with X_t and x_t,

$$A \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B \equiv \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}.$$

[1] This definition of predetermined variables is slightly more general than that of Blanchard and Kahn (1980). They define predetermined variables as having *zero* one-period-ahead forecast errors.

The upper block, (1.2), determines X_{t+1} given X_t, x_t, i_t, and ε_{t+1}. I assume that A_{22} is invertible, so the second block, (1.3), can be written

$$x_t = A_{22}^{-1}(Hx_{t+1|t} - A_{21}X_t - B_2 i_t).$$

The lower block determines x_t given $x_{t+1|t}$, X_t, and i_t.

Assuming that the shocks ε_t only enter the upper block in (1.1) is not restrictive. If some shocks enter the lower block, the model can be modified by defining additional predetermined variables for these shocks and entering those into the upper block, leaving no shocks but additional predetermined variables in the lower block.

Let

$$Y_t = D \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} \tag{1.4}$$

be an n_Y-vector of *target variables*, measured as the deviation from a fixed n_Y-vector of *target levels*, Y^*, where D is a matrix of dimension $n_Y \times (n_X + n_x + n_i)$. Let the *period loss function* be

$$L_t = \frac{1}{2} Y_t' \Lambda Y_t \equiv \frac{1}{2} \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}' W \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}, \tag{1.5}$$

where Λ and $W \equiv D'\Lambda D$ are given symmetric positive semidefinite matrices, where the elements of Λ are the weights on the target variables in the period loss function. Let the intertemporal loss function in period 0 be

$$\mathrm{E}_0 \sum_{t=0}^{\infty} (1-\delta)\delta^t L_t, \tag{1.6}$$

where $0 < \delta < 1$ is a discount factor.

2 Optimal policy under commitment: The commitment equilibrium

Consider minimizing the intertemporal loss function, (1.6), under commitment once-and-for-all in period $t = 0$, subject to (1.1) for $t \geq 0$ and $X_0 = \bar{X}_0$, where $\bar{X}_0$ is given. Variants of this problem are solved in Backus and Driffill (1986), Currie and Levine (1993), Sims (1999), Sims (2002b) and Sims (2000), and Söderlind (1999). The problem can be solved in several ways.

2.1 Set up the Lagrangian, derive the first-order conditions, and solve a difference equation

The standard method is to set up the Lagrangian, derive the first-order conditions, combine these with dynamic equations, and solve the resulting difference equation.

The model (1.1) can be written

$$\bar{H}\begin{bmatrix} X_{t+1} \\ x_{t+1|t} \\ i_{t+1|t} \end{bmatrix} = \bar{A}\begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} + \begin{bmatrix} C \\ 0 \end{bmatrix}\varepsilon_{t+1}, \tag{2.1}$$

where the matrices $\bar{A}$ and $\bar{H}$ are of dimension $(n_X + n_x) \times (n_X + n_x + n_i)$ and given by

$$\bar{A} \equiv \begin{bmatrix} A & B \end{bmatrix}, \qquad \bar{H} \equiv \begin{bmatrix} I & 0 & 0 \\ 0 & H & 0 \end{bmatrix}. \tag{2.2}$$

Set up the Lagrangian,

$$\begin{aligned}
\mathcal{L}_0 &= \mathrm{E}_0 \sum_{t=0}^{\infty}(1-\delta)\delta^t \left\{ L_t + \begin{bmatrix} \xi'_{t+1} & \Xi'_t \end{bmatrix} \left(\bar{H}\begin{bmatrix} X_{t+1} \\ x_{t+1|t} \\ i_{t+1|t} \end{bmatrix} - \bar{A}\begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} - \begin{bmatrix} C \\ 0 \end{bmatrix}\varepsilon_{t+1} \right) \right\} \\
&\quad + \frac{1-\delta}{\delta}\xi'_0(X_0 - \bar{X}_0) \\
&= \mathrm{E}_0 \sum_{t=0}^{\infty}(1-\delta)\delta^t \left\{ L_t + \begin{bmatrix} \xi'_{t+1} & \Xi'_t \end{bmatrix} \left(\bar{H}\begin{bmatrix} X_{t+1} \\ x_{t+1} \\ i_{t+1} \end{bmatrix} - \bar{A}\begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} - \begin{bmatrix} C \\ 0 \end{bmatrix}\varepsilon_{t+1} \right) \right\} \\
&\quad + \frac{1-\delta}{\delta}\xi'_0(X_0 - \bar{X}_0),
\end{aligned}$$

where ξ_{t+1} and Ξ_t are vectors of n_X and n_x Lagrange multipliers of the upper and lower block, respectively, of the model equations. The law of iterated expectations has been used in the second equality. Note that Ξ_t is dated to emphasize that it depends on information available in period t, since the lower block is an equation that determines x_t given information available in period t.

The first-order conditions with respect to X_t, x_t, and i_t for $t \geq 1$ can be written

$$\begin{bmatrix} X'_t & x'_t & i'_t \end{bmatrix} W + \begin{bmatrix} \xi'_t & \Xi'_{t-1} \end{bmatrix} \frac{1}{\delta}\bar{H} - \begin{bmatrix} \xi'_{t+1|t} & \Xi'_t \end{bmatrix} \bar{A} = 0. \tag{2.3}$$

The first-order condition with respect to X_t, x_t, and i_t for $t = 0$ can be written

$$\begin{bmatrix} X'_t & x'_t & i'_t \end{bmatrix} W + \begin{bmatrix} \xi'_t & 0 \end{bmatrix} \frac{1}{\delta}\bar{H} - \begin{bmatrix} \xi'_{t+1|t} & \Xi'_t \end{bmatrix} \bar{A} = 0, \tag{2.4}$$

where $X_0 = \bar{X}_0$. In comparison with (2.3), a vector of zeros enters in place of Ξ_{-1}, since there is no constraint corresponding to the lower block of (2.1) for $t = -1$. By including a fictitious vector of Lagrange multipliers, Ξ_{-1}, equal to zero,

$$\Xi_{-1} = 0, \tag{2.5}$$

in (2.4), the first-order conditions can be written more compactly as (2.3) for all $t \geq 0$ and (2.5).

The system of difference equations (2.3) has $n_X + n_x + n_i$ equations. The first n_X equations can be associated with the Lagrange multipliers ξ_t. The expression $-\frac{1-\delta}{\delta}\xi_t$ can be interpreted as the total marginal losses in period t of the predetermined variables X_t (for $t = 0$, with given X_0, the equations determine ξ_0). They are forward-looking variables: the Lagrange multipliers of the equations for the predetermined variables always are forward-looking, whereas the Lagrange multipliers of the equations for the forward-looking variables always are predetermined. The middle n_x equations can be associated with the Lagrange multipliers Ξ_t. The expression $(1-\delta)\Xi_t' A_{22}$ can be interpreted as the total marginal losses in period t of the forward-looking variables, x_t. I can also interpret $(1-\delta)\Xi_t' H$ as the total marginal loss in period t of the one-period-ahead expectations of the forward-looking variables, $x_{t+1|t}$. The last n_i equations are the first-order equations for the vector of instruments. In the special case when the lower right $n_i \times n_i$ submatrix W_{ii} of W is of full rank, the instruments can be solved in terms of the other variables and eliminated from (2.3), leaving the first $n_X + n_x$ equations involving the Lagrange multipliers and the predetermined and forward-looking variables only.

Rewrite the $n_X + n_x + n_i$ first-order conditions as

$$\bar{A}' \begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = W \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} + \frac{1}{\delta}\bar{H}' \begin{bmatrix} \xi_t \\ \Xi_{t-1} \end{bmatrix}. \tag{2.6}$$

They can be combined with the model equations (2.1) to get a system of $2(n_X + n_x) + n_i$ difference equations for $t \geq 0$,

$$\begin{bmatrix} \bar{H} & 0 \\ 0 & \bar{A}' \end{bmatrix} \left[\begin{array}{c} X_{t+1} \\ x_{t+1|t} \\ i_{t+1|t} \\ \hline \xi_{t+1|t} \\ \Xi_t \end{array}\right] = \begin{bmatrix} \bar{A} & 0 \\ W & \frac{1}{\delta}\bar{H}' \end{bmatrix} \left[\begin{array}{c} X_t \\ x_t \\ i_t \\ \hline \xi_t \\ \Xi_{t-1} \end{array}\right] + \left[\begin{array}{c} C \\ 0 \\ \hline 0 \\ 0 \end{array}\right] \varepsilon_{t+1}. \tag{2.7}$$

Here, X_t and Ξ_{t-1} are predetermined variables ($n_X + n_x$ in total), and x_t, i_t, and ξ_t are nonpredetermined variables ($n_x + n_i + n_X$ in total).

In order to apply the algorithm in appendix A, it is practical to put the $n_{\tilde{X}} \equiv n_X + n_x$ predetermined and $n_{\tilde{x}} \equiv n_x + n_i + n_X$ nonpredetermined variables together as $\tilde{X}_t$ and $\tilde{x}_t$, respectively, where

$$\tilde{X}_t \equiv \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix}, \qquad \tilde{x}_t \equiv \begin{bmatrix} x_t \\ i_t \\ \xi_t \end{bmatrix}. \tag{2.8}$$

Then the system can be written as

$$\tilde{H} \begin{bmatrix} \tilde{X}_{t+1} \\ \tilde{x}_{t+1|t} \end{bmatrix} = \tilde{A} \begin{bmatrix} \tilde{X}_t \\ \tilde{x}_t \end{bmatrix} + \begin{bmatrix} \tilde{C} \\ 0 \end{bmatrix} \varepsilon_{t+1}, \tag{2.9}$$

where

$$\tilde{H} \equiv \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & A'_{22} & 0 & 0 & A'_{12} \\ 0 & 0 & H & 0 & 0 \\ 0 & A'_{21} & 0 & 0 & A'_{11} \\ 0 & B'_2 & 0 & 0 & B'_1 \end{bmatrix}, \qquad \tilde{A} \equiv \begin{bmatrix} A_{11} & 0 & A_{12} & B_1 & 0 \\ W_{xX} & H'/\delta & W_{xx} & W_{xi} & 0 \\ A_{21} & 0 & A_{22} & B_2 & 0 \\ W_{XX} & 0 & W_{Xx} & W_{Xi} & I/\delta \\ W_{iX} & 0 & W_{ix} & W_{ii} & 0 \end{bmatrix}, \qquad \tilde{C} = \begin{bmatrix} C \\ 0 \end{bmatrix}.$$

Partition $\tilde{H}$ and $\tilde{A}$ conformably with $\tilde{X}_t$ and $\tilde{x}$,

$$\tilde{H} \equiv \begin{bmatrix} \tilde{H}_{11} & \tilde{H}_{12} \\ \tilde{H}_{21} & \tilde{H}_{22} \end{bmatrix}, \qquad \tilde{A} \equiv \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix},$$

so

$$\tilde{H}_{11} \equiv \begin{bmatrix} I & 0 \\ 0 & A'_{22} \end{bmatrix}, \qquad \tilde{A}_{22} \equiv \begin{bmatrix} A_{22} & B_2 & 0 \\ W_{Xx} & W_{Xi} & I/\delta \\ W_{ix} & W_{ii} & 0 \end{bmatrix}.$$

Note that $\tilde{H}_{11}$ is invertible, since A_{22} is invertible. Then the top $n_{\tilde{X}}$ equations determine $\tilde{X}_{t+1}$, given $\tilde{X}_t$, $\tilde{x}_t$, and ε_{t+1}. The bottom $n_{\tilde{x}}$ equations determine $\tilde{x}_t$ for given $\tilde{X}_t$ and $\tilde{x}_{t+1|t}$. For this, $\tilde{A}_{22}$ must be invertible.

Under suitable assumptions (see appendix A and Klein (2000), Sims (2002b), and Söderlind (1999)) this system has a unique solution for $t \geq 0$, given X_0 and $\Xi_{-1} = 0$. The solution uses the generalized Schur decomposition. Klein (2000) provides a detailed discussion of how this solution method relates to those of Blanchard and Kahn (1980), Binder and Pesaran (1994) and Binder and

Pesaran (1997), King and Watson (1998) and King and Watson (2002), Sims (2002b), and Uhlig (1995).

The solution assumes the *saddlepoint property* emphasized by Blanchard and Kahn (1980): The number of generalized eigenvalues of the matrices $(\tilde{H}, \tilde{A})$ with modulus larger than unity equals the number of non-predetermined variables, $n_x + n_i + n_X$.[2] Then the solution can be written

$$
\begin{aligned}
x_t &= F_x \tilde{X}_t &(2.10)\\
i_t &= F_i \tilde{X}_t, &(2.11)\\
\tilde{X}_{t+1} &= M \tilde{X}_t + \tilde{C}\varepsilon_{t+1}, &(2.12)
\end{aligned}
$$

and where the matrices F_x, F_i, and M depend on A, B, H, D, Λ, and δ, but are independent of C. This demonstrates the certainty equivalence of the commitment solution: it is independent of the covariance matrix of the shocks to X_t, CC', and the same as when that covariance matrix is zero. There is also a solution for the forward-looking Lagrange multiplier ξ_t,

$$\xi_t = F_\xi \tilde{X}_t,$$

but this solution is not needed here. The matrix F_i can be called the *optimal policy function* or the *optimal reaction function.*

The submatrices of the matrix M, F_x, and F_i,

$$M \equiv \begin{bmatrix} M_{XX} & M_{X\Xi} \\ M_{\Xi X} & M_{\Xi\Xi} \end{bmatrix}, \qquad F_x \equiv \begin{bmatrix} F_{xX} & F_{x\Xi} \end{bmatrix}, \qquad F_i \equiv \begin{bmatrix} F_{iX} & F_{i\Xi} \end{bmatrix},$$

are related according to

$$
\begin{aligned}
M_{XX} &\equiv A_{11} + A_{12}F_{xX} + B_1 F_{iX},\\
M_{X\Xi} &\equiv A_{12}F_{x\Xi} + B_1 F_{i\Xi}.
\end{aligned}
$$

Note that, as is the case for non-predetermined variables, the one-period-ahead forecast errors of i_t and x_t are endogenous,

$$
\begin{aligned}
x_{t+1} - \mathrm{E}_t x_{t+1} &= F_x \tilde{C}\varepsilon_{t+1},\\
i_{t+1} - \mathrm{E}_t i_{t+1} &= F_i \tilde{C}\varepsilon_{t+1},
\end{aligned}
$$

since F_x and F_i are endogenous.

[2] The generalized eigenvalues of the matrices $(\tilde{H}, \tilde{A})$ are the complex numbers λ that satisfy $\det(\lambda\tilde{H} - \tilde{A}) = 0$ (see appendix A).

In a commitment equilibrium,

$$Y_t = D \begin{bmatrix} I & 0 \\ & F_x \\ & F_i \end{bmatrix} \tilde{X}_t \equiv \tilde{D}\tilde{X}_t,$$

$$L_t = \frac{1}{2} Y_t{}' \Lambda Y_t \equiv \frac{1}{2} \tilde{X}_t' \bar{W} \tilde{X}_t,$$

where $\tilde{D}$ is an $n_Y \times (n_X + n_x)$ matrix and $\bar{W} \equiv \tilde{D}' \Lambda \tilde{D}$ is an $(n_X + n_x) \times (n_X + n_x)$ matrix.

The equilibrium loss in any period $t \geq 0$ satisfies

$$\mathrm{E}_t \sum_{\tau=0}^{\infty} (1-\delta)\delta^{\tau} L_{t+\tau} = \frac{1}{2}[(1-\delta)\tilde{X}_t' V \tilde{X}_t + \delta w], \tag{2.13}$$

where V is an $(n_X + n_x) \times (n_X + n_x)$ matrix and w a scalar. The equilibrium loss satisfies the Bellman equation,

$$(1-\delta)\tilde{X}_t' V \tilde{X}_t + \delta w = (1-\delta)\tilde{X}_t' \bar{W} \tilde{X}_t + \delta \mathrm{E}_t[(1-\delta)\tilde{X}_{t+1}' V \tilde{X}_{t+1} + \delta w].$$

From this and (2.12) follows that V satisfies the Lyapunov equation,

$$V = \bar{W} + \delta M' V M, \tag{2.14}$$

and w satisfies

$$w = \mathrm{tr}\left\{ V \tilde{C} \tilde{C}' \right\}. \tag{2.15}$$

Furthermore, from (2.13) and (2.15) it follows that

$$\lim_{\delta \to 1} \mathrm{E}_t \sum_{\tau=0}^{\infty} (1-\delta)\delta^{\tau} L_{t+\tau} = \frac{1}{2} w = \frac{1}{2} \mathrm{tr}\left\{ V \tilde{C} \tilde{C}' \right\}.$$

2.2 Commitment in a timeless perspective

Suppose the commitment is not made in period 0 but far into the past, or alternatively, that any commitment in any period t is restricted as if it had been made far into the past. This kind of commitment has been called a "commitment in a timeless perspective" by Woodford, cf. Svensson and Woodford Svensson and Woodford (2005). Then the condition (2.5) no longer applies, and the first-order condition (2.3) and the solution (2.10)-(2.12) holds for all $t = \ldots -1, 0, 1, \ldots$ As noted by Svensson and Woodford Svensson and Woodford (2005), a simple way of finding the solution for commitment in a timeless perspective is to add the term $\frac{1-\delta}{\delta} \Xi_{t-1}' H x_t$ to the commitment problem in period t, where Ξ_{t-1} is the Lagrange multiplier of the equations for the forward-looking variables from the optimization in period $t-1$. Then, the optimization problem in period t has the intertemporal loss function,

$$\mathrm{E}_t \sum_{\tau=0}^{\infty}(1-\delta)\delta^{\tau} L_{t+\tau} + \frac{1-\delta}{\delta}\Xi_{t-1}' H x_t.$$

When this term is added, optimization under discretion in each period also results in the solution for commitment in a timeless perspective. This term is also related to the recursive saddlepoint method of Marcet and Marimon.

2.3 The recursive saddlepoint method of Marcet and Marimon

The problem to minimize (1.6) subject to (1.1) and (1.4)-(1.5) is not recursive, so the practical dynamic-programming method cannot be used directly. The reason the problem is not recursive is that the forward-looking variables, x_t, depend on expected future forward-looking variables, (1.3). The recursive saddlepoint method of Marcet and Marimon (1998, 2019) provides a simple way to reformulate the problem as a dual recursive saddlepoint problem, so dynamic programming can be applied. The dual problem is then, except for being a saddlepoint problem, isomorphic to the standard backward-looking linear-quadratic problem—the stochastic optimal linear-quadratic regulator (LQR) problem (Ljungqvist and Sargent (2004))—and the standard solution to the LQR problem can be applied.

Rewrite the Lagrangian as

$$\begin{aligned}\mathcal{L}_0 \;\; = \;\; & \mathrm{E}_0 \sum_{t=0}^{\infty}(1-\delta)\delta^t \left[\begin{array}{l} L_t + \Xi_t'(Hx_{t+1} - A_{21}X_t - A_{22}x_t - B_2 i_t) \\ +\,\xi_{t+1}'(X_{t+1} - A_{11}X_t - A_{12}x_t - B_1 i_t - C\varepsilon_{t+1}) \end{array}\right] \\ & + \frac{1-\delta}{\delta}\xi_0'(X_0 - \bar{X}_0).\end{aligned}$$

The reason why the problem is not recursive is that the term Hx_{t+1}, which is dated $t+1$, appears in the upper row of the Lagrangian. However, note that, because of (2.5), the discounted sum of the upper term in the Lagrangian can be written

$$\begin{aligned}&\sum_{t=0}^{\infty}(1-\delta)\delta^t[L_t + \Xi_t'(Hx_{t+1} - A_{21}X_t - A_{22}x_t - B_2 i_t)] \;\; = \\ &\sum_{t=0}^{\infty}(1-\delta)\delta^t[L_t + \Xi_t'(-A_{21}X_t - A_{22}x_t - B_2 i_t) + \frac{1}{\delta}\Xi_{t-1}' H x_t]. \qquad (2.16)\end{aligned}$$

Now all the terms within the bracket on the right side are dated t or earlier. The recursive saddlepoint method is in this case to let this term define the dual period loss. More precisely, the dual period loss is defined as

$$\begin{aligned}\tilde{L}_t &\equiv L_t + \gamma_t'(-A_{21}X_t - A_{22}x_t - B_2 i_t) + \frac{1}{\delta}\Xi_{t-1}' H x_t \\ &\equiv L_t + L_t^1 \\ &\equiv \tilde{L}(X_t, \Xi_{t-1}; x_t, i_t, \gamma_t),\end{aligned} \tag{2.17}$$

where Ξ_{t-1} is a new predetermined variable in period t and γ_t is introduced as a new control, where Ξ_{t-1} and γ_t are related by the dynamic equation,

$$\Xi_t = \gamma_t. \tag{2.18}$$

The problem can then be reformulated as the recursive dual saddlepoint problem,

$$\max_{\{\gamma_t\}_{t\geq 0}} \min_{\{x_t, i_t\}_{t\geq 0}} \mathrm{E}_0 \sum_{t=0}^{\infty} (1-\delta)\delta^t \tilde{L}_t, \tag{2.19}$$

where the optimization is subject to (1.2), (2.18), and to X_0 and $\Xi_{-1} = 0$ given. The value function for the saddlepoint problem, starting in any period t, satisfies the Bellman equation

$$\tilde{V}(X_t, \Xi_{t-1}) \equiv \max_{\gamma_t} \min_{(x_t, i_t)} \{(1-\delta)\tilde{L}(X_t, \Xi_{t-1}; x_t, i_t, \gamma_t) + \delta \mathrm{E}_t \tilde{V}(X_{t+1}, \Xi_t)\},$$

subject to (1.2) and (2.18).

Define

$$\tilde{\imath}_t = \begin{bmatrix} x_t \\ i_t \\ \gamma_t \end{bmatrix},$$

and define $\tilde{W}$, $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ such that

$$\tilde{L}_t = \frac{1}{2}\begin{bmatrix} \tilde{X}_t \\ \tilde{\imath}_t \end{bmatrix}' \tilde{W} \begin{bmatrix} \tilde{X}_t \\ \tilde{\imath}_t \end{bmatrix}, \tag{2.20}$$

$$\tilde{X}_{t+1} = \tilde{A}\tilde{X}_t + \tilde{B}\tilde{\imath}_t + \tilde{C}\varepsilon_{t+1}, \tag{2.21}$$

where $\tilde{X}_t$ is defined as in (2.8). Then, $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ satisfy

$$\tilde{A} \equiv \begin{bmatrix} A_{11} & 0 \\ 0 & 0 \end{bmatrix}, \qquad \tilde{B} \equiv \begin{bmatrix} A_{12} & B_1 & 0 \\ 0 & 0 & I \end{bmatrix}, \qquad \tilde{C} \equiv \begin{bmatrix} C \\ 0 \end{bmatrix}$$

and $\tilde{W}$ satisfies

$$\tilde{W} = \begin{bmatrix} W_{XX} & 0 & W_{Xx} & W_{Xi} & -A'_{21} \\ 0 & 0 & \frac{1}{\delta}H & 0 & 0 \\ W'_{Xx} & \frac{1}{\delta}H' & W_{xx} & W_{xi} & -A'_{22} \\ W'_{Xi} & 0 & W'_{xi} & W_{ii} & -B'_2 \\ -A_{21} & 0 & -A_{22} & -B_2 & 0 \end{bmatrix},$$

where W is partitioned conformably with X_t, x_t, and i_t according to

$$W \equiv \begin{bmatrix} W_{XX} & W_{Xx} & W_{Xi} \\ W'_{Xx} & W_{xx} & W_{xi} \\ W'_{Xi} & W'_{xi} & W_{ii} \end{bmatrix}.$$

The problem (2.19) subject to (2.21) and given $\tilde{X}_t$ is obviously isomorphic to the stochastic LQR problem (Anderson et al. (1995), Ljungqvist and Sargent (2004)), except being a saddlepoint problem. However, the saddlepoint aspect does not affect the first-order conditions. It is easy to show that the first-order conditions of the saddlepoint problem are identical to those of the original problem.

Hence, I can use the standard solution for the LQR problem: The value function for the saddlepoint problem will be quadratic,

$$\tilde{V}(\tilde{X}_t) \equiv \frac{1}{2}[(1-\delta)\tilde{X}'_t\tilde{V}\tilde{X}_t + \delta\tilde{w}],$$

and so the Bellman equation can be written,

$$(1-\delta)\tilde{X}'_t\tilde{V}\tilde{X}_t + \delta\tilde{w} = (1-\delta)\max_{\gamma_t}\min_{(x_t,i_t)}\{\begin{bmatrix} \tilde{X}_t \\ \tilde{\imath}_t \end{bmatrix}'\tilde{W}\begin{bmatrix} \tilde{X}_t \\ \tilde{\imath}_t \end{bmatrix} + \delta(\mathrm{E}_t\tilde{X}'_{t+1}\tilde{V}\tilde{X}_{t+1} + \frac{\delta}{1-\delta}\tilde{w})\} \quad (2.22)$$

subject to (2.21).

The first-order condition with respect to $\tilde{\imath}_t$ is

$$J\tilde{\imath}_t + K\tilde{X}_t = 0,$$

where the matrices J and K are defined as

$$\begin{aligned} J &\equiv R + \delta\tilde{B}'\tilde{V}\tilde{B}, \\ K &\equiv N' + \delta\tilde{B}'\tilde{V}\tilde{A}, \end{aligned}$$

where

$$\tilde{W} \equiv \begin{bmatrix} Q & N \\ N' & R \end{bmatrix},$$

is partitioned conformably with $\tilde{X}_t$ and $\tilde{\imath}_t$, so

$$Q \equiv \begin{bmatrix} W_{XX} & 0 \\ 0 & 0 \end{bmatrix}, \qquad N \equiv \begin{bmatrix} W_{Xx} & W_{Xi} & -A_{21}' \\ \frac{1}{\delta}H & 0 & 0 \end{bmatrix}, \qquad R \equiv \begin{bmatrix} W_{xx} & W_{xi} & -A_{22}' \\ W_{xi}' & W_{ii} & -B_2' \\ -A_{22} & -B_2 & 0 \end{bmatrix}.$$

It follows that the solution for $\tilde{\imath}_t$ can be written

$$\tilde{\imath}_t = F\tilde{X}_t, \tag{2.23}$$

where

$$F \equiv -J^{-1}K. \tag{2.24}$$

Using (2.23) and (2.24) in (2.22) results in the Riccati equation,

$$\tilde{V} = Q + \delta\tilde{A}'\tilde{V}\tilde{A} - K'J^{-1}K.$$

Thus, the solution F can be found by first solving the Riccati equation for $\tilde{V}$ and then using (2.24).

The matrix F provides the solution not only to the saddlepoint problem but also to the original problem. The equilibrium dynamics will then be given by

$$\begin{aligned} \tilde{X}_{t+1} &= M\tilde{X}_t + \tilde{C}\varepsilon_{t+1}, \\ x_t &= F_x\tilde{X}_t, \\ i_t &= F_i\tilde{X}_t, \\ L_t &= \frac{1}{2}\tilde{X}_t'\bar{W}\tilde{X}_t, \end{aligned} \tag{2.25}$$

where

$$M \equiv \tilde{A} + \tilde{B}F,$$

the matrix F is partitioned conformably with x_t, i_t, and γ_t,

$$F \equiv \begin{bmatrix} F_x \\ F_i \\ F_\gamma \end{bmatrix},$$

and

$$\bar{W} \equiv \begin{bmatrix} I & 0 \\ & F_x \\ & F_i \end{bmatrix}' W \begin{bmatrix} I & 0 \\ & F_x \\ & F_i \end{bmatrix}.$$

Whereas the solution to the saddlepoint problem in the form of the matrix F also is the solution to the original problem, the value function of the saddlepoint problem in the form of the matrix $\tilde{V}$ and the scalar $\tilde{w}$ does not directly provide the value function of the original problem. This is because the period loss for the saddlepoint problem, $\tilde{L}_t$, differs from the period loss for the original problem, L_t. Indeed, the matrix $\tilde{V}$ is not positive semidefinite.

In order to find the value function for original problem, I decompose the value function of the saddlepoint problem according to

$$\frac{1}{2}[(1-\delta)\tilde{X}_t'\tilde{V}\tilde{X}_t + \delta\tilde{w})] \equiv \frac{1}{2}[(1-\delta)\tilde{X}_t'V\tilde{X}_t + \delta w] + \frac{1}{2}[(1-\delta)\tilde{X}_t'V^1\tilde{X}_t + \delta w^1],$$

where

$$\frac{1}{2}[(1-\delta)\tilde{X}_t'V\tilde{X}_t + \delta w] \equiv \mathrm{E}_t\sum_{\tau=0}^{\infty}(1-\delta)\delta^\tau\frac{1}{2}\tilde{X}_{t+\tau}'\bar{W}\tilde{X}_{t+\tau} = \mathrm{E}_t\sum_{\tau=0}^{\infty}(1-\delta)\delta^\tau L_{t+\tau},$$

is the value function for the original problem starting in period t with $\tilde{X}_t$ given. The value function for the original problem will satisfy the Bellman equation,

$$\frac{1}{2}[(1-\delta)\tilde{X}_t'V\tilde{X}_t + \delta w] \equiv \frac{1}{2}(1-\delta)\tilde{X}_t'\bar{W}\tilde{X}_t + \frac{1}{2}\delta\mathrm{E}_t[(1-\delta)\tilde{X}_{t+1}'V\tilde{X}_{t+1} + \delta w],$$

and the matrix V will satisfy the Lyapunov equation,

$$V = \bar{W} + \delta M'VM,$$

and be positive semidefinite. The scalar w will satisfy

$$w = \mathrm{tr}(V\tilde{C}\tilde{C}').$$

However, the matrix V and the scalar w can be found in a more direct way from the matrix $\tilde{V}$ and scalar $\tilde{w}$. Note that, by (1.3), (2.17), and (2.18), the identity

$$\frac{1}{2}[(1-\delta)\tilde{X}_t'V\tilde{X}_t + \delta w] \equiv \frac{1}{2}[(1-\delta)\tilde{X}_t'\tilde{V}\tilde{X}_t + \delta\tilde{w}] - (1-\delta)\frac{1}{\delta}\Xi_{t-1}'HF_x\tilde{X}_t$$

must hold. That is,

$$\frac{1}{2}[(1-\delta)\tilde{X}_t'V^1\tilde{X}_t + \delta w^1] \equiv -(1-\delta)\frac{1}{\delta}\Xi_{t-1}'HF_x\tilde{X}_t \equiv -\frac{1}{2}(1-\delta)\frac{1}{\delta}(\Xi_{t-1}'HF_x\tilde{X}_t + \tilde{X}_t'F_x'H'\Xi_{t-1}).$$

Hence, identification of terms implies $w^1 \equiv 0$, so w and V are determined by

$$w = \tilde{w},$$

$$V = \tilde{V} - \begin{bmatrix} 0 & \frac{1}{\delta} F'_{xX} H' \\ \frac{1}{\delta} H F_{xX} & \frac{1}{\delta}(H F_{x\Xi} + F'_{x\Xi} H') \end{bmatrix} \equiv \begin{bmatrix} \tilde{V}_{XX} & \tilde{V}_{X\Xi} - \frac{1}{\delta} F'_{xX} H' \\ \tilde{V}_{\Xi X} - \frac{1}{\delta} H F_{xX} & \tilde{V}_{\Xi\Xi} - \frac{1}{\delta}(H F_{x\Xi} + F'_{x\Xi} H') \end{bmatrix},$$

where $\tilde{V}$ and F_x are partitioned conformably with X_t and Ξ_{t-1} as

$$\tilde{V} \equiv \begin{bmatrix} \tilde{V}_{XX} & \tilde{V}_{X\Xi} \\ \tilde{V}_{\Xi X} & \tilde{V}_{\Xi\Xi} \end{bmatrix}, \qquad F_x \equiv \begin{bmatrix} F_{xX} & F_{x\Xi} \end{bmatrix}.$$

In summary, the original problem is reformulated by incorporating the block of equations for the forward-looking variables, (1.3), in such a way that the resulting saddlepoint problem becomes recursive and isomorphic to the LQR problem. Then the solution to the LQR problem is applied to the original problem, the dynamics of the original problem is specified, and the Lyapunov function for the value function of the original problem is specified and solved (alternatively, the above identification procedure is used to find the value function of the original problem from the value function of the saddlepoint problem).

Appendix B shows that the recursive saddlepoint method can also be applied to problems that are not linear-quadratic.

From (2.16) and (1.3), it follows that

$$\mathrm{E}_t \sum_{\tau=0}^{\infty} (1-\delta)\delta^\tau \tilde{L}_{t+\tau} = \mathrm{E}_t \sum_{\tau=0}^{\infty} (1-\delta)\delta^\tau L_{t+\tau} + \frac{1-\delta}{\delta} \Xi'_{t-1} H x_t, \tag{2.26}$$

the intertemporal loss function for the dual problem equals the intertemporal loss for the original problem plus the second term on the right side of (2.26),

$$\frac{1-\delta}{\delta} \Xi'_{t-1} H x_t. \tag{2.27}$$

It follows that minimizing the right side of (2.26) under discretion will result in the optimal policy under commitment in a timeless perspective. In Svensson and Woodford (2005), a "commitment to continuity and predictability" is interpreted as a central bank optimizing under discretion but taking into account previous expectations and plans in the form of adding (2.27) to its intertemporal loss function. That is, such a commitment means that the central bank applies the appropriate shadow price vector $\frac{1-\delta}{\delta} \Xi'_{t-1}$ from the previous period's optimization to the linear combination $H x_t$ of the current period's forward-looking variables. Such a commitment to a modified loss function then implies that optimization under discretion results in the optimal policy under commitment in a timeless perspective.

2.3.1 Using the recursive saddlepoint method to solve linear difference equations with forward-looking variables

Note that the recursive saddlepoint method can be used to solve linear difference equations with forward-looking variables. Consider the system

$$\begin{aligned} X_{t+1} &= A_{11}X_t + A_{12}x_t + C\varepsilon_{t+1}, \qquad (2.28) \\ \mathrm{E}_t H x_{t+1} &= A_{21}X_t + A_{22}x_t, \qquad (2.29) \end{aligned}$$

and assume that it has a unique solution

$$x_t = F_x X_t. \qquad (2.30)$$

This solution can be found with the generalized Schur decomposition, as in Klein (2000) and demonstrated above.

The solution can also be found with the recursive saddlepoint problem. Let

$$L(X_t, x_t) \equiv \frac{1}{2} \begin{bmatrix} X_t \\ x_t \end{bmatrix}' W \begin{bmatrix} X_t \\ x_t \end{bmatrix},$$

where W is any positive semidefinite matrix, and let

$$\tilde{L}(\tilde{X}_t; \tilde{\imath}_t) \equiv L(X_t, x_t) + \gamma_t'(-A_{21}X_t - A_{22}x_t) + \frac{1}{\delta}\Xi_{t-1}' H x_t \equiv \frac{1}{2} \begin{bmatrix} \tilde{X}_t \\ \tilde{\imath}_t \end{bmatrix}' \tilde{W} \begin{bmatrix} \tilde{X}_t \\ \tilde{\imath}_t \end{bmatrix},$$

$$\tilde{X}_{t+1} = \tilde{A}\tilde{X}_t + \tilde{B}\tilde{\imath}_t + \tilde{C}\varepsilon_t,$$

where

$$\tilde{X}_t \equiv \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix}, \qquad \tilde{\imath}_t \equiv \begin{bmatrix} x_t \\ \gamma_t \end{bmatrix},$$

$$\tilde{W} = \begin{bmatrix} W_{XX} & 0 & W_{Xx} & -A_{21}' \\ 0 & 0 & \frac{1}{\delta}H & 0 \\ W_{Xx}' & \frac{1}{\delta}H' & W_{xx} & -A_{22}' \\ -A_{21} & 0 & -A_{22} & 0 \end{bmatrix},$$

where W is partitioned conformably with X_t and x_t according to

$$W \equiv \begin{bmatrix} W_{XX} & W_{Xx} \\ W_{Xx}' & W_{xx} \end{bmatrix}.$$

Then we can apply the recursive saddlepoint method as above. This will result in the solution

$$\tilde{\imath}_t \equiv \begin{bmatrix} x_t \\ \gamma_t \end{bmatrix} = \tilde{F}\tilde{X}_t \equiv \begin{bmatrix} \tilde{F}_{xX} & \tilde{F}_{x\Xi} \\ \tilde{F}_{\gamma X} & \tilde{F}_{\gamma\Xi} \end{bmatrix} \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix},$$

where

$$\tilde{F}_{xX} \equiv F_x, \qquad \tilde{F}_{x\Xi} \equiv 0. \tag{2.31}$$

Here (2.31) should be demonstrated in detail, but we realize that it must be true when there is a unique (nonbubble) solution (2.30).

Note that, since there are no degrees of freedom for x_t, the solution (2.31) for x_t does not depend on the matrix W. The solution for γ_t will depend on W.

3 Targeting rules and instrument rules

Practical definitions of instrument and targeting rules are developed in Svensson (1999), Rudebusch and Svensson (1999), Svensson and Woodford (2005), and Svensson (2003). A general derivation of a targeting rule in a linear-quadratic model is provided by Giannoni and Woodford (2003).

An *explicit instrument rule* expresses the instrument as function of current and lagged predetermined variables only. In the solution above, (2.11) is the *optimal* explicit instrument rule, also called the optimal policy function or the optimal reaction function. It can be written in three main ways: First, as in (2.11), it can be written as a function of X_t, the predetermined variables, and Ξ_{t-1}, the predetermined Lagrange multipliers of the lower block of (1.1), the block of equations determining the forward-looking variables. Second, since Ξ_{t-1} (in the timeless perspective) can be written as an infinite sum of lagged predetermined variables

$$\Xi_{t-1} = \sum_{\tau=0}^{\infty} M_{\Xi\Xi}{}^{\tau} M_{\Xi X} X_{t-1-\tau},$$

it follows that the explicit instrument rule can also be written as an infinite sum of lagged predetermined variables,

$$i_t = F_{iX} X_t + F_{i\Xi} \sum_{\tau=0}^{\infty} M_{\Xi\Xi}{}^{\tau} M_{\Xi X} X_{t-1-\tau}.$$

Third, since $\tilde{X}_t$ by (2.12) can be written as an infinite sum of lagged shocks,

$$\tilde{X}_t = \sum_{\tau=0}^{\infty} M^{\tau} \tilde{C} \varepsilon_{t-\tau},$$

the explicit instrument rule can also be written as such a sum,

$$i_t = F_i \sum_{\tau=0}^{\infty} M^\tau \tilde{C} \varepsilon_{t-\tau}.$$

An *implicit instrument rule* is an equilibrium condition that involves the instrument and forward-looking variables; it may or may not also involve predetermined variables. Any implicit instrument rule consistent with a given equilibrium is not unique. For the solution above, for any arbitrary $n_i \times n_x$ matrix G, by (2.10), I can construct a nonunique *optimal* implicit instrument rule in the following way,

$$i_t = F_i \tilde{X}_t = F_i \tilde{X}_t + G(x_t - F_x \tilde{X}_t) = (F_i - GF_x)\tilde{X}_t + Gx_t.$$

The different forms of instrument rules may have different determinacy properties, see Svensson and Woodford (2005).

A *targeting rule* is an equilibrium condition involving the target variables only, including forecasts (expected leads) and lags thereof. The optimal targeting rule is the first-order condition for an optimum expressed in terms of forecasts, current values, and lags of the target variables only.[3] The first-order condition (2.6) can be written

$$A' \begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = \frac{1}{\delta} \begin{bmatrix} I & 0 \\ 0 & H' \end{bmatrix} \begin{bmatrix} \xi_t \\ \Xi_{t-1} \end{bmatrix} + [D_X \ D_x]' \Lambda Y_t, \tag{3.1}$$

$$B' \begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = D_i' \Lambda Y_t, \tag{3.2}$$

where the matrix $D \equiv [D_X \ D_x \ D_i]$ is partitioned conformably with X_t, x_t, and i_t.

Consider the $n_X + n_x$ equations in (3.1) as a system of difference equations for the forward-looking variables ξ_t and the predetermined variables Ξ_t, for given realizations of the target variables Y_t. Under suitable assumptions, there exist a solution to this system of the form

$$\xi_t = g_1 \Xi_{t-1} + g_2 \sum_{\tau=0}^{\infty} P_\tau [D_X \ D_x]' \Lambda Y_{t+\tau|t} \equiv g_1 \Xi_{t-1} + g_2 \mathrm{E}_t P(\mathrm{L}^{-1}) [D_X \ D_x]' \Lambda Y_t, \tag{3.3}$$

$$\Xi_t = m_1 \Xi_{t-1} + m_2 \sum_{\tau=0}^{\infty} P_\tau [D_X \ D_x]' \Lambda Y_{t+\tau|t} \equiv m_1 \Xi_{t-1} + m_2 \mathrm{E}_t P(\mathrm{L}^{-1}) [D_X \ D_x]' \Lambda Y_t, \tag{3.4}$$

where the matrices g_1, g_2, m_1, m_2, and $\{P_\tau\}_{\tau=0}^{\infty}$ depend on A, H, and δ, and where $P(\mathrm{L}^{-1}) \equiv \sum_{\tau=0}^{\infty} P_\tau \mathrm{L}^{-\tau}$ denotes a lead polynomial, with L the lag operator and L^{-1} the lead operator, so

[3] This derivation of the targeting rule follows Giannoni and Woodford (2003), except that they restrict the submatrix B_1 of B in (2.2) to be zero.

$\mathrm{L}^{-\tau} Y_t \equiv Y_{t+\tau}$. Express the solution as

$$\begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = \Psi \begin{bmatrix} \xi_{t|t-1} \\ \Xi_{t-1} \end{bmatrix} + \mathrm{E}_t Q(\mathrm{L}^{-1})[D_X \;\; D_x]' \Lambda Y_t, \tag{3.5}$$

where the square matrix Ψ is given by

$$\Psi \equiv \begin{bmatrix} 0 & g_1 m_1 \\ 0 & m_1 \end{bmatrix}, \tag{3.6}$$

and $\{Q_\tau\}_{\tau=0}^{\infty}$ in $Q(\mathrm{L}^{-1}) \equiv \sum_{\tau=0}^{\infty} Q_\tau \mathrm{L}^{-\tau}$ depend on g_1, g_2, m_2, and $\{P_\tau\}_{\tau=0}^{\infty}$.

Furthermore, for any square matrix Ψ, there exist a matrix polynomial $\Phi(\mathrm{L})$ of the same dimension and a scalar polynomial $\alpha(\mathrm{L})$ such that

$$\Phi(\mathrm{L})(I - \Psi \mathrm{L}) = \alpha(\mathrm{L}) I.$$

Here, $\alpha(\mathrm{L}) \equiv \det(I - \Psi \mathrm{L})$, and $\Phi(\mathrm{L})$ is the adjoint of $I - \Psi \mathrm{L}$, that is, the transpose of the matrix of cofactors of $I - \Psi \mathrm{L}$. The order of $\alpha(\mathrm{L})$ is $\mathrm{rank}(\Psi) \leq n_x$. The dimension of Ψ and $\Phi(\mathrm{L})$ is $n_X + n_x$, and the order of $\Phi(\mathrm{L})$ is $\min[\mathrm{rank}(\Psi), n_X + n_x - 1]$.

The solution (3.5) can be written

$$(I - \Psi \mathrm{L}) \begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = \mathrm{E}_t Q(\mathrm{L}^{-1})[D_X \;\; D_x]' \Lambda Y_t.$$

Premultiplying this by $\Phi(\mathrm{L})$ gives

$$\alpha(\mathrm{L}) \begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = \Phi(\mathrm{L}) \mathrm{E}_t Q(\mathrm{L}^{-1})[D_X \;\; D_x]' \Lambda Y_t.$$

Since $\alpha(\mathrm{L})$ is a scalar polynomial, premultiplying this expression by B' and using (3.2) result in

$$B' \Phi(\mathrm{L}) \mathrm{E}_t Q(\mathrm{L}^{-1})[D_X \;\; D_x]' \Lambda Y_t - \alpha(\mathrm{L}) D_i' \Lambda Y_t = 0. \tag{3.7}$$

Thus, this is the general form of the targeting rule, the first-order condition (3.1) and (3.2), where the Lagrange multipliers have been eliminated. It is a condition in terms of forecasts and lags of the target variables only. Although it looks complicated in the general form, in any given model, it is often quite simple, see Svensson (2003).

If $D_i \neq 0$ and $D_i' \Lambda \neq 0$, the target variables depend directly on the the instruments, or the instruments are among the target variables, and the second term in (3.7) enters in the targeting rule. If $D_i = 0$ or $D_i' \Lambda = 0$, the target variables do not depend directly on the instrument and the instruments are not among the target variables. Then the second term in the targeting rule vanishes, and it is

$$B'\Phi(\mathrm{L})\mathrm{E}_t Q(\mathrm{L}^{-1})[D_X \;\; D_x]'\Lambda Y_t = 0. \tag{3.8}$$

As explained in Svensson (2003), the optimal targeting rule (3.7) corresponds to the equality of the marginal rates of substitution and marginal rates of transformation between the target variables,

$$\mathrm{MRS} = \mathrm{MRT}.$$

The marginal rates of substitution between the target variables follow from the loss function, (1.5) and (1.6); the marginal rates of transformation follow from the model equations, (1.1) and (1.4).

The optimal targeting rule provides n_i equations, the same number as the number of instruments, regardless of whether the instruments are among the targeting rules or not. They can be interpreted as n_i equations to be added to the $n_X + n_x$ model equations, (1.1), so as to determine the $n_X + n_x$ predetermined and forward-looking variables and the n_i instruments.

Consider a special case, where there is only one instrument ($n_i = 1$), and where

$$Y_t \equiv \begin{bmatrix} \bar{Y}_t \\ i_t \end{bmatrix} = \begin{bmatrix} \bar{D}_X & \bar{D}_x & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} = \begin{bmatrix} \bar{D}_X X_t + \bar{D}_x x_t \\ i_t \end{bmatrix}.$$

That is, the target variables consist of one set of variables, $\bar{Y}_t$, and the instrument, i_t, and $\bar{Y}_t$ does not depend directly on the instrument. Call $\bar{Y}_t$ the non-instrument target variables. Furthermore, suppose the weighting matrix is

$$\Lambda \equiv \begin{bmatrix} \bar{\Lambda} & 0 \\ 0 & \lambda_i \end{bmatrix},$$

so the period loss function is additively separable in the non-instrument target variables and the instrument,

$$L_t = \frac{1}{2}(\bar{Y}_t'\bar{\Lambda}\bar{Y}_t + \lambda_i i_t^2).$$

Then,

$$\begin{aligned} [D_X \;\; D_x]'\Lambda Y_t &= [\bar{D}_X \;\; \bar{D}_x]'\bar{\Lambda}\bar{Y}_t, \\ D_i'\Lambda Y_t &= \lambda_i i_t. \end{aligned}$$

It follows that (3.7) can be written

$$\alpha(\mathrm{L})\lambda_i i_t = B'\Phi(\mathrm{L})\mathrm{E}_t Q(\mathrm{L}^{-1})[\bar{D}_X \;\; \bar{D}_x]'\bar{\Lambda}\bar{Y}_t. \tag{3.9}$$

This can be seen as an implicit instrument rule, where the instrument in period t is related to lagged values of the instrument and to forecasts, current values, and lags of the non-instrument target variables. But (3.9) is fundamentally a targeting rule, in the sense that the instrument enters there only because it is a target variable. Suppose that $\lambda_i = 0$ (corresponding to $D_i'\Lambda = 0$), so the instrument is no longer (effectively) a target variable. Then, the instrument vanishes from (3.9), and it becomes

$$B'\Phi(\mathrm{L})\mathrm{E}_t Q(\mathrm{L}^{-1})[\bar{D}_X \ \bar{D}_x]'\bar{\Lambda}\bar{Y}_t = 0,$$

identical to (3.8), and involves only the non-instrument target variables.

3.1 Backward-looking model

Consider a backward-looking model, where (1.1) and (1.4) are simplified to

$$\begin{aligned} X_{t+1} &= AX_t + Bi_t + C\varepsilon_{t+1}, \\ Y_t &= [D_X \ D_i]\begin{bmatrix} X_t \\ i_t \end{bmatrix}. \end{aligned}$$

That is, there are no forward-looking variables, so $A \equiv A_{11}$, $B \equiv B_1$, and $D \equiv [D_X \ D_i]$.

For such a model, the first-order conditions (3.1) and (3.2) simplify to

$$A'\xi_{t+1|t} = \frac{1}{\delta}\xi_t + D_X{}'\Lambda Y_t, \tag{3.10}$$

$$B'\xi_{t+1|t} = D_i'\Lambda Y_t. \tag{3.11}$$

The first-order conditions (3.10) (those with respect to X_t) must fulfill,

$$\begin{aligned} \xi_t &= -\delta D_X{}'\Lambda Y_t + \delta A'\xi_{t+1|t} \\ &= -\sum_{\tau=0}^{T-1}(\delta A')^\tau \delta D_X'\Lambda Y_{t+\tau|t} + (\delta A')^T \xi_{t+T|t} \\ &= -\sum_{\tau=0}^{\infty}(\delta A')^\tau \delta D_X'\Lambda Y_{t+\tau|t} \ \equiv \ g_2 \mathrm{E}_t P(L^{-1})D_X'\Lambda Y_t \end{aligned}$$

where the conditions

$$\begin{aligned} \lim_{T\to\infty}(\delta A')^T \xi_{t+T|t} &= 0, \\ \lim_{T\to\infty}\left|\sum_{\tau=0}^{T-1}(\delta A')^\tau \delta D_X'\Lambda Y_{t+\tau|t}\right| &< \infty, \end{aligned}$$

should be fulfilled. It follows that in (3.3) $g_1 \equiv 0$ and that $g_2 P(L^{-1})$ can be identified as

$$g_2 P(L^{-1}) \equiv -\sum_{\tau=0}^{\infty} (\delta A' L^{-1})^\tau \delta.$$

It follows that $\xi_{t+1|t}$ must fulfill

$$\xi_{t+1|t} = -\sum_{\tau=0}^{\infty} (\delta A')^\tau \delta D_X' \Lambda Y_{t+1+\tau|t} \;\; = \;\; -\mathrm{E}_t \sum_{\tau=0}^{\infty} L^{-1} (\delta A' L^{-1})^\tau \delta D_X' \Lambda Y_t. \tag{3.12}$$

Hence, $Q(L^{-1})$ in (3.5) can be identified as

$$Q(L^{-1}) \equiv -\sum_{\tau=0}^{\infty} L^{-1} (\delta A' L^{-1})^\tau \delta.$$

From the first-order conditions (3.11) (those with respect to i_t), it then follows that

$$B' \xi_{t+1|t} = -B' \sum_{\tau=0}^{\infty} (\delta A')^\tau \delta D_X' \Lambda Y_{t+1+\tau|t} = D_i' \Lambda Y_t.$$

Hence, the optimal targeting rule (3.7) can, for a backward-looking model, be written

$$B' \sum_{\tau=0}^{\infty} (\delta A')^\tau \delta D_X' \Lambda Y_{t+1+\tau|t} + D_i' \Lambda Y_t = 0.$$

4 Optimization under discretion: The discretion equilibrium

Let (1.1) be the model equations. Let the period loss function be written as

$$L_t = \frac{1}{2} \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}' W \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}, \tag{4.1}$$

and let the intertemporal loss function in period t be

$$\mathrm{E}_t \sum_{\tau=0}^{\infty} (1-\delta) \delta^\tau L_{t+\tau},$$

where $0 < \delta < 1$. Consider the decision problem to choose i_t in period t to minimize the intertemporal loss function under discretion, that is, subject to (1.1), X_t given, and

$$i_{t+1} = F_{t+1} X_{t+1} \tag{4.2}$$

$$x_{t+1} = G_{t+1} X_{t+1}, \tag{4.3}$$

where F_{t+1} and G_{t+1} are determined by the decision problem in period $t+1$. Both F_{t+1} and G_{t+1} are assumed known in period t; only G_{t+1} will matter for the decision problem in period t.

Oudiz and Sachs (1985) derive an algorithm for the solution of this problem (with $H = I$), which is further discussed in Backus and Driffill (1986), Currie and Levine (1993), and Söderlind (1999).

First, take the expectation in period t of (1.1),

$$\begin{bmatrix} X_{t+1|t} \\ Hx_{t+1|t} \end{bmatrix} = A \begin{bmatrix} X_t \\ x_t \end{bmatrix} + Bi_t. \tag{4.4}$$

Second, using (4.3) and the upper block of (4.4) results in

$$x_{t+1|t} = G_{t+1}X_{t+1|t} = G_{t+1}(A_{11}X_t + A_{12}x_t + B_1 i_t). \tag{4.5}$$

The lower block of (1.1) is

$$Hx_{t+1|t} = A_{21}X_t + A_{22}x_t + B_2 i_t. \tag{4.6}$$

Multiplying (4.5) by H, setting the result equal to (4.6), and solving for x_t gives

$$x_t = \bar{A}_t X_t + \bar{B}_t i_t, \tag{4.7}$$

where

$$\bar{A}_t \equiv (A_{22} - HG_{t+1}A_{12})^{-1}(HG_{t+1}A_{11} - A_{21}), \tag{4.8}$$

$$\bar{B}_t \equiv (A_{22} - HG_{t+1}A_{12})^{-1}(HG_{t+1}B_1 - B_2) \tag{4.9}$$

(I assume that $A_{22} - HG_{t+1}A_{12}$ is invertible). Using (4.7) in the upper block of (1.1) then gives

$$X_{t+1} = \tilde{A}_t X_t + \tilde{B}_t i_t + C\varepsilon_{t+1}, \tag{4.10}$$

where

$$\tilde{A}_t \equiv A_{11} + A_{12}\bar{A}_t, \tag{4.11}$$

$$\tilde{B}_t \equiv B_1 + A_{12}\bar{B}_t. \tag{4.12}$$

Third, using (4.7) in (4.1) leads to

$$L_t = \frac{1}{2}\begin{bmatrix} X_t \\ i_t \end{bmatrix}' \begin{bmatrix} Q_t & N_t \\ N_t' & R_t \end{bmatrix} \begin{bmatrix} X_t \\ i_t \end{bmatrix}, \tag{4.13}$$

where

$$Q_t \equiv W_{XX} + W_{Xx}\bar{A}_t + \bar{A}_t' W_{Xx}' + \bar{A}_t' W_{xx}\bar{A}_t, \tag{4.14}$$

$$N_t \equiv W_{Xx}\bar{B}_t + \bar{A}_t' W_{xx}\bar{B}_t + W_{Xi} + \bar{A}_t' W_{xi}, \tag{4.15}$$

$$R_t \equiv W_{ii} + \bar{B}_t' W_{xx}\bar{B}_t + \bar{B}_t' W_{xi} + W_{xi}'\bar{B}_t. \tag{4.16}$$

Fourth, since the loss function is quadratic and the constraints are linear, it follows that the optimal value of the problem will be quadratic. In period $t+1$ the optimal value will depend on X_{t+1} and can hence be written $\frac{1}{2}[(1-\delta)X'_{t+1}V_{t+1}X_{t+1}+\delta w_{t+1}]$, where V_{t+1} is a positive semidefinite matrix and w_{t+1} is a scalar independent of X_{t+1}. Both V_{t+1} and w_{t+1} are assumed known in period t. Then the optimal value of the problem in period t is associated with the positive semidefinite matrix V_t and the scalar w_t, and satisfies the Bellman equation

$$\frac{1}{2}[(1-\delta)X'_tV_tX_t+\delta w_t] \equiv (1-\delta)\min_{i_t}\left\{L_t+\delta \mathrm{E}_t\frac{1}{2}[X'_{t+1}V_{t+1}X_{t+1}+\frac{\delta}{1-\delta}w_{t+1}]\right\}, \tag{4.17}$$

subject to (4.10) and (4.13). Indeed, the problem has been transformed to a standard LQR problem without forward-looking variables, albeit in terms of X_t and with time-varying parameters. The first-order condition is, by (4.13) and (4.17),

$$\begin{aligned} 0 &= X'_tN_t+i'_tR_t+\delta\mathrm{E}_t[X'_{t+1}V_{t+1}\tilde{B}_t] \\ &= X'_tN_t+i'_tR_t+\delta(X'_t\tilde{A}'_t+i'_t\tilde{B}'_t)V_{t+1}\tilde{B}_t. \end{aligned}$$

The first-order condition can be solved for the reaction function

$$i_t = F_tX_t, \tag{4.18}$$

where

$$F_t \equiv -(R_t+\delta\tilde{B}'_tV_{t+1}\tilde{B}_t)^{-1}(N'_t+\delta\tilde{B}'_tV_{t+1}\tilde{A}_t) \tag{4.19}$$

(I assume that $R_t+\delta\tilde{B}'_tV_{t+1}\tilde{B}_t$ is invertible). Using (4.18) in (4.7) gives

$$x_t = G_tX_t,$$

where

$$G_t \equiv \bar{A}_t+\bar{B}_tF_t. \tag{4.20}$$

Furthermore, using (4.18) in (4.17) and identifying terms result in

$$V_t \equiv Q_t+N_tF_t+F'_tN'_t+F'_tR_tF_t+\delta(\tilde{A}_t+\tilde{B}_tF_t)'V_{t+1}(\tilde{A}_t+\tilde{B}_tF_t). \tag{4.21}$$

Finally, the above equations ((4.8), (4.9), (4.11), (4.12), (4.14)–(4.16), (4.19), (4.20), and (4.21)) define a mapping from (G_{t+1},V_{t+1}) to (G_t,V_t), which also determines F_t. The solution to the problem is a fixed point (G,V) of the mapping and a corresponding F. It can be obtained as the limit of (G_t,V_t) when $t\to-\infty$. The solution thus satisfies the corresponding steady-state matrix equations.

Thus, the instrument i_t and the forward-looking variables x_t will be linear functions,

$$\begin{aligned} i_t &= F X_t, && (4.22) \\ x_t &= G X_t, && (4.23) \end{aligned}$$

where the corresponding F and G satisfy the corresponding steady-state equations. The matrix F can be called the equilibrium policy function or the equilibrium reaction function. The resulting equation for X_t is

$$X_{t+1} = M X_t + C\varepsilon_{t+1}, \tag{4.24}$$

where

$$M \equiv \tilde{A} + \tilde{B} F$$

where $\tilde{A}$ and $\tilde{B}$ is the fixed point of the mapping from $(\tilde{A}_{t+1}, \tilde{B}_{t+1})$ to $(\tilde{A}_t, \tilde{B}_t)$.

It also follows that F, G, $\tilde{A}$, and $\tilde{B}$ depend on A, B, H, W, and δ, but are independent of C. This demonstrates the certainty equivalence of the discretionary equilibrium.

In a discretion equilibrium,

$$Y_t = D \begin{bmatrix} I \\ G \\ F \end{bmatrix} X_t \equiv \tilde{D} X_t,$$

$$L_t = \frac{1}{2} Y_t' \Lambda Y_t \equiv \frac{1}{2} X_t' \bar{W} X_t,$$

where $\tilde{D}$ is an $n_Y \times n_X$ matrix and $\bar{W} \equiv \frac{1}{2}\tilde{D}'\Lambda\tilde{D}$ is an $n_X \times n_X$ matrix.

The equilibrium loss in any period $t \geq 0$ will satisfy

$$\mathrm{E}_t \sum_{\tau=0}^{\infty} (1-\delta)\delta^{\tau} L_{t+\tau} = \frac{1}{2}[(1-\delta) X_t' V X_t + \delta w],$$

where the $n_X \times n_X$ matrix V and the scalar w, the fixed point of the mapping from (V_{t+1}, w_{t+1}) to (V_t, w_t), satisfy

$$\begin{aligned} V &= \bar{W} + \delta M' V M, \\ w &= \mathrm{tr}[VCC']. \end{aligned}$$

The equilibrium loss obviously depends on C.

One might think that the discretion solution can also be found by combining (2.1) with the first-order condition (2.4) for $t \geq 0$. This solution is generally *not* correct. It amounts to treating expectations $x_{t+1|t}$ as exogenous. This is consistent with (4.3) only in the special case of all predetermined variables in the vector X_t being exogenous, in which case $x_{t+1|t} = GX_{t+1|t}$ is independent of i_t. However, if some predetermined variables are endogenous, $X_{t+1|t}$ and thereby $x_{t+1|t}$ will depend on i_t, which is taken into account in the Bellman equation derived above. The reason why the first-order conditions (2.4) for $t \geq 0$ give the correct discretion solution in the model of Svensson and Woodford (2005) is that all predetermined variables are exogenous there.[4]

5 Targeting rules and instrument rules under discretion

The equilibrium *explicit instrument* rule under discretion is (4.22). (Since the discretion equilibrium is not optimal, it is better to refer to the instrument rule as the equilibrium one than the optimal one.) It can also be written with the instrument as a function of current and past shocks,

$$i_t = F \sum_{\tau=0}^{\infty} M^{\tau} C \varepsilon_{t-\tau}.$$

Any equilibrium *implicit instrument rule* can written, for an arbitrary $n_i \times n_x$ matrix K, as

$$i_t = FX_t = FX_t + K(x_t - GX_t) = (F - KG)X_t + Kx_t.$$

In order to find the equilibrium *targeting rule* under discretion, note that, in a discretion equilibrium, I should use the "equilibrium" model in period t, which takes the future discretion equilibrium into account in the form that $x_{t+\tau} = GX_{t+\tau}$ for $\tau \geq 1$. This equilibrium model can be written as

$$X_{t+1} = \tilde{A}X_t + \tilde{B}i_t + C\varepsilon_{t+1}, \tag{5.1}$$

$$x_t = A_{22}^{-1}(HG\tilde{A} - A_{21})X_t + A_{22}^{-1}(HG\tilde{B} - B_2)i_t. \tag{5.2}$$

Here, (5.2) follows from the lower block of (1.1), $Hx_{t+1|t} = HGX_{t+1|t}$, and (5.1).

From (5.2) and (1.4) it follows that the target variables satisfy

$$\begin{aligned} Y_t &= D_X X_t + D_x x_t + D_i i_t \\ &= [D_X + D_x A_{22}^{-1}(HG\tilde{A} - A_{21})]X_t + [D_x A_{22}^{-1}(HG\tilde{B} - B_2) + D_i]i_t \\ &\equiv \hat{D} \begin{bmatrix} X_t \\ i_t \end{bmatrix}. \end{aligned} \tag{5.3}$$

Then the period loss function satisfies

$$L_t = \frac{1}{2} Y_t' \Lambda Y_t \equiv \frac{1}{2} \begin{bmatrix} X_t' & i_t' \end{bmatrix} \hat{W} \begin{bmatrix} X_t \\ i_t \end{bmatrix}, \tag{5.4}$$

[4] I thank Andrew Levin, Eric Swanson and Joseph Pearlman, who have clarified this point (Joseph Pearlman indirectly via an email to Robert Tetlow) and allowed me to correct erroneous statements in previous versions of these notes.

where $\hat{W} \equiv \hat{D}\Lambda\hat{D}$. The problem has become a problem without any forward-looking variables.

Construct the Lagrangian,

$$
\begin{aligned}
\mathcal{L}_0 &= \mathrm{E}_0 \sum_{t=0}^{\infty}(1-\delta)\delta^t \left\{ L_t + \xi'_{t+1}\left(X_{t+1} - \tilde{A}X_t - \tilde{B}i_t - C\varepsilon_{t+1}\right)\right\} + \frac{1-\delta}{\delta}\xi'_0(X_0 - \bar{X}_0) \\
&= \mathrm{E}_0 \sum_{t=0}^{\infty}(1-\delta)\delta^t \left\{ L_t + \xi'_{t+1}\left(\hat{H}\begin{bmatrix} X_{t+1} \\ i_{t+1} \end{bmatrix} - \hat{A}\begin{bmatrix} X_t \\ i_t \end{bmatrix} - \begin{bmatrix} C\varepsilon_{t+1} \\ 0 \end{bmatrix}\right)\right\} \\
&\quad + \frac{1-\delta}{\delta}\xi'_0(X_0 - \bar{X}_0),
\end{aligned}
$$

where ξ_{t+1} is a vector of n_X Lagrange multipliers of the model equations, and

$$
\hat{H} \equiv \begin{bmatrix} I & 0 \end{bmatrix}, \qquad \hat{A} \equiv \begin{bmatrix} \tilde{A} & \tilde{B} \end{bmatrix}.
$$

The first-order conditions with respect to X_t and i_t for $t \geq 0$, taking (5.4) into account, can be written

$$
\begin{bmatrix} X'_t & i'_t \end{bmatrix}\hat{W} + \xi'_t\frac{1}{\delta}\hat{H} - \xi'_{t+1|t}\hat{A} = 0, \tag{5.5}
$$

where $X_0 = \bar{X}_0$. The first-order condition (5.5) can be rewritten as

$$
\begin{aligned}
\tilde{A}'\xi_{t+1|t} &= \frac{1}{\delta}\xi_t + \hat{D}'_X\Lambda Y_t & (5.6) \\
\tilde{B}'\xi_{t+1|t} &= \hat{D}'_i\Lambda Y_t, & (5.7)
\end{aligned}
$$

where $\hat{D} \equiv [\hat{D}_X \;\; \hat{D}_i]$ is partitioned in conformity with X_t and i_t. A similar argument as above for the commitment case will result in the equation

$$
\tilde{B}'\Phi(\mathrm{L})\mathrm{E}_t Q(\mathrm{L}^{-1})\hat{D}'_X\Lambda Y_t - \alpha(\mathrm{L})\hat{D}'_i\Lambda Y_t = 0,
$$

where $\Phi(\mathrm{L})$ and $Q(\mathrm{L}^{-1})$ are $n_X \times n_X$ matrix polynomials and $\alpha(\mathrm{L})$ is a scalar polynomial.

This is the general form of the optimal targeting rule for the discretion case. Again, although it looks complicated in the general form, in any given model it is often quite simple. Again, the optimal targeting rule corresponds to the equality of the marginal rates of substitution and marginal rates of transformation between the target variables,

$$
\mathrm{MRS} = \mathrm{MRT}.
$$

The marginal rates of substitution between the target variables follow from the loss function, (1.5) and (1.6). Under discretion, the relevant marginal rates of transformation are the "equilibrium" ones, which follow from the "equilibrium" model equations (5.1) and (5.3), not the "structural" model equations (1.1) and (1.4).

Appendix

A Solving a system of linear difference equations with nonpredetermined variables

Consider the system

$$H\begin{bmatrix} y_{1,t+1} \\ \mathrm{E}_t y_{2,t+1} \end{bmatrix} = A\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} C \\ 0 \end{bmatrix} \varepsilon_{t+1} \tag{A.1}$$

for $t \geq 0$, where y_{1t} is an n_1-vector of predetermined variables and y_{10} is given, y_{2t} is an n_2-vector of nonpredetermined variables, ε_{t+1} is an iid random n_ε-vector with zero mean and covariance matrix I. The real matrices A and H are $n \times n$, where $n \equiv n_1 + n_2$, and the real matrix C is $n_1 \times n_\varepsilon$. Let the matrices H and A be partitioned conformably with y_{1t} and y_{2t} so

$$H \equiv \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \quad A \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

I assume that H_{11} and A_{22} are nonsingular. Then the upper block of (A.1) determines $y_{1,t+1}$ given y_{1t}, y_{2t}, and ε_{t+1}. The lower block determines y_{2t} given y_{1t} and $\mathrm{E}_t y_{t+1}$.

Take expectations conditional on information in period t and write the system as

$$H\begin{bmatrix} \mathrm{E}_t y_{1,t+1} \\ \mathrm{E}_t y_{2,t+1} \end{bmatrix} = A\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \tag{A.2}$$

Following Klein (2000), Sims (2002b), and Söderlind (1999), use the generalized Schur decomposition of A and H. This decomposition results in the square possibly complex matrices Q, S, T, and Z such that

$$A = Q'TZ', \tag{A.3}$$

$$H = Q'SZ', \tag{A.4}$$

where Q' for a complex matrix denotes the complex conjugate transpose of Q (the transpose of the complex conjugate of Q).[5] The matrices Q and Z are unitary ($Q'Q = Z'Z = I$), and S and T are upper triangular (see Golub and van Loan (1989)). Furthermore, the decomposition is sorted according to ascending modulus of the generalized eigenvalues, so $|\lambda_j| \geq |\lambda_k|$ for $j \geq k$.[6]

[5] If $Q = [q_{jk}]$ has elements $q_{jk} = \mathrm{Re}q_{jk} + i\mathrm{Im}q_{jk}$, the complex conjugate of Q is the matrix $\bar{Q} = [\bar{q}_{jk}]$ with elements $\bar{q}_{jk} = \mathrm{Re}q_{jk} - i\mathrm{Im}q_{jk}$.

[6] The sorting of the eigenvalues was previously often done by two programs written by Sims and available at www.princeton.edu/∼sims, Qzdiv and Qzswitch. Now they are done much faster with the Matlab function ordqz.

(The generalized eigenvalues are the ratios of the diagonal elements of T and S, $\lambda_j = t_{jj}/s_{jj}$ ($j = 1, ..., n$). A generalized eigenvalue is infinity if $t_{jj} \neq 0$ and $s_{jj} = 0$ and zero if $t_{jj} = 0$ and $s_{jj} \neq 0$.)

Assume the *saddlepoint property* (Blanchard and Kahn (1980)): The number of generalized eigenvalues with modulus larger than unity (the unstable eigenvalues) equals the number of non-predetermined variables. Thus, I assume that $|\lambda_j| > 1$ for $n_1 + 1 \leq j \leq n_1 + n_2$ and $|\lambda_j| < 1$ for $1 \leq j \leq n_1$.(for an exogenous predetermined variable with a unit root, I can actually allow $|\lambda_j| = 1$ for some $1 \leq j \leq n_1$).

Define

$$\begin{bmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{bmatrix} \equiv Z' \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}. \tag{A.5}$$

I can interpret $\tilde{y}_{1t}$ as a complex vector of n_1 transformed predetermined variables and $\tilde{y}_{2t}$ as a complex vector of n_2 transformed non-predetermined variables. Premultiply the system (A.2) by Q and use (A.3)-(A.5) to write it as

$$\begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix} \begin{bmatrix} \mathrm{E}_t\tilde{y}_{1,t+1} \\ \mathrm{E}_t\tilde{y}_{2,t+1} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{bmatrix}, \tag{A.6}$$

where S and T have been partitioned conformably with $\tilde{y}_{1t}$ and $\tilde{y}_{2t}$.

Consider the lower block of (A.6),

$$S_{22}\, \mathrm{E}_t\tilde{y}_{2,t+1} = T_{22}\, \tilde{y}_{2t}. \tag{A.7}$$

Since the diagonal terms of S_{22} and T_{22} (s_{jj} and t_{jj} for $n_1 + 1 \leq j \leq n_1 + n_2$) satisfy $|t_{jj}/s_{jj}| > 1$, the diagonal terms of T_{22} are nonzero, the determinant of T_{22} is nonzero, and T_{22} is invertible. Note that S_{22} may not be invertible. I can then solve for $\tilde{y}_{2t}$ as

$$\tilde{y}_{2t} = J\mathrm{E}_t\tilde{y}_{2,t+1} = 0, \tag{A.8}$$

where the complex matrix J is given by

$$J \equiv T_{22}^{-1} S_{22}. \tag{A.9}$$

I exploit that the modulus of the diagonal terms of $T_{22}^{-1} S_{22}$ is less than one. I also assume that $\mathrm{E}_t\tilde{y}_{2,t+\tau}$ is sufficiently bounded. Then $J^\tau \mathrm{E}_t\tilde{y}_{2,t+\tau} \to 0$ when $\tau \to \infty$. Note that J may not be invertible, since S_{22} may not be invertible.

I have, by (A.5),

$$
\begin{aligned}
y_{1t} &= Z_{11}\tilde{y}_{1t}, && \text{(A.10)}\\
y_{2t} &= Z_{21}\tilde{y}_{1t}, && \text{(A.11)}
\end{aligned}
$$

where

$$Z \equiv \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \tag{A.12}$$

is partitioned conformably with y_{1t} and y_{2t}. Under the assumption of the saddlepoint property, Z_{11} is square. I furthermore assume that Z_{11} is *invertible*. Then I can solve for $\tilde{y}_{1t}$ in (A.10),

$$\tilde{y}_{1t} = Z_{11}^{-1} y_{1t}, \tag{A.13}$$

and use this in (A.11) to get

$$y_{2t} = F y_{1t}, \tag{A.14}$$

where the real $n_2 \times n_1$ matrix F is given by

$$F \equiv Z_{21} Z_{11}^{-1}. \tag{A.15}$$

It remains to find a solution for $y_{1,t+1}$. By (A.8), the upper block of (A.6) is

$$S_{11}\mathrm{E}_t\tilde{y}_{1,t+1} = T_{11}\tilde{y}_{1t}.$$

Since the diagonal terms of S_{11} and T_{11} satisfy $|t_{jj}/s_{jj}| < 1$, all diagonal terms of S_{11} must be nonzero, so the determinant of S_{11} is nonzero, and S_{11} is invertible. I can then solve for $\mathrm{E}_t\tilde{y}_{1,t+1}$ as

$$\mathrm{E}_t\tilde{y}_{1,t+1} = S_{11}^{-1} T_{11}\tilde{y}_{1t}.$$

By (A.10),

$$
\begin{aligned}
\mathrm{E}_t y_{1,t+1} &= Z_{11}\mathrm{E}_t\tilde{y}_{1,t+1} \\
&= Z_{11} S_{11}^{-1} T_{11}\tilde{y}_{1t} \\
&= Z_{11} S_{11}^{-1} T_{11} Z_{11}^{-1} y_{1t} && \text{(A.16)}
\end{aligned}
$$

where I have used (A.13).

It follows that I can write the solution as

$$y_{1,t+1} = M y_{1t} + H_{11}^{-1} C \varepsilon_{t+1}, \tag{A.17}$$

where the real matrix M is given by

$$M \equiv Z_{11} S_{11}^{-1} T_{11} Z_{11}^{-1} \tag{A.18}$$

Thus, the solution to the system (A.1) is given by (A.14) and (A.17) for $t \geq 0$.

A.1 Relation to Sims (2002b)

Consider the system of equations (A.1) and introduce the vector of endogenous expectational errors of the nonpredetermined variables,

$$\eta_t \equiv y_{2t} - \mathrm{E}_{t-1} y_{2t},$$

which have the obvious property that $\mathrm{E}_{t-1}\eta_t = 0$. Using this to substitute for $\mathrm{E}_t y_{2,t+1}$ in (A.1), the latter can be written in the form

$$\Gamma_0 y_t = \Gamma_1 y_{t-1} + \Psi \varepsilon_t + \Pi \eta_t, \tag{A.19}$$

where

$$y_t \equiv \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}, \qquad \Gamma_0 \equiv H, \qquad \Gamma_1 \equiv A, \qquad \Psi \equiv \begin{bmatrix} C \\ 0 \end{bmatrix}, \qquad \Pi \equiv H_2,$$

where $H \equiv [H_1 \;\; H_2]$ is partitioned conformably with y_{1t} and y_{2t}. Sims (2002b) uses the generalized Schur decomposition to find a solution to (A.19) with a more elaborate method than the one used above in A, explicitly taking into account the condition $\mathrm{E}_{t-1}\eta_t = 0$ and the somewhat complex restrictions this implies.

Sims also deals with the case when ε_t is an arbitrary exogenous stochastic process and not necessarily a zero-mean iid shocks as above. Then the solution can be expressed as

$$y_t = \Theta_1 y_{t-1} + \Theta_0 \varepsilon_t + \Theta_y \sum_{\tau=0}^{\infty} \Theta_f^{\tau} \Theta_\theta \mathrm{E}_t \varepsilon_{t+1+\tau},$$

where Θ_0 and Θ_1 are real matrices, Θ_y, Θ_f, and Θ_θ are complex matrices, and $\Theta_y \Theta_f^{\tau} \Theta_\theta$ for any integer $\tau \geq 0$ is a real matrix.[7] Svensson (2005, Appendix) solves (A.1) with the method in appendix A when ε_t is an arbitrary exogenous stochastic process and expresses the solutions as

$$\begin{aligned}
y_{2t} &= F y_{1t} + Z_t, \\
y_{1,t+1} &= M y_{1t} + N \mathrm{E}_t Z_{t+1} + P C \mathrm{E}_t \varepsilon_{t+1} + H_{11}^{-1} C (\varepsilon_{t+1} - \mathrm{E}_t \varepsilon_{t+1}), \\
Z_t &\equiv \sum_{\tau=0}^{\infty} R_\tau H_{11}^{-1} C \mathrm{E}_t \varepsilon_{t+1+\tau},
\end{aligned}$$

where F, M, N, P and $\{R_\tau\}_{\tau=0}^{\infty}$ are real matrices of appropriate dimension.

Klein (2000) provides a detailed discussion of how the solution method used in these notes relates to those of Blanchard and Kahn (1980), Binder and Pesaran (1994) and Binder and Pesaran (1997), King and Watson (1998) and King and Watson (2002), Sims (2002a), and Uhlig (1995).

[7] The solution is calculated by his software Gensys, available at www.princeton.edu/~sims.

B The recursive saddlepoint method for a nonlinear problem

Consider the nonlinear problem

$$\min_{\{i_t\}_{t\geq 0}} \mathrm{E}_0 \sum_{t=0}^{\infty} (1-\delta)\delta^t L_t, \tag{B.1}$$

where the period loss function is

$$L_t = L(X_t, x_t, i_t, s_t), \tag{B.2}$$

the constraints are

$$\begin{aligned} X_{t+1} &= A_1(X_t, x_t, i_t, s_{t+1}), & \text{(B.3)} \\ \mathrm{E}_t H(X_{t+1}, x_{t+1}, s_{t+1}) &= A_2(X_t, x_t, i_t, s_t), & \text{(B.4)} \end{aligned}$$

where $\{s_t\}$ is an exogenous Markov process, $A_1(\cdot)$, $A_2(\cdot)$, and $H(\cdot)$ are vector-valued functions of the same dimension as X_t, x_t, and x_t, respectively, and where X_0 and s_0 are given. Equation (B.3) determines X_{t+1} given X_t, x_t, i_t, and s_{t+1}. The function $A_2(X_t, x_t, i_t, s_t)$ is assumed to be invertible with respect to x_t, so equation (B.4) determines x_t given X_t, i_t, s_t, and expectations $\mathrm{E}_t H(X_{t+1}, x_{t+1}, s_{t+1})$.

Write the Lagrangian as

$$\begin{aligned} \mathcal{L}_0 &= \mathrm{E}_0 \sum_{t=0}^{\infty} (1-\delta)\delta^t \left\{ \begin{array}{l} L_t + \Xi_t'[\mathrm{E}_t H(X_{t+1}, x_{t+1}, s_{t+1}) - A_2(X_t, x_t, i_t, s_t)] \\ +\xi_{t+1}'[X_{t+1} - A_1(X_t, x_t, i_t, s_{t+1})] \end{array} \right\} \\ &= \mathrm{E}_0 \sum_{t=0}^{\infty} (1-\delta)\delta^t \left\{ \begin{array}{l} L_t + \Xi_t'[H(X_{t+1}, x_{t+1}, s_{t+1}) - A_2(X_t, x_t, i_t, s_t)] \\ +\xi_{t+1}'[X_{t+1} - A_1(X_t, x_t, i_t, s_{t+1})] \end{array} \right\}, \end{aligned} \tag{B.5}$$

where Ξ_t is the vector of Lagrange multipliers for (B.4), ξ_{t+1} is the vector of Lagrange multipliers for (B.4), and the second equality follows from the law of iterated expectations. The problem is not recursive, since the term $H(X_{t+1}, x_{t+1}, s_{t+1})$, which depends on the forward-looking variable x_{t+1}, appears in the first line of of (B.5). However, note that the discounted sum of the first line of (B.5) can be written

$$\sum_{t=0}^{\infty} (1-\delta)\delta^t \{L_t + \Xi_t'[H(X_{t+1}, x_{t+1}, s_{t+1}) - A_2(X_t, x_t, i_t, s_t)]\} =$$

$$\sum_{t=0}^{\infty} (1-\delta)\delta^t \{L_t - \Xi_t' A_2(X_t, x_t, i_t, s_t) + \frac{1}{\delta}\Xi_{t-1}' H(X_t, x_t, s_t)\},$$

where $\Xi_{-1} = 0$. Now all the terms within the curly brackets on the right side are dated t or earlier. The recursive saddlepoint method is in this case to let this expression within the curly brackets define the dual period loss. More precisely, the dual period loss is defined as

$$\tilde{L}_t \equiv L_t - \gamma_t' A_2(X_t, x_t, i_t, s_t) + \frac{1}{\delta}\Xi_{t-1}' H(X_t, x_t, s_t) \equiv \tilde{L}(X_t, \Xi_{t-1}; x_t, i_t, \gamma_t; s_t),$$

where Ξ_{t-1} is a predetermined variable in period t and γ_t is an additional control variable, and where Ξ_{t-1} and γ_t are related by the dynamic equation,

$$\Xi_t = \gamma_t. \tag{B.6}$$

Marcet and Marimon (1998, 2019) show that the problem can then be reformulated as the recursive saddlepoint problem,

$$\max_{\{\gamma_t\}_{t\geq 0}} \min_{\{x_t, i_t\}_{t\geq 0}} \mathrm{E}_0 \sum_{t=0}^{\infty} (1-\delta)\delta^t \tilde{L}_t,$$

where the optimization is subject to (B.3), (B.6), and X_0 and Ξ_{-1} given. The value function for the saddlepoint problem, starting in any period t, satisfies the Bellman equation

$$\tilde{V}(\tilde{X}_t; s_t) \equiv \max_{\gamma_t} \min_{(x_t, i_t)} \{(1-\delta)\tilde{L}(\tilde{X}_t; \tilde{\imath}_t; s_t) + \delta \mathrm{E}_t \tilde{V}(\tilde{X}_{t+1}; s_{t+1})\},$$

subject to (B.3) and (B.6), where $\tilde{X}_t \equiv (X_t', \Xi_{t-1}')'$ and $\tilde{\imath}_t \equiv (x_t', i_t', \gamma_t')'$. The optimal policy function for the saddlepoint problem will be

$$\tilde{\imath}_t = F(\tilde{X}_t, s_t) \equiv \begin{bmatrix} F_x(\tilde{X}_t, s_t) \\ F_i(\tilde{X}_t, s_t) \\ F_\gamma(\tilde{X}_t, s_t) \end{bmatrix}.$$

It follows that the solution for the original problem is

$$\begin{aligned}
x_t &= F_x(\tilde{X}_t, s_t), \\
i_t &= F_i(\tilde{X}_t, s_t), \\
L_t &= L[X_t, F_x(\tilde{X}_t, s_t), F_i(\tilde{X}_t, s_t), s_t] \;\equiv\; \bar{L}(\tilde{X}_t, s_t), \\
\tilde{X}_{t+1} &= \begin{bmatrix} A_1[X_t, F_x(\tilde{X}_t, s_t), F_i(\tilde{X}_t, s_t), s_{t+1}] \\ F_\gamma(\tilde{X}_t, s_t) \end{bmatrix} \;\equiv\; M(\tilde{X}_t, s_t, s_{t+1}).
\end{aligned}$$

The value function for the original problem satisfies the Bellman equation

$$V(\tilde{X}_t; s_t) \equiv (1-\delta)\bar{L}(\tilde{X}_t, s_t) + \delta \mathrm{E}_t V[M(\tilde{X}_t, s_t, s_{t+1}), s_{t+1}].$$

This value function is related to the value function of the dual problem by

$$V(\tilde{X}_t; s_t) \equiv \tilde{V}(\tilde{X}_t; s_t) - \frac{1-\delta}{\delta}\Xi_{t-1}' H[X_t, F_x(\tilde{X}_t, s_t), s_t].$$

Expressing the constraint as (B.4) is not very restrictive. Suppose, for instance, that the constraint (B.4) is replaced by a constraint of the form

$$\mathrm{E}_t \sum\nolimits_{\tau=1}^{\infty} \delta^\tau H(X_{t+\tau}, x_{t+\tau}, s_{t+\tau}) = A_2(X_t, x_t, i_t, s_t) \tag{B.7}$$

(this case is also treated in Marcet and Marimon (1998, 2019)). The constraint can easily be rewritten of the form (B.4). First, introduce the additional forward-looking variable

$$x_{2,t} \equiv \mathrm{E}_t \sum\nolimits_{\tau=0}^{\infty} \delta^\tau H(X_{t+\tau}, x_{t+\tau}, s_{t+\tau}) = H(X_t, x_t, s_t) + \delta \mathrm{E}_t x_{2,t+1}.$$

Second, replace (B.7) by the two constraints

$$\begin{aligned} \delta \mathrm{E}_t x_{2,t+1} &= -H(X_t, x_t, s_t) + x_{2t}, \\ 0 &= A_2(X_t, x_t, i_t, x_t) + H(X_t, x_t, s_t) - x_{2t}. \end{aligned}$$

These two constraints are obviously of the same form as (B.4), since they can be written

$$\mathrm{E}_t \tilde{H}(X_{t+1}, x_{t+1}, x_{2,t+1}, s_{t+1}) = \tilde{A}_2(X_t, x_t, x_{t2}, i_t, s_t),$$

where

$$\begin{aligned} \tilde{H}(X_t, x_t, x_{2t}, s_t) &\equiv \begin{bmatrix} \delta x_{2t} \\ 0 \end{bmatrix}, \\ \tilde{A}_2(X_t, x_t, x_{t2}, i_t, s_t) &\equiv \begin{bmatrix} -H(X_t, x_t, s_t) + x_{2t} \\ A_2(X_t, x_t, i_t, x_t) + H(X_t, x_t, s_t) - x_{2t} \end{bmatrix}. \end{aligned}$$

Note that it should be possible to use the recursive saddlepoint method to solve nonlinear difference equations with forward-looking variables, as in the linear case.

References

Anderson, Evan W., Lars Peter Hansen, Ellen R. McGrattan, and Thomas J. Sargent (1995), "Mechanics of Forming and Estimating Dynamic Linear Economies," working paper, homepages.nyu.edu/ ts43.

Backus, David, and John Driffill (1986), "The Consistency of Optimal Policy in Stochastic Rational Expectations Models," CEPR Discussion Paper 124, https://cepr.org/active/publications/discussion_papers/dp.php?dpno=124.

Binder, Michael, and M. Hashem Pesaran (1994), "Multivariate Rational Expectations Models: A Review and Some New Results," in Pesaran, M.H., and M. Wickens (eds.), *Handbook of Applied Econometrics: Macroeconomics*, Basil Blackwell, Oxford, pages 139–187.

Binder, Michael, and M. Hashem Pesaran (1997), "Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and Their Fully Recursive Computation," *Econometric Theory* 13, 877–888.

Blanchard, Olivier J., and Charles M. Kahn (1980), "The Solution of Linear Difference Models under Rational Expectations," *Econometrica* 48, 1305–1311.

Currie, David, and Paul Levine (1993), *Rules, Reputation and Macroeconomic Policy Coordination*, Cambridge University Press, Cambridge.

Giannoni, Marc P., and Michael Woodford (2003), "Optimal Interest-Rate Rules: I. General Theory," NBER Working Paper No. 9419, https://www.nber.org/papers/w9419.

Golub, Gene H., and Charles F. van Loan (1989), *Matrix Computations, Second Edition*, Johns Hopkins University Press, Baltimore, MD.

King, Robert G., and Mark W. Watson (1998), "The Solution of Singular Difference Systems under Rational Expectations," *International Economic Review* 39, 1015–1026.

King, Robert G., and Mark W. Watson (2002), "System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations," *Computational Economics* 20, 57–86.

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

Ljungqvist, Lars, and Thomas J. Sargent (2004), *Recursive Macroeconomic Theory, Second Edition*, MIT Press, Cambridge, MA.

Marcet, Albert, and Ramon Marimon (1998), "Recursive Contracts," University of Pompeu Fabra Economics Working Paper 337.

Marcet, Albert, and Ramon Marimon (2019), "Recursive Contracts," *Econometrica* 87(5), 1589–1631.

Oudiz, Gilles, and Jeffrey Sachs (1985), "International Policy Coordination in Dynamic Macroeconomic Models," in Buiter, William H., and Richard C. Marston (eds.), *International Economic Policy Coordination*, Cambridge University Press, Cambridge.

Rudebusch, Glenn, and Lars E.O. Svensson (1999), "Policy Rules for Inflation Targeting," in Taylor, John B. (ed.), *Monetary Policy Rules*, The University of Chicago Press, pages 203–246, http://www.nber.org/chapters/c7417.pdf.

Sims, Christopher A. (1999), "Random Lagrange Multipliers and Transversality," lecture notes.

Sims, Christopher A. (2000), "Stochastic Lagrange Multipliers for Problems with Lagged Expectations," lecture notes.

Sims, Christopher A. (2002a), "Principal Components, Some Standard Distributions," lecture notes.

Sims, Christopher A. (2002b), "Solving Linear Rational Expectations Models," *Computational Economics* 20(1-2), 1–20.

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

Svensson, Lars E.O. (1999), "Inflation Targeting as a Monetary Policy Rule," *Journal of Monetary Economics* 43(3), 607–654.

Svensson, Lars E.O. (2003), "What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules," *Journal of Economic Literature* 41(2), 426–477.

Svensson, Lars E.O. (2005), "Monetary Policy with Judgment: Forecast Targeting," *International Journal of Central Banking* 1(1), 1–54, https://larseosvensson.se/2005/05/01/monetary-policy-with-judgment-forecast-targeting/.

Svensson, Lars E.O., and Michael Woodford (2005), "Implementing Optimal Policy through Inflation-Forecast Targeting," in Bernanke, Ben S., and Michael Woodford (eds.), *The Inflation-Targeting Debate*, chap. 2, University of Chicago Press, pages 19–83, https://www.nber.org/chapters/c9556.pdf.

Uhlig, Harald (1995), "A Toolkit for Analyzing Nonlinear Economic Dynamic Models Easily," Discussion Paper 101, Federal Reserve Bank of Minneapolis.