MONETARY POLICY WITH JUDGMENT: FORECAST TARGETING

Lars E.O. Svensson

Monetary Policy with Judgment: Forecast Targeting
Lars E.O. Svensson
NBER Working Paper No. 11167
March 2005
JEL No. E42, E52, E58

ABSTRACT

"Forecast targeting," forward-looking monetary policy that uses central-bank judgment to construct optimal policy projections of the target variables and the instrument rate, may perform substantially better than monetary policy that disregards judgment and follows a given instrument rule. This is demonstrated in a few examples for two empirical models of the U.S. economy, one forward looking and one backward looking. A complicated infinite-horizon central-bank projection model of the economy can be closely approximated by a simple finite system of linear equations, which is easily solved for the optimal policy projections. Optimal policy projections corresponding to the optimal policy under commitment in a timeless perspective can easily be constructed. The whole projection path of the instrument rate is more important than the current instrument setting. The resulting reduced-form reaction function for the current instrument rate is a very complicated function of all inputs in the monetary-policy decision process, including the central bank's judgment. It cannot be summarized as a simple reaction function such as a Taylor rule. Fortunately, it need not be made explicit.

Lars E. O. Svensson
Department of Economics
Fisher Hall
Princeton University
Princeton, NJ 08544-1021
and NBER
svensson@princeton.edu

1. Introduction

On a general level, this paper is motivated by a desire to provide a better theory of modern monetary policy, both from a descriptive and a normative point of view, than much of the current literature on monetary policy. The current literature to a large extent applies a one-line modeling of monetary policy, such as when the instrument rate is assumed to be a given function of a few variables, for instance, "monetary policy is assumed to follow a Taylor rule."

I believe that the theory that I develop here is better from a *descriptive* point of view, since it takes into account some crucial aspects of monetary-policy decisions, such as the collection, processing, and analysis of large amounts of data, the construction of projections of the target variables, the use of considerable amounts of judgment, and the desire to achieve (mostly) relatively specific objectives. The modern monetary-policy process I have in mind can be concisely described as "forecast targeting," meaning "setting the instrument rate such that the forecasts of the target variables look good," where "look good" refers to the objectives of monetary policy, such as a given target for inflation and a zero target for the output gap.[1] I believe this view of the monetary-policy process is also helpful from a *normative* point of view, for instance, in evaluating the performance of and suggesting improvements to existing monetary policy.[2]

On a more specific level, this paper is motivated by a desire to demonstrate the crucial and beneficial role of *judgment*—information, knowledge, and views outside the scope of a particular model—in modern monetary policy and, in particular, to demonstrate that the appropriate use of good judgment can dramatically improve monetary-policy performance, even when compared to policy that is optimal in all respects except for incorporating judgment.[3] As will be explained in detail below, judgment will be represented as the central-bank's conditional mean estimate of arbitrary multidimensional stochastic "deviations"—"add factors"—to the model equations.[4] I also

---

[1] Bernanke [3] discusses and compares forecast targeting (which he refers to as "forecast-based policies") and simple instrument rules (which he refers to as "simple feedback policies"). He states that "the Federal Reserve relies primarily on the forecast-based approach for making policy" and cites Greenspan's [9] speech, entitled "Risk and Uncertainty in Monetary Policy," as evidence. He also notes "that not only have most central banks chosen to rely most heavily on forecast-based policies but also that the results, at least in recent years, have generally been quite good, as most economies have enjoyed low inflation and overall economic stability."

[2] See Svensson [22] and Svensson, Houg, Solheim, and Steigum [28] for examples of evaluations of monetary policy in New Zealand and Norway, respectively, with this perspective.

[3] Svensson [25] also emphasizes the role of judgment in monetary policy but does not provide any direct comparision of the performance of monetary policy with and without judgment.

[4] Svensson and Tetlow [31] show how central-bank judgment can be extracted according to the method of Optimal Policy Projections (OPP). This is a method to provide advice on optimal monetary policy while taking policymakers' judgment into account. An early version of the method was developed by Robert Tetlow for a mostly backward-looking variant of the Federal Reserve Board's FRB/US model. The resulting projections have been referred to at the Federal Reserve Board as "policymaker perfect-foresight projections"—somewhat misleadingly. The paper demonstrates the usefulness of OPP with a few example projections for two Greenbook forecasts and the FRB/US model.

wish to demonstrate the benefits of regarding the whole projection paths of the target variables rather than forecasts at some specific horizon, such as 8 quarters, as the relevant objects in the monetary-policy decision process. In particular, I believe that it is important to emphasize the whole projection of future instrument rates rather than just the current instrument rate. Furthermore, the modern view of the transmission mechanism of monetary policy emphasizes that monetary-policy actions have effects on the economy and the central bank's target variables almost exclusively through the private-sector expectations of the future paths of inflation, output, and interest rates that these actions give rise to; therefore, monetary policy is really the management of private-sector expectations (Woodford [35]). From this follows that effective implementation of monetary policy requires the effective communication to the private sector of the central bank's preferred projections, including the instrument-rate projection. The most obvious communication of these projections is to explicitly announce and motivate them. Finally, I want to demonstrate the benefits of the approximation of an inherently rather complex infinite-horizon central-bank projection models of the economy to much simpler finite-horizon projection models that are much easier to use but still arbitrarily close approximations to the infinite-horizon models.

The decision process of modern monetary policy has several distinct characteristics (see Brash [4], Sims [19], and Svensson [22]):

1. Large amounts of data about the state of the economy and the rest of the world, including private-sector expectations and plans, are collected, processed, and analyzed before each major decision.

2. Because of lags in the transmission process, monetary-policy actions affect the economy with a lag. For this reason alone, good monetary policy must be forward-looking, aim to influence the future state of the economy, and therefore rely on forecasts—projections. Central-bank staff and policymakers make projections of the future development of a number exogenous variables, such as foreign developments, import supply, export demand, fiscal policy, productivity growth, and so forth. They also construct projections of a number of endogenous variables, quantities and prices, under alternative assumptions, including alternative assumptions about the future path of instrument rates. The policymakers are presented with projections of the most important variables, including target variables such as inflation and output, often under alternative assumptions about exogenous variables and, in particular, the instrument rate (such as the instrument rate being constant, following market expectations, following some

arbitrary reaction function, or being optimal relative to a specific objective function).

3. Throughout this process, a considerable amount of judgment is applied to assumptions and projections. Projections and monetary-policy decisions cannot rely on models and simple observable data alone. All models are drastic simplifications of the economy, and data give a very imperfect view of the state of the economy. Therefore, judgmental adjustments in both the use of models and the interpretation of their results—adjustments due to information, knowledge, and views outside the scope of any particular model—are a necessary and essential component in modern monetary policy.

4. Based on this large amount of information and analysis, the policymakers decide on a current instrument rate, such that the corresponding projections of the target variables look good relative to the central bank's objectives. Since the projections of the target variables depend insignificantly on the current instrument-rate setting and mainly on the whole path of future instrument rates, the policymakers, explicitly or implicitly, actually choose an instrument-rate projection—an instrument-rate plan—and the current instrument-rate decision can be seen as the first element of that plan.

5. Finally, the current instrument rate is announced and implemented. In many cases, the corresponding projections for inflation and output or the output gap are also announced. In a few cases, an instrument-rate projection is announced as well.[5]

This process makes the current instrument-rate decision a very complex function of the large amounts of data and judgment that have entered into the process. I believe that it is not very helpful to summarize this function as a simple reaction function such as a Taylor rule. Furthermore, the resulting complex reaction function is a *reduced form*, which depends on the central-bank objectives, its view of the transmission mechanism of monetary policy, the data the central bank has collected, and the judgment it has exercised. It is the endogenous complex result of a complex process. In no way is this reaction function *structural*, in the sense of being invariant to the central bank's view of the transmission mechanism and private-sector behavior, or the amount of information and judgmental adjustments. Still, much current literature treats monetary policy as characterized by a given reaction function that is essentially structural and invariant to changes in the model of the economy. Treating the reaction function as a reduced form is a first step in a sensible theory of

---

[5] The Reserve Bank of New Zealand has published an instrument-rate projection for many years. The Bank of Norway is increasingly providing more information about the future path of the instrument rate.

monetary policy. But, fortunately, this complex reduced-form reaction function need not be made explicit. It is actually not needed in the modern monetary-policy process.

However, there is a convenient, more structural representation of monetary policy, namely in the form of a *targeting rule*, as advocated recently in some detail in Svensson and Woodford [30] and Svensson [25] and earlier more generally in Svensson [21]. An *optimal* targeting rule is a first-order condition for optimal monetary policy. It corresponds to the standard efficiency condition of equality between the marginal rates of substitution and the marginal rates of transformation between the target variables, the former given by the monetary-policy loss function, the latter given by the transmission mechanism of monetary policy. An optimal targeting rule is invariant to everything else in the model, including additive judgment and the stochastic properties of additive shocks. Thus, it is a compact, robust, and structural representation of monetary policy, and much more robust than the optimal reaction function. A *simple* targeting rule can potentially be a practical representation of robust monetary policy, a robust monetary policy that performs reasonably well under different circumstances.[6]

Optimal targeting rules remain a practical way of representing optimal monetary policy in the small models usually applied for academic monetary-policy analysis. However, for the larger and higher-dimensional operational macromodels used by many central banks in constructing projections, the optimal targeting rule becomes more complex and arguably less practical as a representation of optimal monetary policy. In this paper, it is demonstrated that optimal policy projections, the projections corresponding to optimal policy under commitment in a timeless perspective, can easily be derived directly with simple numerical methods, without reference to any optimal targeting rule.[7] For practical optimal monetary policy, policymakers actually need not know the optimal targeting rule. Even less do they need to know any reaction function. They only need to ponder the graphs of the projections of the target variables that are generated in the policy process and choose the projections of the target variables and the instrument rate that look best relative to the central bank's objectives.

The paper is organized in the following way. Section 2 lays out a reasonably general infinite-horizon model of the transmission mechanism and the central bank's objectives; defines projections, judgment and optimal policy projections; and specifies how the optimal policy can be implemented

---

[6] McCallum and Nelson [15] have recently criticized the advocacy of targeting rules in Svensson [25]. Svensson [26] rebuts this criticism and gives references to a rapidly growing literature that applies targeting rules to monetary-policy analysis. Walsh [34] shows a case of equivalence between targeting rules and robust control.

[7] Nevertheless, a general form of an optimal targeting rule is derived in appendix F, for the finite-horizon approximation of the projection model.

and what information the private sector needs from the central bank. The section also presents a simple model of judgment, when the deviation is a version of a finite-order moving average. Then judgment can be seen as the accumulation of information over time and allows for a recursive but high-dimensional representation of the dynamics of the deviation and judgment. Finally, the section represents the optimal policy projections as the solution to a somewhat complex system of difference equations, while taking judgment into account. It also makes the point that, fortunately, the complex reduced-form reaction function need not be made explicit. Section 3 presents a convenient finite-horizon model for the construction of optimal policy projections, for both forward- and backward-looking models. This finite-horizon model can be written as a simple finite system of linear equations. Nevertheless, it is an exact or arbitrarily close approximation to the infinite-horizon model and is easily solved for the optimal policy projections taking judgment into account. Section 4 discusses and specifies monetary policy that disregards judgment and follows different instrument rules, such as variants of the Taylor rule or more complex instrument rules that are optimal in the absence of judgment. Section 5 gives examples of and compares monetary policy with and without judgment, for two different empirical models of the U.S. economy: the backward-looking model of Rudebusch and Svensson [18] and the forward-looking New Keynesian model of Lindé [13]. In these examples, monetary policy with judgment results in substantially better performance than monetary policy without judgment. This is also the case when monetary policy without judgment is represented as a Taylor rule where the instrument rate responds to forward-looking variables that incorporate private-sector judgment (although, as emphasized below, there are serious principal and practical problems in implementing such an instrument rule). Section 6 presents conclusions.

A separate and extensive appendix contains numerous technical details.[8] These include a general solution of the policy problem and the related system of difference equations with forward-looking variables when the deviation is an arbitrary stochastic process; a specification of the model when the deviation and judgment are finite-order moving-average processes and the application of the practical method of Marcet and Marimon [14] to that case; the precise mathematical structure of the finite-horizon approximation model, including the optimal targeting rule; and details on the empirical backward- and forward-looking models.

---

[8] The appendix is available at www.princeton.edu/~svensson.

2. A model of the policy problem with judgment

Consider the following linear model of an economy with a private sector and a central bank,[9] in a form that allows for both predetermined and forward-looking variables as well as judgment,

$$\begin{bmatrix} X_{t+1} \\ Cx_{t+1|t} \end{bmatrix} = A \begin{bmatrix} X_t \\ x_t \end{bmatrix} + Bi_t + \begin{bmatrix} z_{t+1} \\ 0 \end{bmatrix}. \tag{2.1}$$

Here, X_t is a (column) n_X-vector of *predetermined* variables (one of these may be unity to conveniently incorporate constants in the model) in period t; x_t is an n_x-vector of *forward-looking* variables; i_t is an n_i-vector of central-bank *instruments* (the forward-looking variables and the instruments are the *nonpredetermined* variables); z_t is an exogenous n_X-vector stochastic process and called the *deviation* in period t; and A, B, and C are matrices of the appropriate dimension. For any variable q_t, I let $q_{t+\tau|t}$ denote private-sector expectations of the realization in period $t+\tau$ of $q_{t+\tau}$ conditional on private-sector information available in period t. I assume that the private sector has rational expectations, given its information.

For increased generality, the model is formulated in terms of an arbitrary number of instruments, n_i. In most practical applications, monetary policy can be seen as having only one instrument—a short interest rate, the instrument rate—so then $n_i = 1$.

The upper block of (2.1) provides n_X equations determining the n_X-vector X_{t+1} in period $t+1$ for given X_t, x_t, i_t and z_{t+1},

$$X_{t+1} = A_{11}X_t + A_{12}x_t + B_1i_t + z_{t+1}, \tag{2.2}$$

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

$$A \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B \equiv \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}. \tag{2.3}$$

The realization of the deviation and the predetermined variables in each period occurs and is observed by the private sector and the central bank in the beginning of the period (the realization of z_{t+1} can be inferred from X_{t+1}, X_t, x_t and i_t and (2.2)).[10]

The lower block of (2.1) provides n_x equations determining the n_x-vector x_t in period t for given $x_{t+1|t}$, X_t, and i_t,

$$x_t = A_{22}^{-1}(Cx_{t+1|t} - A_{21}X_t - B_2i_t); \tag{2.4}$$

---

[9] For simplicity, there is no explicit fiscal authority in the model, but such an authority and its behavior can be included in the model (2.1).

[10] See Svensson and Woodford [29] for an analysis of optimal policy in a model with forward-looking variables when the current state of the economy is imperfectly observed and inferred from observed indicators.

I hence assume that the $n_x \times n_x$ submatrix A_{22} is invertible. The realization of X_t is observed by the private sector and the central bank in the beginning of period t; the central bank then sets the instruments, i_t; after observing the instruments, the private sector forms its expectations, $x_{t+1|t}$; and this finally determines the forward-looking variables x_t.

To assume that the deviation appears only in the upper block of (2.1) is not restrictive. Suppose that I have a model where the deviation appears in both blocks,

$$\begin{bmatrix} X^o_{t+1} \\ Cx_{t+1|t} \end{bmatrix} = \begin{bmatrix} A^o_{11} & A^o_{12} \\ A^o_{21} & A^o_{22} \end{bmatrix} \begin{bmatrix} X^o_t \\ x_t \end{bmatrix} + \begin{bmatrix} B^o_1 \\ B^o_2 \end{bmatrix} i_t + \begin{bmatrix} z^o_{1,t+1} \\ z^o_{2,t} \end{bmatrix},$$

By adding the vector z^o_t to the predetermined variables, I can always form a new model of the form (2.1), where

$$X_t \equiv \begin{bmatrix} X^o_t \\ z^o_{2t} \end{bmatrix}, \qquad A \equiv \begin{bmatrix} A^o_{11} & 0 & A^o_{12} \\ 0 & 0 & 0 \\ A^o_{21} & 0 & A^o_{22} \end{bmatrix}, \qquad B \equiv \begin{bmatrix} B^o_1 \\ 0 \\ B^o_2 \end{bmatrix}, \qquad z_t \equiv \begin{bmatrix} z^o_{1t} \\ z^o_2 \end{bmatrix},$$

and there is no deviation in the lower block.

As in Svensson [25], the deviation represents additional determinants—determinants outside the model—of the variables in the economy, the difference between the actual value of a variable and the value predicted by the model. It can be interpreted as model perturbations, as in the literature on robust control.[11] The central bank's mean estimate of future deviations will be identified with the central bank's *judgment*. It represents the unavoidable judgment always applied in modern monetary policy. Any existing model is always an approximation of the true model of the economy, and monetary-policy makers always find it necessary to make some judgmental adjustments to the results of any given model. Such judgmental adjustments could refer to future fiscal policy, productivity, consumption, investment, international trade, foreign-exchange and other risk premia, raw-material prices, private-sector expectations, and so forth. The "add factors" applied to model equations in central-bank projections (Reifschneider, Stockton, and Wilcox [17]) are an example of central-bank judgment. Given this interpretation of judgment and the deviation z_{t+1}, it would be completely misleading to make a simplifying assumption such as the deviation being a simple autoregressive process. In that case, it could just be incorporated among the predetermined variables. Thus, I will refrain from such an assumption and instead leave the dynamic properties of z_{t+1} unspecified, except in a special case when the deviation is a version of a finite-order moving-average process. Generally, the focus will be on the central bank's judgment of the whole sequence of future deviations.

---

[11] See, for instance, Hansen and Sargent [10]. However, that literature deals with the more complex case when the model perturbations are endogenous and chosen by nature to correspond to a worst-case scenario.

More precisely, let the infinite-dimensional period-t random vector $\zeta^t \equiv (z'_{t+1}, z'_{t+2}, ...)'$ (where $'$ denotes the transpose) denote the vector of the (in period t) unknown random vectors z_{t+1}, z_{t+2},... Let the central bank's beliefs in period t about the random vector ζ^t be represented by the infinite-dimensional probability distribution Φ_t with distribution function $\Phi_t(\zeta^t)$. The probability distribution Φ_t may itself be time-varying and stochastic. The central bank is assumed to know the matrices A, B, C, D, and W and the discount factor δ (D, W, and δ refer to the central bank's objectives and are defined below). The private sector is assumed to know the matrices A, B, and C, but may or may not know the central bank's objectives (that is, D, W, and δ). The private sector may or may not have the same beliefs about the future deviations as the central bank.

Let Y_t be an n_Y-vector of *target* variables. For simplicity, these target variables are measured as the difference from a fixed n_Y-vector Y^* of *target levels*. This vector of target levels is held fixed throughout this paper. In order to examine the consequences of shifting target levels, one only needs to replace Y^t by $Y^t - Y^*$ throughout the paper. Let the target variables be given by

$$Y_t = D \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}, \tag{2.5}$$

where D is an $n_Y \times (n_X + n_x + n_i)$ matrix.

Let the central bank's intertemporal loss function in period t be

$$\mathrm{E}_t \sum_{\tau=0}^{\infty} \delta^\tau L_{t+\tau} \equiv \int \sum_{\tau=0}^{\infty} \delta^\tau L_{t+\tau} d\Phi_t(\zeta^t), \tag{2.6}$$

where $0 < \delta \leq 1$ is a discount factor, L_t is the period loss given by

$$L_t = \frac{1}{2} Y_t' W Y_t, \tag{2.7}$$

and W is a symmetric positive semidefinite $n_Y \times n_Y$ matrix. That is, in period t the central bank wants to minimize the expected discounted sum of current and future losses, where the expectation E_t is with respect to the distribution Φ_t.

Since this is a linear model with a quadratic loss function and the random deviations enter additively, the conditions for certainty equivalence are satisfied. Then, as shown in detail in appendix A, the optimal policy in period t need only consider central-bank mean forecasts—*projections*—of all variables, including the infinite-dimensional mean forecast, z^t, of the random vector ζ^t,

$$z^t \equiv \mathrm{E}_t \zeta^t \equiv \int \zeta^t d\Phi_t(\zeta^t).$$

The central-bank projection in period t of the realization of the deviation in period $t+\tau$, $z_{t+\tau}$, is denoted $z_{t+\tau,t}$, so $z^t \equiv (z'_{t+1,t}, z'_{t+2,t}, ...)'$. The projection z^t is identified with the central bank's *judgment*. Under the assumed certainty equivalence, the projection z^t is, for optimal policy, a sufficient statistic for the distribution Φ_t. Although there is genuine uncertainty about the future random deviations, ζ^t, the only thing that matters for policy is the mean, the judgment, z^t. The second and higher moments of ζ^t—the variance, skew, kurtosis, and so forth—do not matter for policy.[12] The judgment can itself be seen as an exogenous infinite-dimensional random vector that is realized in the beginning of each period and summarizes the central bank's relevant information in that period about expected future deviations.

Let $q^t \equiv (q'_{t,t}, q'_{t+1,t}, ...)'$ denote a central-bank projection in period t for any vector of variables $q_{t+\tau}$ ($\tau \geq 0$) (with the exception of ε^t and $\varepsilon_{t+\tau,t}$, to be introduced below), a central-bank mean forecast conditional on central-bank information in period t. (Thus, for variables other than the deviation, the projection also includes also the current value, $q_{t,t} = q_t$.) The central bank then constructs various projections of the endogenous variables to be used in its decision process. These projections of endogenous variables may be conditional on various assumptions. In order to keep private-sector expectations and central-bank projections conceptually distinct, I denote the former by $q_{t+\tau|t}$ and the latter by $q_{t+\tau,t}$ for any variable q_t.

For given judgment, z^t, the projection model of the central bank for the projections (X^t, x^t, i^t, Y^t) in period t—the model the central bank uses in the decision process to consider alternative projections—will be

$$\begin{bmatrix} X_{t+\tau+1,t} \\ Cx_{t+\tau+1,t} \end{bmatrix} = A \begin{bmatrix} X_{t+\tau,t} \\ x_{t+\tau,t} \end{bmatrix} + Bi_{t+\tau,t} + \begin{bmatrix} z_{t+\tau+1,t} \\ 0 \end{bmatrix}, \tag{2.8}$$

$$Y_{t+\tau,t} = D \begin{bmatrix} X_{t+\tau,t} \\ x_{t+\tau,t} \\ i_{t+\tau,t} \end{bmatrix} \tag{2.9}$$

for $\tau \geq 0$, where $X_{t,t}$ satisfies

$$X_{t,t} = X_t, \tag{2.10}$$

since the realization of the predetermined variables is assumed to be observed in the beginning of period t.

In order to introduce more compact notation, let the $(n_X + n_x + n_i)$-vector $s_t \equiv (X'_t, x'_t, i'_t)'$ denote the state of the economy in period t, and let the vector $s_{t+\tau,t} \equiv (X'_{t+\tau,t}, x'_{t+\tau,t}, i'_{t+\tau,t})'$ denote the

---

[12] The variance of the future deviations will add a term to the intertemporal loss, but that term is independent of policy.

9

projection in period t of the state of the economy in period $t+\tau$. Finally, let the infinite-dimensional vector $s^t \equiv (s'_{t,t}, s'_{t+1,t}, s_{t+2,t}, ...)'$ denote a projection in period t of the (current and future) states of the economy. By (2.9), I can write the projection of the target variables in a compact way, as a linear function of the projection of the states of the economy, as

$$Y^t = \tilde{D}s^t, \tag{2.11}$$

where $\tilde{D}$ is an infinite-dimensional block-diagonal matrix with the $\tau + 1$-th diagonal block equal to D for $\tau \geq 0$.

The *set of feasible projections of the states of the economy* in period t, $\mathcal{S}_t$, can now be defined as the set of projections s^t that satisfy (2.8)-(2.10) for given X_t and z^t.

The intertemporal loss function (2.6) with (2.7) induces an intertemporal loss function for the target-variable projection,[13]

$$\mathcal{L}(Y^t) \equiv \sum_{\tau=0}^{\infty} \delta^\tau Y'_{t+\tau,t} W Y_{t+\tau,t}, \tag{2.12}$$

The *policy problem* in period t is to find the optimal policy projection $(\hat{s}^t, \hat{Y}^t)$, the projection that minimizes (2.12) over the set of feasible projections of the states of the economy, that is, subject to (2.8)-(2.11) for $\tau \geq 0$ for given X_t and z^t. More compactly,

$$\hat{s}^t = \arg \min_{s^t \in \mathcal{S}_t} \mathcal{L}(\tilde{D}s^t), \tag{2.13}$$

$\hat{Y}^t \equiv \tilde{D}\hat{s}$; and the optimal policy projection $(\hat{X}^t, \hat{x}^t, \hat{i}^t)$ of the predetermined variables, forward-looking variables, and instruments can be extracted from $\hat{s}^t$.

The policy problem will be further specified below to correspond to commitment in a "timeless perspective," in order to avoid any time-consistency problems (see Woodford [36] and Svensson and Woodford [30]).

2.1. Implementation and what information the private sector needs

The *implementation* of the optimal policy in period t involves announcing the optimal policy projection and setting the instruments in period t equal to the first element of the instrument projection,

$$i_t = \hat{i}_{t,t}.$$

---

[13] Note that $\min \mathrm{E}_t \sum_{\tau=0}^{\infty} \delta^\tau L_{t+\tau} = \min\{\mathcal{L}(Y^t) + \mathrm{E}_t \sum_{\tau=0}^{\infty} \delta^\tau (Y_{t+\tau} - Y_{t+\tau,t})' W (Y_{t+\tau} - Y_{t+\tau,t})\} = \min\{\mathcal{L}(Y^t) + \sum_{\tau=0}^{\infty} \delta^\tau \mathrm{trace}(W\mathrm{Cov}_t Y_{t+\tau})\}$. By certainty equivalence, $\mathrm{Cov}_t Y_{t+\tau} \equiv \mathrm{E}_t (Y_{t+\tau} - Y_{t+\tau,t})(Y_{t+\tau} - Y_{t+\tau,t})' \equiv \int (Y_{t+\tau} - Y_{t+\tau,t})(Y_{t+\tau} - Y_{t+\tau,t})' d\Phi_t(\zeta^t)$ is independent of policy, so minimizing (2.12) in period t implies the same policy as minimizing (2.6) in period t.

Furthermore, note that, since $\mathrm{trace}(W\mathrm{Cov}_t Y_{t+\tau})$ will normally be strictly positive, (2.6) will normally converge only for $\delta < 1$, whereas (2.12) will normally converge also for $\delta = 1$.

In period $t + 1$, conditional on new realizations of the predetermined variables, X_{t+1}, and the judgment, z^{t+1}, a new optimal policy projection, $(\hat{X}^{t+1}, \hat{x}^{t+1}, \hat{\imath}^{t+1}, \hat{Y}^{t+1})$, is found and announced together with a new instrument setting,

$$i_{t+1} = \hat{\imath}_{t+1,t+1}.$$

In a forward-looking model, the private sector (including policymakers other than the central bank) will need to know at least parts of the aggregate projections $\hat{X}^t, \hat{x}^t$, and $\hat{\imath}^t$, in order to make decisions consistent with these and make the rational-expectations equilibrium in the economy correspond to the central bank's optimal policy projection. If the private sector knows the matrices A, B, C, D, and W and the discount factor δ and has the same judgment z^t as the central bank, it can in principle compute the optimal policy projection itself—assuming that it has the same computational capacity as the central bank.

However, the private sector actually needs to know less. An assumption maintained through this paper is that the private sector knows the model (2.1), in the sense of knowing the matrices A, B, and C. Furthermore, it observes X_t (determined by X_{t-1}, x_{t-1}, and i_{t-1} in period $t - 1$ and the realization of z_t in the beginning of period t according to (2.2)) in the beginning period t, then observes $i_t = \hat{\imath}_{t,t}$ set by the central bank, thereafter forms one-period-ahead expectations $x_{t+1|t}$, and finally determines (and thereby knows) x_t; after this, period t ends. In order to make decisions in period t consistent with the optimal policy projection—that is, decisions resulting in $x_t = \hat{x}_{t,t}$ from (2.4)—the private sector needs be able to form expectations $x_{t+1|t} = \hat{x}_{t+1,t}$. The most direct way is if the central bank announces $\hat{x}_{t+1,t}$ and the private sector believes the announcement. Formally, $\hat{x}_{t+1,t}$ is the minimum additional information the private sector needs. However, the central bank may have to provide the whole optimal policy projection, and also motivate the underlying judgment, in order to demonstrate the optimal policy projections are *internally consistent* with the model (2.1). In particular, the private sector may not believe $\hat{x}_{t+1,t}$ unless it is apparently consistent with the whole projection $\hat{\imath}^t$. Furthermore, the private sector will need to know the central bank's loss function—D, W, and δ—in order to judge whether the projections announced are really optimal relative to the central bank's loss function and thereby *incentive-compatible*, *credible*. Only then may the central bank be able to *convince* the private sector to form expectations according to the optimal policy projection.[14] Indeed, the private sector completely trusting the central bank's isolated announcement of $\hat{x}_{t+1,t}$ could invite misleading announcements, given the time-consistency

---

[14] Being explicit about the loss function and announcing the optimal policy projection also seem to take care of the criticism of real-world inflation targeting expressed by Faust and Henderson [6].

problem discussed below.[15] [16]

2.2. Judgment as a finite-order moving average

Consider the special case when the deviation is a version of a moving-average stochastic process with a given finite order $T > 0$ (where T could be relatively large),

$$z_{t+1} = \varepsilon_{t+1} + \sum_{j=1}^{T} \varepsilon_{t+1,t+1-j}, \tag{2.14}$$

where $\tilde{\varepsilon}_t \equiv (\varepsilon'_t, \varepsilon^{t'})' \equiv (\varepsilon'_t, \varepsilon'_{t+1,t}, ..., \varepsilon'_{t+T,t})'$ is a zero-mean iid random $(T+1)n_X$-vector realized in the beginning of period t and called the innovation in period t.[17] For $T = 0$, z_{t+1} is a simple iid disturbance. For $T > 0$, the deviation is a version of a moving-average process.

It follows that the central-bank judgment $z_{t+\tau,t}$ $(\tau \geq 1)$ is also a finite-order moving average and satisfies

$$z_{t+\tau,t} \equiv \mathrm{E}_t z_{t+\tau} = \sum_{j=\tau}^{T} \varepsilon_{t+\tau,t+\tau-j} = \varepsilon_{t+\tau,t} + \sum_{j=\tau+1}^{T} \varepsilon_{t+\tau,t+\tau-j} = \varepsilon_{t+\tau,t} + z_{t+\tau,t-1}.$$

Hence, $\varepsilon_{t+\tau,t} = z_{t+\tau,t} - z_{t+\tau,t-1}$ can be interpreted as the innovation in period t to the previous judgment $z_{t+\tau,t-1}$, the new information the central bank receives in period t about the realization of $z_{t+\tau}$ in period $t + \tau$. Hence, the judgment $z_{t+\tau,t}$ in period t is the sum of current and previous information about $z_{t+\tau}$. For horizons larger than T, the central-bank judgment is constant and, without loss of generality, equal to zero,

$$z_{t+\tau,t} = 0 \qquad (\tau > T). \tag{2.15}$$

The dynamics of the deviation z_t and the judgment z^{t+1} can then be written as

$$\begin{bmatrix} z_{t+1} \\ z^{t+1} \end{bmatrix} = A_z \begin{bmatrix} z_t \\ z^t \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \varepsilon^{t+1} \end{bmatrix}, \tag{2.16}$$

where the $(T+1)n_X \times (T+1)n_X$ matrix A_z is defined as

$$A_z \equiv \begin{bmatrix} 0_{n_X \times n_X} & I_{n_X} & 0_{n_X \times (T-1)n_X} \\ 0_{(T-1)n_X \times n_X} & 0_{(T-1)n_X \times n_X} & I_{(T-1)n_X} \\ 0_{n_X \times n_X} & 0_{n_X \times n_X} & 0_{n_X \times (T-1)n_X} \end{bmatrix},$$

---

[15] See Geraats [7] for such examples.

[16] In a much noted contribution, Morris and Shin [16] and Amato, Morris, and Shin [1] have emphasized the possibility that public information may be bad and reduce social welfare by crowding out private information. Svensson [27] scrutinizes this results and shows that, in the model considered by Morris and Shin, public information actually increases social welfare for reasonable parameters.

[17] Note that $\varepsilon^t \equiv (\varepsilon'_{t+1,t}, \varepsilon'_{t+2,t}, ..., \varepsilon'_{t+T,t})'$ here denotes a random vector realized in the beginning of period t and not the projection in period t of the random variables $\varepsilon_{t+1}, \varepsilon_{t+2}, ..., \varepsilon_{t+T}$. That projection is always zero under the above assumption of ε^t being a zero-mean iid random variable.

where $0_{m \times n}$ and I_m denote an $m \times n$ zero matrix and an $n \times n$ identity matrix, respectively. Thus, the dynamics of the deviation and the judgment has a convenient linear and recursive representation.

The modeling of the dynamics of the deviation, z_t, and the additive judgment, z^t, in (2.16) allows for a relatively flexible accumulation of information about future deviations. Whereas the stochastic process for the deviation is not a simple Markov process in terms of itself but a finite-order moving-average process, it can be written as a higher-dimensional AR(1) process. The restriction imposed is that the innovation is zero-mean and iid across periods. There is no restriction of the variance and covariance of the elements of $\tilde{\varepsilon}_t$ within the period. It follows that, for instance, $\varepsilon_{t+\tau,t}$ may have a variance that is decreasing in τ, corresponding to a situation where there is less information about the mean projection of deviations further into the future; by assumption, there is no specific information about the deviation for $\tau > T$. For given t, there may be serial correlation of $\varepsilon_{t+\tau,t}$ across τ, corresponding to new information about serially correlated future deviations.

2.3. Representing optimal policy projections

Without the judgment terms (or, alternatively, with the deviation being an iid zero-mean process or an autoregressive process with iid shocks), the above infinite-horizon linear-quadratic problem with forward-looking variables is a well-known problem, examined and solved in Backus and Driffill [2], Currie and Levine [5], and Söderlind [20]. The traditional way to find a solution to this problem is to derive the first-order conditions for an optimum and combine the first-order conditions with the model (2.1) to form a system of difference equations with an infinite horizon. The solution can then also be expressed as a difference equation. Furthermore, Marcet and Marimon [14] have shown a new practial way of reformulating the problem with forward-looking variables as a recursive saddlepoint problem (see appendix D).

A new element here is the solution with the judgment. For the case when the deviation is a finite-order moving average, the dynamics of the deviation and the judgment, (2.16), can be incorporated with (2.1), the vector of predetermined variables can be expanded to include z^t, and the standard solution can be applied directly.[18] The details for this case are provided in appendices C and D. When the judgment is a realization of an infinite-dimensional random vector, the standard solution has to be modified to take that into account. The details of that solution in the form of a difference equation are explained in appendices A and B. Here I shall first report the solution in the form of an infinite-horizon difference equation and later develop a very convenient finite-horizon version of

---

[18] Since z_t is incorporated in X_t, one does not need to add z_t as a separate predetermined variable.

the model.

Under the assumption of optimization under commitment, one way to describe the optimal policy projection is by the following difference equations,

$$
\begin{bmatrix} \hat{x}_{t+\tau,t} \\ \hat{\imath}_{t+\tau,t} \end{bmatrix} = F \begin{bmatrix} \hat{X}_{t+\tau,t} \\ z^{t+\tau,t} \\ \Xi_{t+\tau-1,t} \end{bmatrix}, \tag{2.17}
$$

$$
\begin{bmatrix} \hat{X}_{t+\tau+1,t} \\ \Xi_{t+\tau,t} \end{bmatrix} = M \begin{bmatrix} \hat{X}_{t+\tau,t} \\ z^{t+\tau,t} \\ \Xi_{t+\tau-1,t} \end{bmatrix}, \tag{2.18}
$$

for $\tau \geq 0$, where $\hat{X}_{t,t} = X_t$. When the deviation is a finite-order moving average and the judgment is finite-dimensional, $z^{t+\tau,t}$ denotes the Tn_X-vector $(z'_{t+\tau+1,t}, z'_{t+\tau+2,t}, ..., z'_{t+\tau+T,t})'$, where $z_{t+\tau+j,t} = 0$ for $j + \tau > T$. When the judgment is infinite-dimensional, $z^{t+\tau,t}$ denotes the infinite-dimensional vector $(z'_{t+\tau+1,t}, z'_{t+\tau+2,t}, ...)'$. In the former case, F and M are finite-dimensional matrices. In the latter case, F and M include a linear operator R on $z^{t+\tau,t}$ (an infinite-dimensional matrix) of the form $\sum_{j=0}^{\infty} R_j z_{t+1+\tau+j,t}$, where $\{R_j\}_{j=0}^{\infty}$ is a sequence of matrices. The matrices F, M, and $\{R_j\}_{j=0}^{\infty}$ depend on A, B, C, D, W, and δ, but they are independent of the second and higher moments of the deviation. The n_X-vector $\Xi_{t+\tau,t}$ consists of the Lagrange multipliers of the lower block of (2.8), the block determining the projection of the forward-looking variables.

As discussed in appendix A, the value of the initial Lagrange multiplier, $\Xi_{t-1,t}$, is zero, if there is commitment from scratch in period t, that is, disregarding any previous commitments. This reflects a time-consistency problem when there is reoptimization and recommitment in later periods, as is inherently the case in practical monetary policy. Instead, I assume that the optimization is under commitment in a timeless perspective. Then, if the optimization, and reoptimization, under commitment in a timeless perspective started in an earlier period and has occurred since then, the initial value of the Lagrange multiplier satisfies

$$
\Xi_{t-1,t} = \Xi_{t-1,t-1}, \tag{2.19}
$$

where $\Xi_{t-1,t-1}$ denotes the Lagrange multiplier of the lower block of (2.8) for the determination of $x_{t-1,t-1}$ in the decision problem in period $t-1$. The dependence of the optimal policy projection in period t on this Lagrange multiplier from the decision problem in the previous period makes the optimal policy projection depend on previous projections and illustrates the history dependence of optimal policy under commitment in a forward-looking model shown in Backus and Driffill [2] and Currie and Levin [5] and especially examined and emphasized in Woodford [36].

14

It follows from (2.17)-(2.19) and (2.11) that the optimal policy projection of the states of the economy, the target variables, and the instruments will be linear functions of X_t, z^t, and $\Xi_{t-1,t-1}$, which can be written in a compact way as

$$\hat{s}^t = H \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}, \qquad \hat{Y}^t = \tilde{D}H \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}, \qquad \hat{\imath}^t = H_i \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}, \qquad (2.20)$$

where H is an appropriately formed infinite-dimensional matrix, and H_i is an infinite-dimensional submatrix of H consisting of the rows corresponding to the instruments. In particular, the instrument setting in period t will be given by

$$i_t = \hat{\imath}_{t,t} = h \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}, \qquad (2.21)$$

where the finite- or infinite-dimensional matrix h consists of the n_i first rows of the matrix H_i.

As explained in Svensson and Woodford [30], a simple way of imposing the timeless perspective is to add a term to the intertemporal loss function (2.12),

$$\mathcal{L}(Y^t) + \Xi_{t-1,t-1} \frac{1}{\delta} C x_{t,t}. \qquad (2.22)$$

In the policy problem in period $t-1$, $\Xi_{t-1,t-1}C$ can be interpreted as the marginal loss in period $t-1$ of a change in the one-period-ahead projection of the forward-looking variables, $x_{t,t-1}$. The time-consistency problem arises from disregarding that marginal loss in the policy problem in period t. Adding the corresponding term to the loss function in period t as in (2.22) handles the time-consistency problem, and the optimal policy under commitment in the timeless perspective will result from minimizing (2.22) subject to (2.8)-(2.10) for given X_t, z^t, and $\Xi_{t-1,t-1}$.[19] Since $x_{t,t}$ is an element of the projection s^t, the optimal policy projection $\hat{s}^t$ is then defined as

$$\hat{s}^t = \arg\min_{s^t \in \mathcal{S}_t} \{\mathcal{L}(\tilde{D}s^t) + \Xi_{t-1,t-1} \frac{1}{\delta} C x_{t,t}\}, \qquad (2.23)$$

for given X_t, z^t, and $\Xi_{t-1,t-1}$.

From (2.18) follows that the Lagrange multiplier $\Xi_{t,t}$, to be used in the decision problem in period $t+1$, will be given by

$$\Xi_{t,t} = H_\Xi \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}, \qquad (2.24)$$

---

[19] Alternatively, as discussed in Giannoni and Woodford [8] and Svensson and Woodford [30], one can impose the constraint

$$x_{t,t} = F_x \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix},$$

where F in (2.17) is suitably partitioned. In the present context, it is more practical to add the term to the intertemporal loss function as in (2.22).

where H_Ξ is a finite- or infinite-dimensional matrix.

Let the *set of feasible target-variable projections* in period t, $\mathcal{Y}_t$, be defined as the set of target-variable projections satisfying (2.11) for projections s^t in the set S_t for given X_t and z^t. In the special case where the forward-looking variables, x_t, happen to be target variables and elements in Y_t, so $x_{t,t}$ is an element of Y^t, the optimal target-variable projection, $\hat{Y}^t$, can be defined as the target-variable projection Y^t that minimizes (2.22) on the set $\mathcal{Y}_t$, for given X_t, z^t, and $\Xi_{t-1,t-1}$,

$$\hat{Y}^t = \arg \min_{Y^t \in \mathcal{Y}_t} \{\mathcal{L}(Y^t) + \Xi_{t-1,t-1} \frac{1}{\delta} C x_{t,t}\}.$$

However, in the more general case when some or all forward-looking variables are not target variables, $x_{t,t}$ is not an element of Y^t, and the optimal policy projection has to be found by optimization over the set $\mathcal{S}_t$, as in (2.23).

2.4. Backward-looking model

In a backward-looking model, there are no forward-looking variables: $n_x = 0$. There is no lower block in (2.1) and (2.8), and there are no forward-looking variables in (2.5) and (2.9). There are no projections of forward-looking variables and Lagrange multipliers in (2.17), (2.18), (2.20), and (2.21). There is no time consistency problem and no need to consider commitment in a timeless perspective.

Hence, for a backward-looking model, the optimal target-variable projection can always be found by minimizing (2.12) over the set of feasible target-variable projections,

$$\hat{Y}^t = \arg \min_{Y^t \in \mathcal{Y}_t} \mathcal{L}(Y^t),$$

for given X_t and z^t.

2.5. The complex reduced-form reaction function need not be made explicit

The compact notation for the determination of the period-t instrument i_t in (2.21) and the Lagrange multiplier $\Xi_{t,t}$ in (2.24) may have given the impression that optimal monetary policy is just a matter of calculating the finite- or infinite-dimensional matrices h and H_Ξ once and for all; then, in each period, first observe X_t, form z^t, and recall $\Xi_{t-1,t-1}$ from last period's decision; then simply compute, announce, and implement i_t from (2.21); and finally compute $\Xi_{t,t}$ to be used in next period's decision.

This is a misleading impression, though. First, h and H_{Ξ} are indeed high- or infinite-dimensional and therefore difficult to grasp and interpret. Second, z^t is also high- or infinite-dimensional. It is difficult to conceive of policymakers or even staff pondering pages and pages, or computer screens and computer screens, of huge arrays of numbers in small print, arguing and debating about adjustments of the numbers of z^t, such as the numbers in rows 220–250 and 335–385. Third, no central bank (certainly no central bank that I have any more thorough information about) behaves in that way, and is ever likely to behave that way. Instead, the practical presentation of information and options to policymakers is always in the form of multiple graphs, modest-size tables, and modest amounts of text.

Fourth, the intertemporal loss function $\mathcal{L}(Y^t)$ has the projections of the target variables as its argument. What matters for the construction of the target variables is the whole projection path of the instruments, not the current instrument setting. The obvious conclusion is that the relevant objects of importance in the decision process are the whole projection paths of the target variables and the instruments, not the current instrument setting or projections of the target variables at some fixed horizon. These projection paths are most conveniently illustrated as graphs. Indeed, graphs of projections are prominent in the existing monetary-policy reports where projections are reported. The analytical techniques discussed in this paper should predominantly be seen as techniques for computer-generated graphs of whole projection paths. Pondering such graphs is an essential part of the monetary-policy decision process. Importantly, policymakers need not know the underlying detailed high- or infinite-dimensional matrices behind the construction of those graphs. Therefore, the complex reduced-form reaction functions embedded in these matrices need not be made explicit.

Fifth, in the discussion in section 2.1, there was no reference to the reaction function, only to the optimal policy projection. Given X_t and i_t, the private sector needs to be able to form the expectations $x_{t+1|t}$ in order to make decisions in period t. The minimum for this is the central bank's announcement of $\hat{x}_{t+1,t}$. In order to make that announcement credible, the central bank may have to announce the complete optimal policy projection and motivate its judgment. But it does not need to announce any reaction function. In principle, given the reaction function, the private sector could combine the reaction function with the model and solve for the optimal policy projection, but that is an overwhelmingly complicated and roundabout way.

3. A finite-horizon projection model

Regardless of whether the judgment is finite- or infinite-dimensional (that is, whether (2.15) holds or not), the problem of minimizing the intertemporal loss function is an infinite-horizon problem. From a practical and computational point of view, it is convenient to transform the infinite-horizon policy problem above to a finite-horizon one. When the judgment satisfies (2.15), this can be done in a simple and approximate, but arbitrarily close to exact, way for the forward-looking model, and in a simple and exact way for the backward-looking model. The finite-horizon model also makes it very easy to incorporate any arbitrary constraints on the projections, for instance, a particular form of the instrument projection, such as a constant instrument for some periods. Then, all the relevant projection paths are computed in one simple step.

3.1. Forward-looking model

Suppose that the estimate of the deviation is constant beyond a fixed horizon T. Without loss of generality, assume that the constant is zero.[20] That is, I assume (2.15).

Start by writing the projection model (2.8) and (2.10) for $\tau = 0, ..., T-1$ as

$$X_{t,t} = X_t, \tag{3.1}$$

$$-\tilde{A}s_{t+\tau,t} + \begin{bmatrix} X_{t+\tau+1,t} \\ Cx_{t+\tau+1,t} \end{bmatrix} = \begin{bmatrix} z_{t+\tau+1,t} \\ 0 \end{bmatrix} \qquad (\tau = 0, ..., T-1). \tag{3.2}$$

where $\tilde{A}$ is the $(n_X + n_x) \times (n_X + n_x + n_i)$ matrix defined by $\tilde{A} \equiv [A \ B]$. The first n_X equations of the last block of $n_X + n_x$ equations in (3.2) determine $X_{t+T,t}$ for given $X_{t+T-1,t}$, $x_{t+T-1,t}$, $i_{t+T-1,t}$, and $z_{t+T,t}$. The last n_x equations of that block are

$$-A_{21}X_{t+T-1,t} - A_{22}x_{t+T-1,t} - B_2 i_{t+T-1,t} + Cx_{t+T,t} = 0.$$

They determine $x_{t+T-1,t}$ for given $X_{t+T-1,t}$ and $i_{t+T-1,t}$, and, importantly, for given $x_{t+T,t}$. A problem is that n_x equations determining $x_{t+T,t}$ are lacking. I will assume that $x_{t+T,t}$ is determined by the assumption that $x_{t+T+1,t}$ is equal to its steady-state level. That is, I assume that the optimal policy projection has the property that, for (2.15), it approaches a steady state when $T \to \infty$. This is true for the models and loss functions considered here. Without loss of generality, I assume that the steady-state values for the forward-looking variables are zero,

$$x_{t+T+1,t} = 0. \tag{3.3}$$

---

[20] If the estimate of the deviation from horizon T on is constant but nonzero, it can be incorporated among other constants in the model. If the estimate of the deviation from horizon T on is not constant but follows an autoregressive process (for instance, if it is assumed to gradually approach a constant), it can be incorporated among the predetermined variables.

From this follows that $X_{t+T,t}$, $x_{t+T,t}$, and $i_{t+T,t}$ must satisfy

$$-A_{21}X_{t+T,t} - A_{22}x_{t+T,t} - B_2 i_{t+T,t} = 0, \tag{3.4}$$

which gives me the desired n_x equations for $x_{t+T,t}$.

Let s^t, the projection of the states of the economy, now denote the *finite-dimensional* $(T+1) \times (n_X + n_x + n_i)$-vector $s^t \equiv (s'_{t,t}, s'_{t+1,t}, ..., s'_{t+T,t})'$. Similarly, let all projections q^t for $q = X$, x, i and Y now denote the finite-dimensional vector $q^t \equiv (q'_{t,t}, q'_{t+1,t}, ..., q'_{t+T,t})'$. Finally, let z^t be the Tn_X-vector $z^t \equiv (z'_{t+1,t}, z'_{t+2,t}, ..., z'_{t+T,t})'$ (recall that z^t does not include the component z_t).

The finite-horizon projection model for the projection of the states of the economy, s^t, then consists of (3.1), (3.2) and (3.4). It can be written compactly as

$$G s^t = g^t, \tag{3.5}$$

where G is the $(T+1)(n_X + n_x) \times (T+1)(n_X + n_x + n_i)$ matrix formed from the matrices on the left side of (3.1), (3.2), and (3.4), and g^t is a $(T+1)(n_X + n_x)$-vector formed from the right side of (3.1), (3.2), and (3.4) as $g^t \equiv (X'_t, z'_{t+1,t}, 0', z'_{t+2,t}, 0', ..., z'_{t+T,t}, 0', 0')'$ (where zeros denote zero vectors of appropriate dimension).

Since Y^t now denotes the finite-dimensional $(T+1)n_Y$-vector $Y^t \equiv (Y'_{t,t}, Y'_{t+1,t}, ..., Y'_{t+T,t})'$, I can write

$$Y^t = \tilde{D} s^t, \tag{3.6}$$

where $\tilde{D}$ now denotes a finite-dimensional $(T+1)n_Y \times (T+1)(n_X + n_x + n_i)$ block-diagonal matrix with the matrix D in each diagonal block.

The set of feasible projections, $\mathcal{S}_t$, is then defined as the finite-dimensional set of s^t that satisfy (3.5) and (3.6) for a given g^t, that is, for a given X_t and z^t.

It remains to specify the intertemporal loss function in the forward-looking model in the finite-horizon case. In the forward-looking model, under assumption (2.15), the minimum loss from the horizon $T+1$ on depends on the projection of the predetermined variables for period $t+T+1$, $X_{t+T+1,t}$, and the Lagrange multipliers $\Xi_{t+T,t}$ according to the quadratic form

$$\frac{1}{2} \delta^{T+1} \begin{bmatrix} X'_{t+T+1,t} & \Xi'_{t+T,t} \end{bmatrix} V \begin{bmatrix} X_{t+T+1,t} \\ \Xi_{t+T,t} \end{bmatrix},$$

where V is a symmetric positive semidefinite matrix that depends on the matrices A, B, C, D, and W and the discount factor δ (see appendix A). It follows from (2.18) and (2.15) that this quadratic

form can be written as a function $X_{t+T,t}$ and $\Xi_{t+T-1,t}$ as[21]

$$\frac{1}{2}\delta^{T+1} \begin{bmatrix} X'_{t+T,t} & \Xi'_{t+T-1,t} \end{bmatrix} M'VM \begin{bmatrix} X_{t+T,t} \\ \Xi_{t+T-1,t} \end{bmatrix}. \tag{3.7}$$

In principle, I could use (2.18) to keep track of $\Xi_{t+T-1,t}$. However, a simpler way is to extend the horizon T so far that $X_{t+T,t}$ and $\Xi_{t+T-1,t}$ are arbitrarily close to their steady-state levels. Without loss of generality, I assume that the steady-state levels are zero, in which case the above quadratic form is zero, and the loss from horizon T can be disregarded. Checking that $X_{t+T,t}$ is close to zero is straightforward; I will show a practical way to check that $\Xi_{t+T-1,t}$ is also close to zero.[22]

Under this assumption, it follows from (2.9), (2.12), and (3.6) that the intertemporal loss function can be written as a function of s^t as the finite-dimensional quadratic form

$$\frac{1}{2}s^{t\prime}\Omega s^t, \tag{3.8}$$

where Ω is a symmetric positive semidefinite block-diagonal $(T+1)(n_X + n_x + n_i)$ matrix with its $(\tau+1)$-th diagonal block being $\delta^\tau D'WD$ for $0 \leq \tau \leq T$. However, in order to impose the timeless perspective, as explained in section 2.3, I need to add the term

$$\Xi_{t-1,t-1}\frac{1}{\delta}Cx_{t,t}$$

to the loss function, where $\Xi_{t-1,t-1}$ is the relevant Lagrange multiplier from the policy problem in period $t-1$. This term can be written $\omega'_{t-1}s^t$, with the appropriate definition of the $(T+1)(n_X + n_x + n_i)$-vector ω_{t-1} as $\omega_{t-1} \equiv (0,'0', (\Xi_{t-1,t-1}\frac{1}{\delta}C)', 0', ..., 0')'$ (where the zeros denote zero vectors of appropriate dimension). Thus, the intertemporal loss function with the added term is

$$\frac{1}{2}s^{t\prime}\Omega s^t + \omega'_{t-1}s_t. \tag{3.9}$$

Then, the policy problem is to find the optimal policy projection $\hat{s}^t$ that minimizes (3.9) subject to (3.5). The Lagrangian for this problem is

$$\frac{1}{2}s^{t\prime}\Omega s^t + \omega'_{t-1}s_t + \Lambda^{t\prime}(Gs^t - g^t), \tag{3.10}$$

where Λ^t is the $(T+1)(n_X + n_x)$-vector of Lagrange multipliers of (3.5). The first-order condition is

$$s^{t\prime}\Omega + \omega'_{t-1} + \Lambda^{t\prime}G = 0.$$

---

[21] The matrix M appearing in (3.7) is the matrix M in (2.18) with the columns corresponding to z^t deleted.

[22] Appendix E presents an iterative numerical procedure that will provide a projection arbitrarily close to the optimal policy projection without requiring such a long horizon that $X_{t+T,t}$ and $\Xi_{t+T-1,t}$ are close to their steady-state levels.

Combining this with (3.5) gives the linear equation system

$$
\begin{bmatrix} G & 0 \\ \Omega & G' \end{bmatrix} \begin{bmatrix} s^t \\ \Lambda^t \end{bmatrix} = \begin{bmatrix} g^t \\ -\omega_{t-1} \end{bmatrix}.
\tag{3.11}
$$

The optimal policy projection $\hat{s}^t$ and Lagrange multiplier Λ^t are then given by the simple matrix inversion[23]

$$
\begin{bmatrix} \hat{s}^t \\ \Lambda^t \end{bmatrix} = \begin{bmatrix} G & 0 \\ \Omega & G' \end{bmatrix}^{-1} \begin{bmatrix} g^t \\ -\omega_{t-1} \end{bmatrix}.
\tag{3.12}
$$

The optimal target-variable projection then follows from (3.6). The optimal policy projection is a linear function of X_t, z^t, and $\Xi_{t-1,t-1}$, and it can be written compactly as in section 2.3,

$$
\hat{s}^t = H \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}, \qquad \hat{Y}^t = \tilde{D} H \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}, \qquad \hat{\imath}^t = H_i \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix},
$$

except that the matrices H and H_i and the vector z^t now are finite-dimensional. The matrices can be directly extracted from (3.12). The period-t instrument setting can be written

$$
i_t = \hat{\imath}_{t,t} = h \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix},
\tag{3.13}
$$

where the finite-dimensional matrix h consists of the first n_i rows of the matrix H_i. Under assumption (2.15) and a sufficiently long horizon T, the finite-horizon projections here are arbitrary close to the optimal infinite-horizon policy projections for $\tau = 0, ..., T$ in section 2.3.

The Lagrange multiplier Λ^t can be written $\Lambda^t \equiv (\frac{1}{\delta}\xi'_{t,t}, \xi'_{t+1,t}, \Xi'_{t,t}, \delta\xi'_{t+2,t}, \delta\Xi'_{t+1,t}, ..., \delta^T \xi'_{t+T,t}, \delta^T \Xi'_{t+T-1,t})'$, where $\xi_{t+\tau,t}$ is the vector of Lagrange multipliers for the block of equations in (3.1), (3.2), and (3.4) determining $X_{t+\tau,t}$ and $\Xi_{t+\tau-1,t}$ is the vector of Lagrange multipliers for the block of equations determining $x_{t+\tau-1,t}$. Hence, extraction of $\Xi_{t+T-1,t}$ from Λ^t allows me to check that the assumption made above of $\Xi_{t+T-1,t}$ being close to zero is satisfied. If the assumption is not satisfied, the horizon T can be extended until the assumption is satisfied.[24] Furthermore, $\Xi_{t,t}$ can be extracted from Λ^t in order to form the vector ω_t to be used in the loss function for the policy problem in period $t+1$, to ensure the timeless perspective. The Lagrange multiplier needed in the loss function in period $t+1$, $\Xi_{t,t}$, can be written

$$
\Xi_{t,t} = H_\Xi \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix},
\tag{3.14}
$$

---

[23] Numerically, it is faster to solve the system of linear equations (3.11) by other methods than first inverting the left-side matrix, see Judd [11].

[24] In practice, the horizon T is extended until the optimal projection $\hat{s}^t$ is insensitive to variations in T.

where the finite-dimensional matrix H_Ξ can be extracted from (3.12).

Again, as noted above in section 2.3, in spite of the compact notation for the instrument i_t and Lagrange multiplier $\Xi_{t,t}$ in (3.13) and (3.14), these analytical techniques should predominantly be seen as techniques for computer-generated graphs to be pondered by the policy makers, and the matrices never need to made explicit to the policymakers. Although the matrices are now formally finite-dimensional, they are still high-dimensional and somewhat difficult to interpret.

3.2. Backward-looking model

Make the same assumption (2.15) as for the forward-looking model. The projection in period t of the state of the economy in period $t + \tau$, $s_{t+\tau,t}$, is now defined as the $(n_X + n_i)$-vector $s_{t+\tau,t} \equiv (X'_{t+\tau,t}, i'_{t+\tau,t})'$ for $\tau \geq 0$, in which case I can write, for the backward-looking model,

$$X_{t+\tau+1,t} = \tilde{A} s_{t+\tau,t} \qquad (\tau \geq T). \tag{3.15}$$

The projection model with horizon T can now be written

$$X_{t,t} = X_t, \tag{3.16}$$

$$-\tilde{A} s_{t+\tau,t} + X_{t+\tau+1,t} = z_{t+\tau+1,t} \qquad (0 \leq \tau \leq T - 1), \tag{3.17}$$

where X_t and z^t are given. The projection of the states of the economy, s^t, is now a $(T+1)(n_X+n_i)$-vector. Then the projection model can be written as (3.5), where G is a $(T+1)n_X \times (T+1)(n_X+n_i)$ matrix formed from the left side of (3.16) and (3.17), and g^t is a $(T+1)n_X$-vector formed from the right side of (3.16) and (3.17) as $g^t \equiv (X'_t, z^{t\prime})'$.

It is a standard result for a linear-quadratic backward-looking model that the minimum loss from the horizon $T + 1$ on depends on the projection of the predetermined variables for period $t + T + 1$, $X_{t+T+1,t}$, according to the quadratic form

$$\frac{1}{2}\delta^{T+1} X'_{t+T+1,t} V X_{t+T+1,t}, \tag{3.18}$$

where V is a symmetric positive semidefinite matrix that depends on the matrices A, B, D, and W and the discount factor δ (see appendix A). I could now, as for the forward-looking model, assume that the predetermined variables approach a steady-state level for large T, without loss of generality assume that the steady-state level is zero, and extend the horizon T so far that the predetermined variables are arbitrarily close to zero, and the loss from period T on is arbitrarily close to zero. I could then form the finite-horizon loss function as for the forward-looking model,

22

and this together with (3.16) and (3.17) would form the finite-horizon model, which would be an arbitrarily close approximation to the infinite-horizon model for sufficiently large T.

However, the absence of the time-consistency problem and the need to keep track of the Lagrange multiplier $\Xi_{t+T-1,t}$ allows a simple approach, which is exact also for small T, as long as assumption (2.15) holds for that T. I follow this approach here.

From (3.15) follows that the quadratic form (3.18) can be written as a function of $s_{t+T,t}$ as

$$\frac{1}{2}\delta^{T+1}s'_{t+T,t}\tilde{A}'V\tilde{A}s_{t+T,t}.$$

The finite-horizon intertemporal loss function can then be written

$$\frac{1}{2}\sum_{\tau=0}^{T}\delta^{\tau}s'_{t+\tau,t}D'WDs_{t+\tau,t} + \frac{1}{2}\delta^{T+1}s'_{t+T,t}\tilde{A}'V\tilde{A}s_{t+T,t}.$$

The intertemporal loss function can now be written more compactly as the quadratic form (3.8), where Ω now is a symmetric positive-semidefinite block-diagonal $(T+1)(n_X+n_i)$-matrix, whose $(\tau+1)$-th diagonal block is $\delta^{\tau}D'WD$ for $0 \leq \tau \leq T-1$ and whose $(T+1)$-th diagonal block now is $\delta^{T}(D'WD + \delta\tilde{A}'V\tilde{A})$. Thus, it differs from the matrix Ω for the forward-looking model by the addition of that last term, $\frac{1}{2}\delta\tilde{A}'V\tilde{A}$.

The finite-horizon policy problem is now to find the optimal policy projection $\hat{s}^{t}$ that minimizes (3.8) subject to (3.5), for given g^{t}, that is, for given X_{t} and z^{t}. The corresponding optimal target-variable projection $\hat{Y}^{t}$ then follows from (3.6).

The Lagrangian for this problem is

$$\frac{1}{2}s^{t\prime}\Omega s^{t} + \Lambda^{t\prime}(Gs^{t} - g^{t}),$$

where Λ^{t} is a vector of Lagrange multipliers for (3.5). The first-order condition is

$$s^{t\prime}\Omega + \Lambda^{t\prime}G = 0.$$

Combining this with (3.5) gives the linear equation system

$$\begin{bmatrix} G & 0 \\ \Omega & G' \end{bmatrix} \begin{bmatrix} s^{t} \\ \Lambda^{t} \end{bmatrix} = \begin{bmatrix} g^{t} \\ 0 \end{bmatrix}.$$

The optimal policy projection $\hat{s}^{t}$ is then given by the simple matrix inversion,

$$\begin{bmatrix} \hat{s}^{t} \\ \Lambda^{t} \end{bmatrix} = \begin{bmatrix} G & 0 \\ \Omega & G' \end{bmatrix}^{-1} \begin{bmatrix} g^{t} \\ 0 \end{bmatrix}. \tag{3.19}$$

The optimal target-variable projection then follows from (3.6).

The optimal policy projection is obviously a linear function of X_t and z^t, and I can write,

$$\hat{s}^t = H \begin{bmatrix} X_t \\ z^t \end{bmatrix}, \qquad \hat{Y}^t = \tilde{D}H \begin{bmatrix} X_t \\ z^t \end{bmatrix}, \qquad \hat{i}^t = H_i \begin{bmatrix} X_t \\ z^t \end{bmatrix},$$

where the finite-dimensional matrices H and H_i can be extracted from (3.19). The instrument setting for period t can be written

$$i_t = \hat{i}_{t,t} = h \begin{bmatrix} X_t \\ z^t \end{bmatrix}, \tag{3.20}$$

where the finite-dimensional matrix h consists of the first n_i rows of the matrix H_i.

3.3. Other considerations

A finite-horizon projection model has several advantages. One is that it is very easy to incorporate any restrictions on the projections. Any equality restriction on X^t, x^t, i^t, or Y^t can be written

$$\bar{R}s^t = \bar{s}^t, \tag{3.21}$$

where the number of rows of the matrix $\bar{R}$ and the dimension of the given vector $\bar{s}^t$ equal the number of restrictions. This makes it easy to incorporate any restriction on the instrument projection, for instance, that it shall be constant or of a particular shape for some periods.

Transforming the model into a finite system of equations may be particularly practical from a computational point of view for a nonlinear model. It may then also be easy to handle commitment in a timeless perspective for a nonlinear model.

4. Monetary policy without judgment

Modern monetary policy inherently to a large extent relies on judgment. Previous sections of this paper have attempted to model this dependence on judgment in a simple but specific way. This section attempts to specify monetary policy without judgment, in order to compare monetary policy with and without judgment.

There are several alternatives in modelling monetary policy without judgment. Above, monetary policy with judgment has been modeled as forecast targeting, finding an instrument projection such that the corresponding projection of the target variables minimizes a loss function. This procedure uses all information available to the central bank, including central-bank judgment. This results in a complex reduced-form reaction function, which fortunately never needs to be made explicit. When modelling monetary policy without judgment, however, the most obvious route is to consider

24

monetary policy as a more mechanical process—a commitment to a particular reaction function, a *commitment to a particular instrument rule.*[25]

Instrument rules can be divided into two categories, explicit instrument rules and implicit instrument rules (Svensson and Woodford [30], Svensson [25]). An *explicit instrument rule* is a reaction function where the instrument responds to predetermined variables only. Its implementation then consists of the central bank observing the predetermined variables in the beginning of the period, and then calculating, announcing, and setting the instrument according to this instrument rule. The implementation obviously requires that the relevant predetermined variables must be observed by the central bank, but since the predetermined variables in a particular period are independent of the instrument setting in that period, no further complications arise. An *implicit* instrument rule is a relation between the current instrument and some of the current forward-looking variables. Then, since the forward-looking variables depend on the instrument setting, the instrument and the forward-looking variables are simultaneously determined. Thus, an implicit instrument rule is actually an *equilibrium condition*, a relation that holds in equilibrium. The implementation of an implicit instrument rule, that is, how to get to the desired equilibrium, is not trivial but a complex issue. This fact has largely been overlooked in the literature, except, for instance, in Svensson and Woodford [30] and Svensson [25].

Here, I shall discuss how monetary policy without judgment can be modeled as a commitment to an instrument rule, either an explicit instrument rule or an implicit instrument rule, with some discussion of the implementation of the latter. I will start from the reduced-form reaction function for the instrument setting that follows from the forecast targeting modeled above. For simplicity, I now consider the realistic case when there is only one instrument, the instrument rate, so $n_i = 1$.

4.1. Explicit instrument rules

The construction of the optimal instrument projection, $\hat{\imath}^t$, in the forward-looking model results in an optimal reduced-form reaction function for the current instrument setting, (3.13) (or (2.21)), which is repeated here as

$$i_t = h_X X_t + h_z z^t + h_\Xi \Xi_{t-1,t-1}, \tag{4.1}$$

where the row vector h is partitioned conformably with X_t, z^t, and $\Xi_{t-1,t-1}$ as $h \equiv [h_X \ \ h_z \ \ h_\Xi]$ (h is now a row vector, since $n_i = 1$). The optimal reaction function implies a particular instrument

---

[25] This is an approach that a large part of the literature has taken, for instance, most papers in the conference volume Taylor [33]. The approach is criticized in Svensson [25] and [26].

response to the predetermined variables, X_t; the judgment, z^t; and the Lagrange multiplier for the equations for $x_{t-1,t-1}$ in the policy problem in period $t-1$, $\Xi_{t-1,t-1}$. The discussion here refers to the forward-looking model; for the backward-looking model, I can simply delete the term $h_\Xi \Xi_{t-1,t-1}$.

I want to model the central bank setting its instrument in a mechanical way, via a commitment to a particular instrument rule, while disregarding judgment. Considering (4.1) as a potential instrument rule, it is natural that disregarding judgment means that the central bank behaves as if it believes that $z^t = 0$ and hence disregards the judgment term in (4.1), $h_z z^t$. Thus, *monetary policy without judgment is modeled as the central bank erroneously believing that the expected future deviations equal zero*, for instance, the central bank believing that z_t is a zero-mean iid process. A first possibility is then that the central bank also disregards the term with the Lagrange multiplier from the previous policy problem. For one thing, if the central bank did set the instrument mechanically in period $t-1$ rather than by explicit optimization, it may not be aware of the Lagrange multiplier and its value. This leaves responding to the current predetermined variables only,

$$i_t = h_X X_t. \tag{4.2}$$

Such a policy would be inefficient for two reasons, even if the response coefficients to X_t are those of the optimal reaction function (4.1). First, it disregards judgment, the term $h_z z^t$. Second, it also disregards the response to lagged predetermined variables implied by the response to the Lagrange multiplier, $\Xi_{t-1,t-1}$, as indicated in (2.19). Indeed, optimization under discretion would result in a reduced-form reaction function where the instrument responds only to the current predetermined variables.[26] [27]

The response to the Lagrange multipliers in (4.1) implies a response to lagged predetermined variables. Disregarding judgment, the Lagrange multiplier in (3.14) (or (2.24)) follows

$$
\begin{aligned}
\Xi_{t-1,t-1} &= H_{\Xi X} X_{t-1} + H_{\Xi\Xi} \Xi_{t-2,t-2} \\
&= \sum_{j=0}^{\infty} H_{\Xi\Xi}^j H_{\Xi X} X_{t-1-j}
\end{aligned}
$$

---

[26] The resulting reduced-form reaction function resulting from optimization under discretion would have different coefficients than the optimal h_X. Because of this, and because of the missing response to lagged predetermined variables, the response is suboptimal and results in so-called stabilization bias relative to the commitment policy (Svensson and Woodford [30]).

[27] Given a particular restricted class of instrument rules, for instance, simple instrument rules with only a few arguments, one can find the optimal instrument rule in that restricted class, see appendix G. The optimal instrument rule in a restricted class will depend on the stochastic properties of the disturbances to the economy. Many papers of Taylor [33] provide examples of such optimal restricted instrument rules.

where H_Ξ is partitioned conformably with X_t, z^t, and $\Xi_{t-1,t-1}$ as $H_\Xi \equiv [H_{\Xi X} \quad H_{\Xi z} \quad H_{\Xi\Xi}]$. A second possibility for policy disregarding judgment is then that the policy responds with the optimal coefficients h_Ξ to these lagged predetermined variables, resulting in the reaction function

$$i_t = h_X X_t + h_\Xi \sum_{j=0}^{\infty} H_{\Xi\Xi}^j H_{\Xi X} X_{t-1-j}. \tag{4.3}$$

This would seem to be an instrument rule corresponding to a rather sophisticated policy, commitment to the reaction function resulting from optimal policy under commitment in a timeless perspective, while disregarding judgment.[28]

4.2. Implicit instrument rules

Another apparent possibility would be an implicit instrument rule, where the instrument responds to the forward-looking variables, x_t. This might seem advantageous, since, in a rational-expectations equilibrium, the forward-looking variables might be affected by private-sector expectations of future deviations. Then, by responding to forward-looking variables, the central bank might indirectly take judgment into account, although in this case private-sector judgment. Thus, one might want to consider an ad hoc implicit instrument rule of the form

$$i_t = f_X X_t + f_x x_t, \tag{4.4}$$

where f_X and f_x are row vectors of given response coefficients.

As mentioned above, there is, however, a specific problem with the central bank responding to forward-looking variables, something largely overlooked in the literature (except, for instance, Svensson and Woodford [30] and Svensson [25]). Since the forward-looking variables depend on the central bank's instrument setting, a simultaneity problem arises. The central bank cannot observe the forward-looking variables before it sets the instrument, and the private sector needs to observe the instrument setting before it determines the forward-looking variables. A relation such

---

[28] Note that, only if h_Ξ is invertible, can this be written on the "instrument smoothing" form

$$
\begin{aligned}
i_t &= h_X X_t + h_\Xi H_{\Xi X} X_{t-1} + h_\Xi H_{\Xi\Xi} h_\Xi^{-1} h_\Xi \sum_{j=0}^{\infty} H_{\Xi\Xi}^j H_{\Xi X} X_{t-1-j} \\
&= h_X X_t + h_\Xi H_{\Xi X} X_{t-1} + h_\Xi H_{\Xi\Xi} h_\Xi^{-1} (i_{t-1} - h_X X_{t-1}) \\
&= h_X X_t + h_\Xi (H_{\Xi X} - H_{\Xi\Xi} h_\Xi^{-1} h_X) X_{t-1} + h_\Xi H_{\Xi\Xi} h_\Xi^{-1} i_{t-1}.
\end{aligned}
$$

If there is only one instrument, h_Ξ is invertible only if $\Xi_{t-1,t-1}$ is a scalar, that is, if there is only one forward-looking variable. The point is that the optimal reaction function under commitment usually cannot be written as an instrument rule involving current predetermined variable and the lagged instrument, unless i_{t-1} happens to be one of the predetermined variables.

as (4.4) is actually an equilibrium condition, where i_t and x_t are simultaneously determined. The implementation of such an equilibrium condition is not straightforward.

A sophisticated way to implement (4.4) is for the central bank to construct projections (X^t, x^t, i^t) that satisfy (4.4). The central bank can amend the relation

$$i_{t+\tau,t} = f_X X_{t+\tau,t} + f_x x_{t+\tau,t} \tag{4.5}$$

for $\tau \geq 0$ to its projection model (3.5) (or (2.8)). It can then solve for the projection $(\tilde{X}^t, \tilde{x}^t, \tilde{i}^t)$ for given X_t, *disregarding the judgment* (setting $z_{t+\tau,t} = 0$ for $\tau \geq 1$). These projections will by construction satisfy (4.5). The corresponding instrument setting, $i_t = \tilde{i}_{t,t}$, will then be a linear function of X_t,

$$i_t = \tilde{f}_X X_t. \tag{4.6}$$

The central bank can then set this instrument level according to the reduced-form reaction function (4.6).[29]

However, if the private sector understands that the central bank is effectively implementing the reaction function (4.6); has rational expectations of future variables; and, in particular, has expectations of future nonzero deviations, $z_{t+\tau|t} \neq 0$ ($\tau \geq 1$); the resulting market-determined forward-looking variables, x_t, will deviate from the central-bank projection, $\tilde{x}_{t,t}$. Thus, although the instrument setting will satisfy

$$i_t = f_X X_t + f_x \tilde{x}_{t,t},$$

for the central-bank projection $\tilde{x}_{t,t}$, it will *not* satisfy (4.4) for the market-determined forward-looking variables x_t.

---

[29] In the context of the finite-horizon projection model, relation (4.5) can be written as $(T+1)n_i$ equations,

$$\bar{R}s^t = 0,$$

where $\bar{R}$ is an $(T+1)n_i \times (T+1)(n_X + n_x + n_i)$ matrix. Combining these with (3.5) for $z^t = 0$ gives an equation system

$$\left[\begin{array}{c} G \\ \bar{R} \end{array} \right] s^t = \left[\begin{array}{c} \tilde{g}^t \\ 0 \end{array} \right],$$

where $\tilde{g}^t$ is the $(T+1)(n_X + n_x)$-vector $(X_t', 0')'$. Under the assumption that the matrix on the left side has full rank, s^t is given by

$$\tilde{s}^t \equiv \left[\begin{array}{c} G \\ \bar{R} \end{array} \right]^{-1} \left[\begin{array}{c} \tilde{g}^t \\ 0 \end{array} \right]. \tag{4.7}$$

This results in

$$\tilde{x}_{t,t} = \tilde{H}_{x0} X_t,$$

where the matrix $\tilde{H}_{x0}$ can be extracted from the right side of (4.7) and

$$\tilde{i}_{t,t} = f_X X_t + f_x \tilde{x}_{t,t} = (f_X + f_x \tilde{H}_{x0}) X_t \equiv \tilde{f}_X X_t.$$

28

For relation (4.4) to be satisfied for the market-determined forward-looking variables, the central bank has to amend the relation (4.5) for $\tau \geq 0$ to its production model (3.5) (or (2.8)) and solve for the projection (X^t, x^t, x^t) for given X_t, *taking the judgment into account.*[30] This results in

$$
\begin{aligned}
x_{t,t} &= \tilde{H}_{xX0}X_t + \tilde{H}_{xz0}z^t \\
i_{t,t} &= f_X X_t + f_x x_{t,t} = (f_X + f_x \tilde{H}_{xX0})X_t + f_x \tilde{H}_{xz0}z^t \\
&\equiv \tilde{f}_X X_t + \tilde{f}_z z^t.
\end{aligned}
$$

Thus, the resulting reduced-form reaction function is

$$
i_t = \tilde{f}_X X_t + \tilde{f}_z z^t.
$$

If the private sector understands that this is the reaction function followed by the central bank and has rational expectations corresponding to the same judgment; the market-determined forward-looking variables, x_t, will equal the central-bank projection, $x_{t,t}$; and the relation (4.4) will be satisfied in equilibrium.

This is of course an example of a central bank explicitly taking judgment into account, not an example of a central bank disregarding judgment. But instead of finding the optimal policy projection, $(\hat{X}^t, \hat{x}^t, \hat{i}^t, \hat{Y}^t)$, that minimizes its loss function, it finds the projection (X^t, x^t, i^t, Y^t) that satisfies the ad hoc relation (4.4). Since the latter is no easier than the former and, in particular, suboptimal, this behavior seems a bit far-fetched.

Thus, I can model monetary policy without judgment as following either the explicit instrument rule (4.2), where both judgment and lagged predetermined variables are ignored, or the more sophisticated explicit instrument rule (4.3), or perhaps some intermediate case of (4.3) where the summation is over a finite past periods. Using the implicit instrument rule (4.4) is somewhat problematic, since its implementation is complex and open to alternative very different interpretations,

---

[30] The equation system is then

$$
\begin{bmatrix} G \\ R \end{bmatrix} s^t = \begin{bmatrix} g^t \\ 0 \end{bmatrix},
$$

where g^t is the $(T+1)(n_X + n_x)$-vector $(X_t', z_{t,t}', 0', z_{t+1,t}', 0', ..., z_{t+T,t}', 0', 0')'$ specified in section 3.1. Under the assumption that the matrix on the left side has full rank, s^t is given by

$$
s^t = \begin{bmatrix} G \\ R \end{bmatrix}^{-1} \begin{bmatrix} g^t \\ 0 \end{bmatrix}. \tag{4.8}
$$

This results in

$$
x_{t,t} = \tilde{H}_{xX0}X_t + \tilde{H}_{xz0}z^t,
$$

where $\tilde{H}_{xX0}$ and $\tilde{H}_{xz0}$ can be extracted from the right side of (4.8) and

$$
i_{t,t} = f_X X_t + f_x x_{t,t} = (f_X + f_x \tilde{H}_{x0})X_t + f_x \tilde{H}_{xz0}z^t \equiv \tilde{f}_X X_t + \tilde{f}_z z^t.
$$

with very different resulting equilibria.

4.3. Taylor rules

In the literature, a number of simple ad hoc instrument rules have been used to discuss and evaluate monetary policy. The most common are variants of the Taylor rule.[31] One variant of the Taylor rule with instrument smoothing (meaning, in this context, a response to the lagged instrument rate) is

$$i_t = (1 - f_i)(f_\pi \pi_t + f_y y_t) + f_i i_{t-1};$$

where π_t denotes a measure of the difference of inflation from a given inflation target; y_t denotes a measure of the output gap; f_π and f_y are given positive coefficients that can be interpreted as the long-run response to inflation and the output gap, respectively; and the coefficient f_i satisfies $0 \leq f_i \leq 1$. If inflation and the output gap are predetermined, this is an explicit instrument rule, and its implementation only requires that the central bank can observe or estimate current inflation and the output gap. If inflation and/or the output gap are forward-looking variables, this is an implicit instrument rule, where the instrument and the forward-looking variable are simultaneously determined. As noted above, such an instrument rule is somewhat problematic and its implementation may need to be further specified.

One variant of the Taylor rule, a so-called forecast-based or forward-looking Taylor rule, can be written

$$i_t = (1 - f_i)[f_\pi \pi_{t+J,t} + f_y y_{t+K,t}] + f_i i_{t-1};$$

where $\pi_{t+J,t}$ denotes a projection of the difference of inflation from an inflation target at horizon $J \geq 0$ and $y_{t+K,t}$ denote a projection of the output gap at horizon $K \geq 0$, where at least one of J or K is positive. Such an instrument rule is an explicit or an implicit instrument rule depending on how the projections are constructed. If the projections are constructed with information that is predetermined in period t, the projections are predetermined and the instrument rule is explicit. If the projections are constructed with information that includes simultaneously determined forward-looking variables, the instrument rule is implicit and hence an equilibrium condition. Again, the implementation of such an instrument rule is not trivial and open to alternative interpretations.[32]

---

[31] Koizicki [12] provides a discussion of the usefuleness of Taylor rules.
[32] Svensson [24] discusses additional serious problems with forecast-based instrument rules.

5. Examples

In this section, I discuss examples of monetary policy with and without judgment in two different empirical models of the U.S. economy: a backward-looking model due to Rudebusch and Svensson [18] and a forward-looking model due to Lindé [13].

5.1. Backward-looking model

The backward-looking empirical model of Rudebusch and Svensson [18] has two equations (with estimates rounded to two decimal points)

$$\pi_{t+1} = 0.70\,\pi_t - 0.10\,\pi_{t-1} + 0.28\,\pi_{t-2} + 0.12\,\pi_{t-3} + 0.14\,y_t + z_{\pi,t+1} \tag{5.1}$$

$$y_{t+1} = 1.16\,y_t - 0.25\,y_{t-1} - 0.10\left(\frac{1}{4}\Sigma_{j=0}^{3}i_{t-j} - \frac{1}{4}\Sigma_{j=0}^{3}\pi_{t-j}\right) + z_{y,t+1}. \tag{5.2}$$

The period is a quarter, and π_t is quarterly GDP inflation measured in percentage points at an annual rate, y_t is the output gap measured in percentage points, and i_t is the quarterly average of the federal-funds rate, measured in percentage points at an annual rate. All variables are measured as differences from their means, their steady-state levels. The deviations $z_{\pi,t+1}$ and $z_{y,t+1}$ for inflation and the output gap have been substituted for the shocks of the original Rudebusch-Svensson model. The predetermined variables are $(\pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-2}, y_t, y_{t-1}, i_{t-1}, i_{t-2}, i_{t-3})$. See appendix H for details.

The target variables are inflation, the output gap, and the first-difference of the federal funds rate. The period loss function is

$$L_t = \frac{1}{2}[\pi_t^2 + \lambda y_t^2 + \nu(i_t - i_{t-1})^2], \tag{5.3}$$

where π_t is measured as the difference from the inflation target, which is equal to the steady-state level. The discount factor, δ, and the relative weights on the output-gap stabilization, λ, and interest-rate smoothing, ν, is set so $\delta = 1$, $\lambda = 1$, and $\nu = 0.2$.

Let me emphasize that there may be considerable uncertainty about the future deviations, ζ^t, in this case $\{z_{\pi,t+\tau}, z_{y,t+\tau}\}_{\tau=1}^{\infty}$. Consider a simple case, when the distribution Φ_t is such that there is uncertainty only about $z_{\pi,t+\tau}$ and only for a finite number of quarters, $1 \le \tau \le T$. Then, I can take ζ^t to be the random T-vector $\zeta^t = (z_{\pi,t+1}, ..., z_{\pi,t+T})$. Suppose furthermore that there are only four possible events with realizations $\zeta^t(j)$ $(j = 1, 2, 3, 4)$, and that these are as follows:

1. With probability 0.4: $\zeta^t(1) = (0, 0, ...)'$, no deviation.

2. With probability 0.3: $\zeta^t(2) = (0, 1.0, 1.0, 0.0, 0.0, 0, ...)'$, a short sequence of large "cost push" shocks.

3. With probability 0.2, $\zeta^t(3) = (0, 0.2, 0.2, 0.2, 0.2, 0, ...)'$, a long sequence of small cost-push shocks.

4. With probability 0.1, $\zeta^t(4) = (0, 1.0, 1.0, 1.0, 1.0, 0, ...)'$, a long sequence of large cost-push shocks.

The resulting judgment is the mean of the future deviations, the T-vector $z^t = 0.4\,\zeta^t(1)' + 0.3\,\zeta^t(2)' + 0.2\,\zeta^t(3)' + 0.1\,\zeta^t(4)' = (0, 0.44, 0.44, 0.14, 0.14, 0, ...)'$.
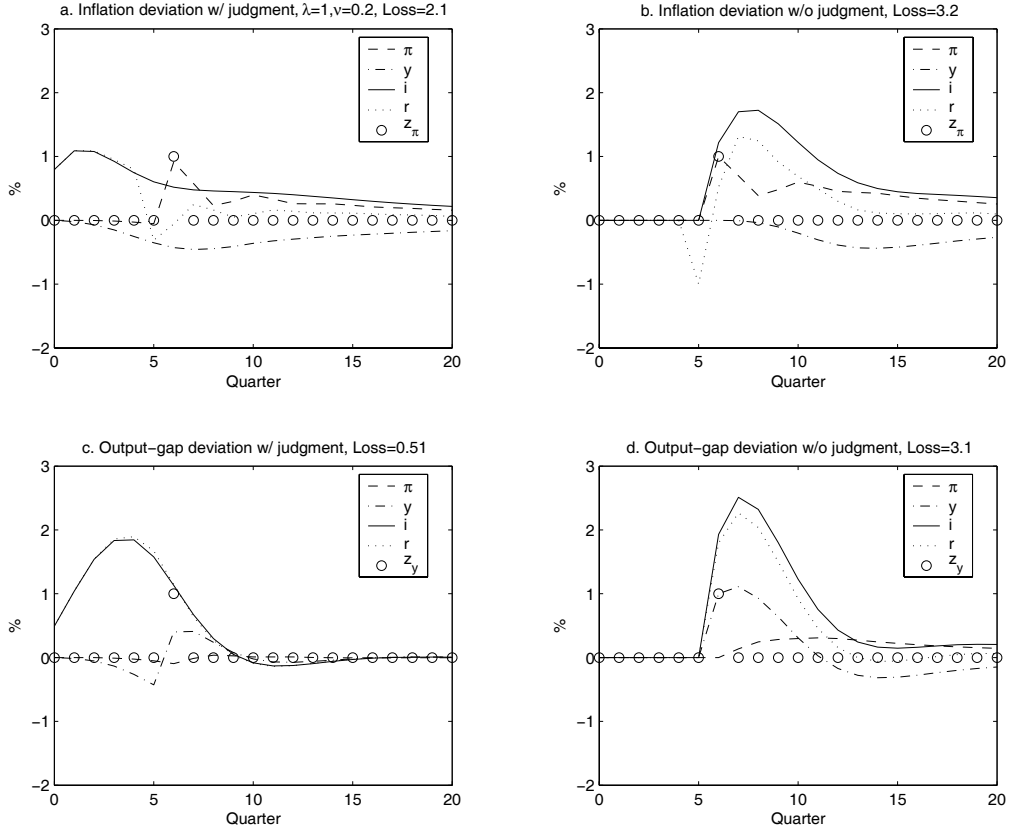
Note that the same judgment arises if the probabilities are the same for the four events but the first event is that all components $\tau = 1, ..., T$ of ζ^t have independent uniform distributions between -1 and 1; the second event is that all components have the same distributions as for the first event except that component $\tau = 2$ and 3 have independent uniform distributions between 0 and 2; the third event is that all components have the same distributions as for the first event except that component $\tau = 2$, 3, 4, and 5 have independent uniform distributions between -0.8 and 1.2; and the fourth event is that all components have the same distribution as for the first event except that component $\tau = 2$, 3, 4, and 5 have independent uniform distributions between 0 and 2. Thus, behind a given judgment vector can be a distribution Φ_t involving considerable uncertainty. Still, only the mean of that distributions matters.

Figure 5.1 shows a situation where the predetermined variables, inflation and the output gap, and the deviations are assumed to have been equal to their steady-state levels, zero, up to quarter 0. Furthermore, in previous quarters, the central bank's judgment, z^t ($t < 0$) has been zero: The central bank's expected future inflation and output-gap deviations have been zero (although possibly with large variances).

In panel a, the central bank's judgment in quarter, z^0, changes from that in previous quarters, such that the central bank's expected inflation deviation equals 1 percentage point for quarter 6, whereas it is still zero in all other quarters; the expected output-gap deviation is still zero for all quarters.[33] Again, behind these means may be a distribution Φ_0 corresponding to considerable uncertainty. The expected inflation deviation, denoted z_π, is marked by circles with no connecting line. The panel shows the optimal policy projection in quarter 0, (π^0, y^0, i^0, r^0), of inflation, π (the

---

[33] In terms of the modeling of the deviation as a mean-average process in section 2.2, panel a shows the impulse respons to ε^0.

Figure 5.1: Monetary policy with and without judgment: Backward-looking model

dashed line); the output gap, y (the dashed-dotted line); the instrument rate, i (the solid line); and the short real interest rate, r (the dotted line).[34]

Panel a has two main interpretations. The first interpretation is that the panel just shows the judgment z^0 and the optimal policy projection $(\hat{\pi}^0, \hat{y}^0, \hat{\imath}^0, \hat{r}^0)$ in quarter $t = 0$ for the future quarters $\tau \geq 1$ and thereby the realization of z_0, $\hat{\pi}_0$, $\hat{y}_0$, $\hat{\imath}_0$, and $\hat{r}_0$ in quarter 0. Conditional on the actual realization of π_1 and y_1 (in turn depending on the realization of $z_{\pi 1}$ and z_{y1}) and the realization of z^1 in quarter $t = 1$, a new optimal policy projection $(\hat{\pi}^1, \hat{y}^1, \hat{\imath}^0, \hat{r}^0)$ has to be plotted in quarter 1 for future quarters $\tau \geq 2$), and so forth for $t = 2, 3, \ldots$

The second interpretation is that the panel shows the probability-zero event that the future realizations of the random deviation z_t for $t \geq 1$ are exactly equal to the judgment z^0 in quarter 0. That is, the realizations of $z_{\pi t}$ for $t \geq 1$ are zero, except in quarter 6 when it is 1 percentage point, and the realizations of z_{yt} for $t \geq 1$ are all zero. In this interpretation, the panel also shows the optimal policy projection $(\hat{\pi}^t, \hat{y}^t, \hat{\imath}^t, \hat{r}^t)$ for each quarter $t \geq 1$. These optimal policy projections are

---

[34] The short real interest rate is defined as $r_t \equiv i_t - \pi_{t+1|t}$.

then simply the continuation of the optimal policy projection of the previous quarter. Furthermore, the actual future realization of inflation, the output gap, the instrument rate, and the real interest rate are then equal to the previous optimal policy projections.

Thus, in the first interpretation, panel a just shows a particular realization of the judgment z^0 and the corresponding optimal policy projection $(\hat{\pi}^0, \hat{y}^0, \hat{\imath}^0, \hat{r}^0)$. In the second interpretation, panel a in addition shows a time series of a particular realization of the future deviation—namely the realization that is exactly equal to the judgment in quarter 0—as well as the resulting realizations over time of inflation, the output gap, the instrument rate, and the real interest rate. Clearly, the probability of the future realizations of the deviation being exactly equal to the previous judgment is generally zero.

Panel a shows that, when the central bank expects a 1 percentage-point inflation deviation in quarter 6, it chooses an optimal instrument-rate projection such that the instrument rate is raised to about 1 percentage point during the first few quarters and then gradually lowered back to its steady-state level. As a result, the projected output gap gradually falls to about -0.5 percentage in quarter 7 and then very gradually rises back towards zero. The inflation projection shows inflation falling slightly before it is hit by the inflation-deviation in quarter 6, then rising to almost 1 percentage point, and finally falling back towards its steady-state level after quarter 6. Thus, the optimal policy projection is a clear example of preemptive monetary policy: The instrument rate is raised and the output gap is reduced well before the expected inflation-deviation shock, so as to efficiently control inflation and bring it back to target after the shock. The optimal policy projection in quarter 0 results in an intertemporal loss of 2.1 units.[35] In the second interpretation, when panel a shows the actual realization of the deviation, the accumulated realized loss over time, discounted to quarter 0, is also the same 2.1 units (since $\delta = 1$, the discounting does not affect the loss).

Panel b shows the time series of the variables for the same particular realization of the future deviations when the inflation deviation equals 1 percentage point in quarter 6 and is zero elsewhere. However, in this panel, the central bank in each quarter disregards judgment, while still responding optimally to the predetermined variables. That is, the central bank responds in the same way to the predetermined variables as for the optimal policy, but it does not respond to any expected future deviation. It behaves as if it believes that the deviation is a serially uncorrelated zero-mean process,

---

[35] Given how the target variables are measured, with the loss function (5.3) and $\delta = 1$, an expected difference of inflation from target of one (two) annualized percentage point(s) for a single quarter gives rise to an intertemporal loss of one (four) units.

so its expected future deviations are zero. This corresponds to a commitment to the instrument rule (4.2) (recall that there is no optimal response to Lagrange multipliers or lagged predetermined variables for the backward-looking model). The central bank then keeps the instrument rate at its steady-state level through quarter 5. Accordingly, inflation and the output gap stay at the steady-state levels through quarter 5. In quarter 6, the inflation shock hits and inflation jumps to 1 percentage point, while the predetermined output gap still stays at zero. In this situation, once the inflation shock has hit, the optimal monetary-policy response is to raise the instrument rate substantially, to more than 1.5 percentage points above the steady-state level during the following few quarters. This reduces the output gap to almost −0.5 percentage points during the next 8–9 quarters. The instrument rate is gradually lowered back to the steady-state level, and inflation and the output gap return to their steady-state levels very slowly. The absence of any preemption requires a larger instrument-rate response when the shock occurs, the output-gap is nevertheless reduced with a considerable lag, and inflation stays above target for a long time. The resulting loss intertemporal is 3.2, 1.1 units higher than when monetary policy relies on judgment.

Panel c is analogous to panel a, except that it shows a situation when the central bank's judgment in quarter 0, z^0, is such that the central bank expects an output-gap deviation of 1 percentage point in quarter 6, whereas no other deviations are expected. The expected output-gap deviation, z_y, is denoted by circle markers with no connecting line. Again, panel c has two interpretations; the first is that the panel just shows the judgment and optimal policy projection in quarter 0; the second is that is also shows the time series of the variables if the future realizations of the output-gap deviation are exactly equal to the judgment in quarter 0. For this expected output-gap deviation, the optimal policy projection shows a substantial increase in the instrument rate to almost 2 percentage points above the steady-state level in quarter 3–5 and then a rather quick reduction back to the steady-state level. As a result, the output-gap projection shows output falling by almost −0.5 percentage points before the expected output-gap deviation hits, after which it rises to less than 0.5 percentage points and then relatively quickly comes back to the steady-level. The resulting movements in the inflation projection are small. A modest loss of 0.51 units results from this preemptive optimal policy projection.

Panel d is analogous to panel b, except that it shows the time series of the variables for the particular realization of the output-gap deviation when it equals 1 percentage-point in quarter 6 and is zero in other quarter. The central bank disregards judgment and only responds to current and lagged inflation and the output gap (although again optimally so, according to the instrument rule

(4.2)). Then the central bank keeps the instrument rate at its steady-state level until the output-gap shock hits in quarter 6. Once the shock has hit, it is optimal to raise the instrument rate even more, to more than 2 percentage points for a few quarters, before it is lowered back to the steady-state level. The output gap stays up around 1 percentage point for several quarters. This causes inflation to rise and only very slowly return to target. The output gap has to undershoot the steady-state level significantly in order to bring inflation back. Clearly, inflation and the output gap deviate substantially more than when the central bank uses its judgment. The resulting intertemporal loss is 3.1, 2.6 units higher than the loss for the optimal policy projection with judgment.

This example shows a substantial difference between monetary policy with and without judgment, with substantial differences in the development of the target variables and corresponding intertemporal losses.

5.2. Forward-looking model

The forward-looking New Keynesian model of Lindé [13] has two equations. I use the following parameter estimates,

$$\pi_t = 0.457\,\pi_{t+1|t} + (1-0.457)\pi_{t-1} + 0.048y_t + z_{\pi t},$$

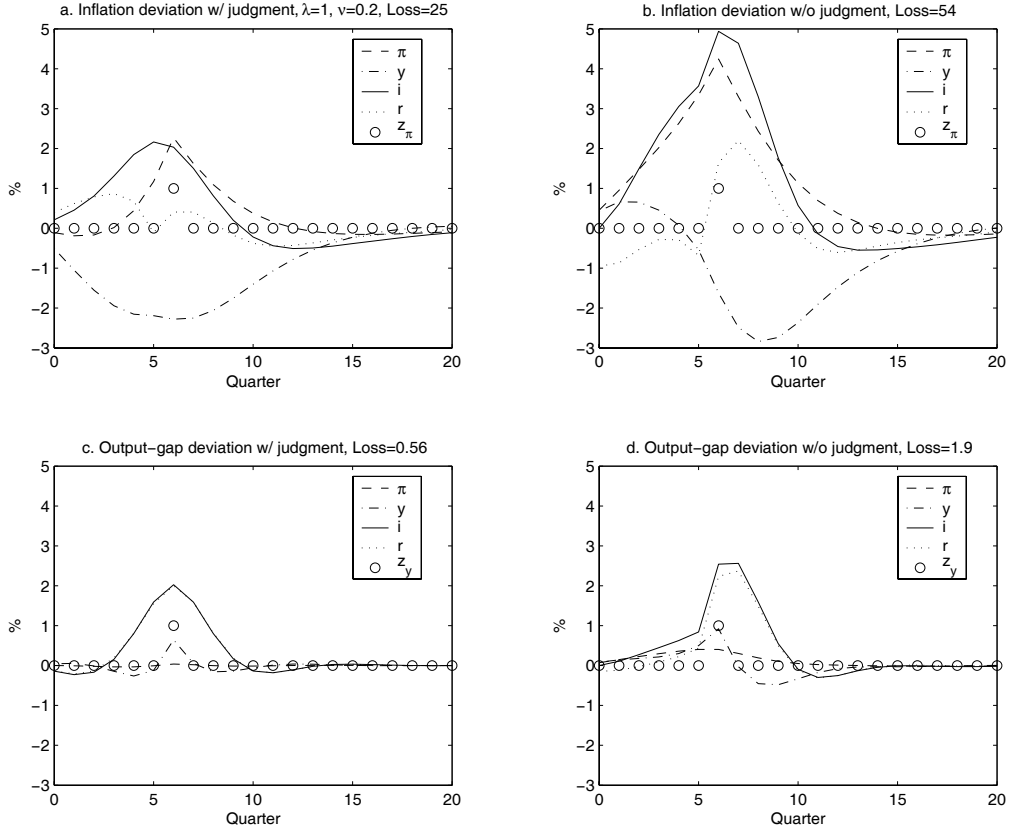$$y_t = 0.425\,y_{t+1|t} + (1-0.425)y_{t-1} - 0.156(i_t - \pi_{t+1|t}) + z_{yt}.$$

The variables are measured as for the backward-looking model. The predetermined variables are $(\pi_{t-1}, y_{t-1}, i_{t-1}, z_{\pi t}, z_{yt})$ (the lagged instrument rate enters because it enters into the loss function, and the two deviations are entered in order to write the model on the form (2.1), see section 2). The forward-looking variables are (π_t, y_t). See appendix I for details. The loss function and its parameters used in the experiment below are the same as for the backward-looking model.[36]

Figure 5.2, panel a-d, shows the same experiments as figure 5.1, but for the forward-looking model. Thus, before quarter 0, the variables are equal to their steady-state levels, zero, and the central bank does not expect any future inflation and output-gap deviations.

In panel a, in quarter 0, while the inflation and the output-gap deviations in that quarter are still zero, the central bank's judgment, z^0, changes, such that the central bank expects an inflation deviation equal to 1 percentage point in quarter 6, whereas it still expects zero inflation deviations for all other quarters and zero output-gap deviations for all quarters. Again, behind these

---

[36] I find it very unrealistic to consider inflation and output in the current quarter as forward-looking variables. I believe it makes more sense to have current inflation and the output gap predetermined, and to have one-quarter-ahead inflation, output-gap, and instrument-rate plans be determined by the model above. Such a variant of the newkeynesian model is used in Svensson and Woodford [30] and Svensson [25].

Figure 5.2: Monetary policy with and without judgment: Forward-looking model

a. Inflation deviation w/ judgment, $\lambda=1$, $\nu=0.2$, Loss=25

b. Inflation deviation w/o judgment, Loss=54

c. Output–gap deviation w/ judgment, Loss=0.56

d. Output–gap deviation w/o judgment, Loss=1.9

expected deviation could be a probability distribution Φ_0 corresponding to substantial uncertainty. As above, panel a has two interpretations. In the first interpretation, it just shows the judgment z^0 and the optimal policy projection $(\hat{\pi}^0, \hat{y}^0, \hat{\imath}^0, \hat{r}^0)$. In the second interpretation, it shows the time series of inflation, the output-gap, the instrument rate, and the real interest rate, for the particular realizations of the future deviations that are exactly equal to the central-bank's judgment in quarter 0. In this interpretation, I also assume that the private sector has sufficient information—cf. the discussion in section 2.1—to form expectations consistent with the optimal policy projection.

The optimal policy projection in panel a shows that the central bank plans to raise the instrument rate to about 2 percentage points above the steady-state level in the quarters before and including the time of the inflation shock. This makes the output-gap projection fall to more than -2 percentage points at the time of the expected inflation deviation. The inflation projection rises before and up to the expected inflation deviation, because private-sector expectations are forward-looking and consistent with the optimal inflation projection. After the expected inflation deviation, the instrument rate, the output gap, and inflation are projected to return to their steady-state lev-
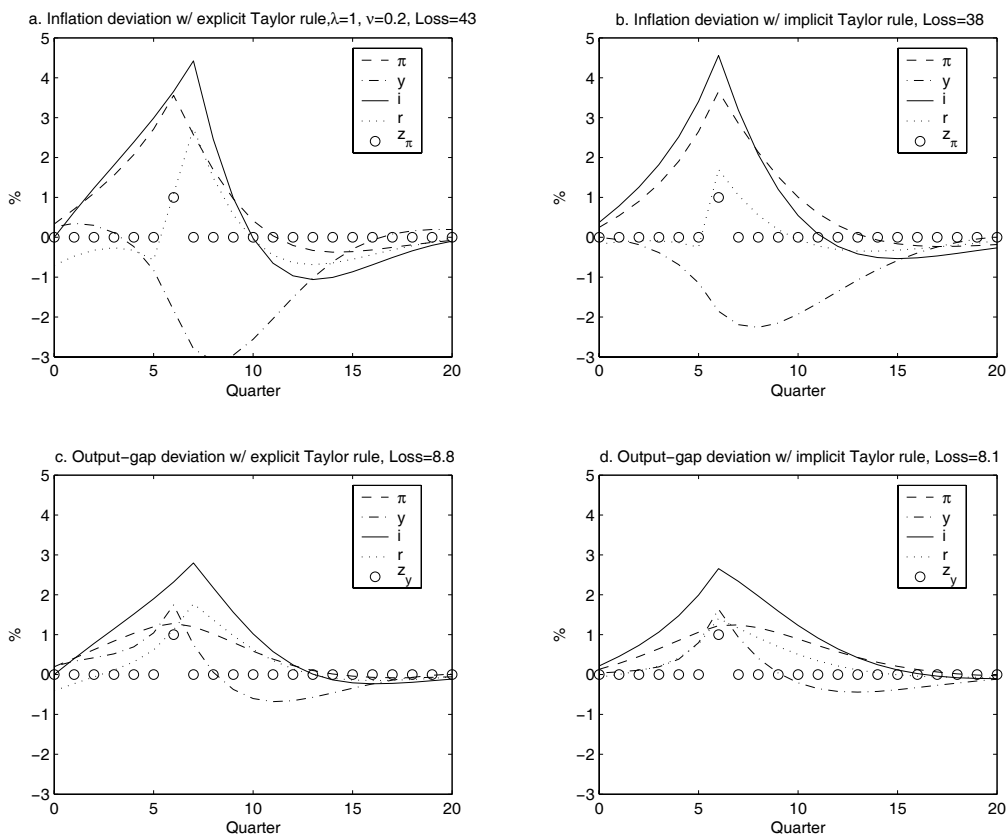
els. Again, there is a considerable amount of preemption in the optimal policy with judgment, with a projected positive real interest rate and negative output gap before the expected inflation deviation. A substantial intertemporal loss of 25 units results from the optimal policy projection.

Panel b shows the realizations over time of these variables when the realizations of the inflation deviation is equal to 1 percentage point in quarter 6 and zero in other quarters and the central bank in each quarter disregards judgment while still responding optimally to current and lagged inflation and output gap. In this case, the central bank is assumed to respond optimally to both the predetermined variables *and* the lagged predetermined variables, as if the central bank had committed itself to the optimal policy under commitment while ignoring its judgment. Hence, the central bank behaves according to instrument rule (4.3) and responds optimally to the current deviation but expects zero future deviations. However, the private sector is assumed to have rational expectations of the future inflation shock. These expectations will increase inflation to more than 4 percentage points at the time of the inflation shock. The central bank's optimal response to current and predetermined variables induces it to raise the instrument rate in line with inflation, but it is nevertheless behind the curve in the sense that the real interest rate becomes negative and the output gap becomes positive in the first few quarters. The central bank's response eventually leads to a high positive real interest rate, a negative output gap, and a fall in inflation. In comparison with panel a, inflation rises earlier and more, and the output gap falls later, than under the optimal monetary policy with judgment. The intertemporal loss is 54, a substantial increase of 29 units above the loss for monetary policy with judgment.

Panel c shows the situation where the central bank's judgment in quarter 0 is such that it expects an output-gap deviation of 1 percentage point in quarter 6 and otherwise zero deviations. The optimal policy projection, taking this judgment into account, is to raise the instrument rate before the expected output-gap deviation, which moderates the expected impact on the output gap. The inflation projection remains very flat, and the projections of the real interest rate and the instrument rate are almost identical. The resulting intertemporal loss is small, 0.56.

Panel d shows the realizations over time of the variables in the situation where the realization of the output-gap deviation is 1 percentage point in quarter 6 and zero in other quarters and the central bank disregards judgment and only responds to current and lagged predetermined variables, although again optimally so, corresponding to the instrument rule (4.3). In comparison with the second interpretation of panel c, when the panel shows the actual realization of the variables for the same realization of the deviations, the central bank ends up raising the instrument rate later

Figure 5.3: Monetary policy with Taylor rules: Forward-looking model

and more, and there is more movement in both the output gap and inflation. The intertemporal loss is 1.9, 1.3 units above the loss for optimal policy under judgment.

Again, there are substantial differences between monetary policy with and without judgment and corresponding intertemporal losses.

5.2.1. Taylor rules

I also examine two variants of Taylor rules for the forward-looking model, an explicit instrument rule for which the instrument rate responds to lagged inflation and the output gap,

$$i_t = 1.5\,\pi_{t-1} + 0.5\,y_{t-1},$$

and an implicit instrument rule for which the instrument rate responds to the forward-looking current inflation and output,

$$i_t = 1.5\,\pi_t + 0.5\,y_t.$$

As noted in section 4, the implementation of an implicit instrument rule is somewhat complex. I disregard these complications here, and just assume that it is somehow implemented. Figure 5.3 shows the realizations over time of the variables when the central bank implementing the two Taylor rules for the two cases of either an inflation deviation or an output-gap deviation only in quarter 6.

Panels a and b show the result of the explicit and implicit Taylor rule, respectively, when there is an inflation deviation in quarter 6 and the private sector has rational expectations of that deviation. The resulting intertemporal losses are substantial, 43 and 38, respectively—18 and 13 units, respectively, above the loss for optimal monetary policy with judgment, 25. Interestingly, the intertemporal loss with either of the two Taylor rules is less than the policy without judgment that responds optimally to current and lagged predetermined variables, panel b in figure 5.2, which has an intertemporal loss of 54. One possible interpretation of this is that history dependence in the form of responding to the Lagrange multipliers is not always advantageous, when these multipliers do not take into account the expected future deviations. The loss for the implicit Taylor rule is lower than for the explicit one. One interpretation is that the implicit Taylor rule takes private-sector expectations better into account, and therefore indirectly takes the expected future deviation better into account.

Panels c and d show the result of the two Taylor rules when there is an output-gap deviation in quarter 6. Here, the intertemporal loss is substantially higher than the small loss for monetary policy without judgment in panel d of figure 5.2. In this case, the optimal response to current and lagged predetermined variables does much better than the two Taylor rules.

I conclude that the two Taylor rules perform considerably worse than the optimal policy with judgment, especially when there are expected future output-gap deviations.

6. Conclusions

The decision process of modern monetary policy that can be called "forecast targeting"—finding a projection of the current and future instrument rate such that the projection paths of the target variables "look good" relative to the central bank's objectives—is formalized in this paper as a technique that provides projections of the instrument rate and the target variables that minimize an intertemporal loss function. The paper shows how this technique can easily incorporate central-bank judgment, a necessary ingredient in modern monetary policy. In two empirical models of the U.S. economy, a few examples are shown in which forecast targeting that incorporates judgment provides significantly better monetary-policy performance than a policy that follows an instrument

rule and disregards judgment. The paper shows how the policy problem, normally treated as an infinite-horizon problem, can be reformulated as a convenient finite-horizon decision problem, which is either an exact or a very close approximation to the infinite-horizon problem. This approximation makes the policy problem much easier to handle numerically. The paper also shows how the time-consistency problem can be easily managed and the resulting projections made to be optimal under commitment in a timeless perspective. In particular, the paper shows that it is not necessary to be explicit about the underlying complex reduced-form reaction function of monetary policy. The policymakers only need to ponder the projections of the target variables and the instrument rate under alternative assumptions, and these projections can be presented as graphs.

Several of the ideas and techniques presented here are already applied to various extents by different central banks. I hope the presentation here will be useful for attempts to apply them more extensively and systematically.

If policymakers hesitate to make the parameters of their loss function explicit (for instance, the weight on output-gap stabilization relative to inflation stabilization), the techniques presented here can still be very useful. For instance, the policymakers can ask the staff to provide optimal policy projections of the target variables for a range of loss-function parameters. These projections then provide one way to illustrate the available tradeoffs among the target variables, the set of feasible projections of the target variables from which the policy makers should choose their optimal policy projection.

The framework used here is one where mean projections are sufficient for optimal decisions, what can be called *mean* forecast targeting, which is sufficient under the assumptions that result in certainty equivalence. If these assumptions are not satisfied, the principal ideas in this paper can be extended to a situation when the projections are probability *distributions* rather than means, and the intertemporal losses can be computed by numerical integration over those distributions. This I have previously called *distribution* forecast targeting (Svensson [23]). The details in such an undertaking remain to be completed, and the practical differences between mean and distribution forecast targeting remain to be clarified. Svensson and Williams [32] examines distribution forecast targeting in a situation when where genuine model uncertainty implies that certainty equivalence does not hold.

References

[1] Amato, Jefferey D., Stephen Morris, and Hyun Song Shin (2002), "Communication and Monetary Policy," *Oxford Review of Economic Policy* 18, 495–503.

[2] Backus, David, and John Driffill (1986), "The Consistency of Optimal Policy in Stochastic Rational Expectations Models," CEPR Discussion Paper No. 124.

[3] Bernanke, Ben S. (2004), "The Logic of Monetary Policy," speech in Washingtond, DC, December 2, Federal Reserve Board, www.federalreserve.gov.

[4] Brash, Donald T. (2000), "Making Monetary Policy: A Look Behind the Curtains," speech in Christchurch, January 26, Reserve Bank of New Zealand, www.rbnz.govt.nz.

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

[6] Faust, Jon, and Dale W. Henderson (2004), "Is Inflation Targeting Best-Practice Monetary Policy?" *Federal Bank of St. Louis Economic Review* 86(4), 117–143.

[7] Geraats, Petra M. (2002), "Central Bank Transparency," *Economic Journal* 112, F532–F556.

[8] Giannoni, Marc P., and Michael Woodford (2002), "Optimal Interest-Rate Rules: I. General Theory," NBER Working Paper No. 9419.

[9] Greenspan, Alan (2004), "Risk and Uncertainty in Monetary Policy," speech in San Diego, CA, January 3, www.federalreserve.gov.

[10] Hansen, Lars Peter, and Thomas J. Sargent (1998), "Alternative Representation of Discounted Robust Linear Quadratic Control," working paper.

[11] Judd, Kenneth L. (1998), *Numerical Methods in Economics*, MIT Press, Cambridge, MA.

[12] Kozicki, Sharon (1999), "How Useful Are Taylor Rules for Monetary Policy?" *Federal Reserve Bank of Kansas City Economic Review*, Second Quarter 1999, 5–33.

[13] Lindé, Jesper (2002), "Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach," Sveriges Riksbank Working Paper No. 129.

[14] Marcet, Albert, and Ramon Marimon (1998), "Recursive Contracts," working paper, www.econ.upf.es.

[15] McCallum, Bennett T., and Edward Nelson (2004), "Targeting Rules vs. Instrument Rules for Monetary Policy," working paper.

[16] Morris, Stephen, and Hyun Song Shin (2002), "The Social Value of Public Information," *American Economic Review* 92, 1521–1534.

[17] Reifschneider, David L., David J. Stockton, and David W. Wilcox (1997), "Econometric Models and the Monetary Policy Process," *Carnegie-Rochester Conference Series on Public Policy* 47, 1–37.

[18] Rudebusch, Glenn D., and Lars E.O. Svensson (1999), "Policy Rules for Inflation Targeting,", in John B. Taylor (ed.), *Monetary Policy Rules*, University of Chicago Press.

[19] Sims, Christopher A. (2002), "The Role of Models and Probabilites in the Monetary Policy Process," *Brookings Papers on Economic Activity* 2:2002, 1–62.

[20] Söderlind, Paul (1999), "Solution and Estimation of RE Macromodels with Optimal Policy," *European Economic Review* 43, 813–823. Related software available at home.tiscalinet.ch/paulsoderlind.

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

[22] Svensson, Lars E.O. (2001a), *Independent Review of the Operation of Monetary Policy in New Zealand: Report to the Minister of Finance*, available at www.princeton.edu/∼svensson.

[23] Svensson, Lars E.O. (2001b), "Price Stability as a Target for Monetary Policy: Defining and Maintaining Price Stability," in Deutsche Bundesbank, ed., *The Monetary Transmission Process: Recent Developments and Lessons for Europe,* Palgrave, New York, 60–102.

[24] Svensson, Lars E.O. (2001c), "Requiem for Forecast-Based Instrument Rules," working paper, www.princeton.edu/∼svensson.

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

[26] Svensson, Lars E.O. (2004), "Targeting Rules vs. Instrument Rules for Monetary Policy: What Is Wrong with McCallum and Nelson?" working paper, www.princeton.edu/∼svensson.

[27] Svensson, Lars E.O. (2005), "Social Value of Public Information: Morris and Shin (2002) Is Actually Pro Transparency, Not Con," working paper, www.princeton.edu/∼svensson.

[28] Svensson, Lars E.O., Kjetil Houg, Haakon Solheim, and Erling Steigum (2002), "An Independent Review of Monetary Policy and Institutions in Norway," Norges Bank Watch 2002, Centre for Monetary Economics, Norwegian School of Management BI, www.princeton.edu/∼svensson.

[29] Svensson, Lars E.O., and Michael Woodford (2003), "Indicator Variables for Optimal Policy," *Journal of Monetary Economics* 50, 691–720.

[30] Svensson, Lars E.O., and Michael Woodford (2005), "Implementing Optimal Policy through Inflation-Forecast Targeting," in Bernanke, Ben S., and Michael Woodford, eds., *Inflation Targeting*, University of Chicago Press, forthoming, www.princeton.edu/∼svensson.

[31] Svensson, Lars E.O., and Robert J. Tetlow (2005), "Optimal Policy Projections," working paper, www.princeton.edu/∼svensson.

[32] Svensson, Lars E.O., and Noah Williams (2005), "Monetary Policy with Model Uncertainty: Distribution Forecast Targeting," in progress.

[33] Taylor, John B., ed. (1999), *Monetary Policy Rules*, University of Chicago Press.

[34] Walsh, Carl E. (2004), "Robustly Optimal Instrument Rules and Robust Control: An Equivalence Result," *Journal of Money, Credit, and Banking*, forthcoming.

[35] Woodford, Michael (2003a), "Inflation Targeting and Optimal Monetary Policy," presented at the 28th Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 16-17, 2003, www.stlouisfed.org.

[36] Woodford, Michael (2003b), *Interest Rates and Prices*, Princeton University Press.

<div align="center">

Appendix to

Monetary Policy with Judgment: Forecast Targeting

Lars E.O. Svensson

February 2005

</div>

Contents

A. Optimal policy under commitment with the deviation being an arbitrary stochastic process

Let the model equations for $t \geq 0$ be (2.1). A common special case is when the matrix $C = I$, but in general C need not be invertible. This system can be written

$$
\tilde{C} \begin{bmatrix} X_{t+1} \\ \mathrm{E}_t x_{t+1} \\ \mathrm{E}_t i_{t+1} \end{bmatrix} = \tilde{A} \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} + \begin{bmatrix} z_{t+1} \\ 0 \end{bmatrix},
\tag{A.1}
$$

where $\mathrm{E}_t q_{t+\tau} \equiv \int q_{t+\tau} d\Phi_t(\zeta^t)$ for any variable $q_{t+\tau}$ $(\tau \geq 0)$, the matrices $\tilde{A}$ and $\tilde{C}$ are of dimension $(n_X + n_x) \times (n_X + n_x + n_i)$ and given by

$$
\tilde{A} \equiv \begin{bmatrix} A & B \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \end{bmatrix}, \quad \tilde{C} \equiv \begin{bmatrix} I & 0 & 0 \\ 0 & C & 0 \end{bmatrix}.
$$

where A and B are partitioned according to (2.3).

The target variables are defined by (2.5). The intertemporal loss function in period 0 is

$$
\mathrm{E}_0 \sum_{t=0}^{\infty} \delta^t L_t,
$$

<div align="center">1</div>

where the period loss function, (2.7), can be written as

$$L_t = \frac{1}{2} \begin{bmatrix} X_t' & x_t' & i_t' \end{bmatrix} D'WD \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}.$$

Consider minimizing this intertemporal loss function under once-and-for-all commitment in period $t = 0$, for given $X_0 = \bar{X}_0$. For convergence, when the variance of z_{t+1} is nonzero, I need $0 < \delta < 1$.

Variants of this problem are solved in Backus and Driffill [2], Currie and Levine [5], and Söderlind [20], when the deviation is an iid shock. The focus here is on the case when the deviation is an arbitrary stochastic process.

Construct the Lagrangian,

$$
\begin{aligned}
\mathcal{L}_0 &= \mathrm{E}_0 \sum_{t=0}^{\infty} \delta^t \left\{ L_t + \begin{bmatrix} \xi_{t+1}' & \Xi_t' \end{bmatrix} \left(\tilde{C} \begin{bmatrix} X_{t+1} \\ \mathrm{E}_t x_{t+1} \\ \mathrm{E}_t i_{t+1} \end{bmatrix} - \tilde{A} \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} - \begin{bmatrix} z_{t+1} \\ 0 \end{bmatrix} \right) \right\} \\
&\quad + \xi_0'(X_0 - \bar{X}_0)/\delta \\
&= \mathrm{E}_0 \sum_{t=0}^{\infty} \delta^t \left\{ L_t + \begin{bmatrix} \xi_{t+1}' & \Xi_t' \end{bmatrix} \left(\tilde{C} \begin{bmatrix} X_{t+1} \\ x_{t+1} \\ i_{t+1} \end{bmatrix} - \tilde{A} \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} - \begin{bmatrix} z_{t+1} \\ 0 \end{bmatrix} \right) \right\} \\
&\quad + \xi_0'(X_0 - \bar{X}_0)/\delta,
\end{aligned}
$$

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 (A.1). The law of iterated expectations has been used in the second equality, $\mathrm{E}_0\mathrm{E}_t = \mathrm{E}_0$ for $t \geq 0$. Note that Ξ_t is dated to emphasize that it depends on information available in period t.

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

$$\begin{bmatrix} X_t' & x_t' & i_t' \end{bmatrix} D'WD + \begin{bmatrix} \xi_t' & \Xi_{t-1}' \end{bmatrix} \frac{1}{\delta} \tilde{C} - \begin{bmatrix} \mathrm{E}_t \xi_{t+1}' & \Xi_t' \end{bmatrix} \tilde{A} = 0. \tag{A.2}$$

For $t = 0$, the first-order condition with respect to X_0, x_0, and i_0 can be written

$$\begin{bmatrix} X_t' & x_t' & i_t' \end{bmatrix} D'WD + \begin{bmatrix} \xi_t' & 0 \end{bmatrix} \frac{1}{\delta} \tilde{C} - \begin{bmatrix} \mathrm{E}_t \xi_{t+1}' & \Xi_t' \end{bmatrix} \tilde{A} = 0, \tag{A.3}$$

where $X_0 = \bar{X}_0$. In comparison with (A.2), a vector of zeros enters in place of Ξ_{-1}, since there is no constraint corresponding to the lower block of (A.1) for $t = -1$. By including a fictitious vector of Lagrange multipliers, Ξ_{-1}, equal to zero,

$$\Xi_{-1} = 0, \tag{A.4}$$

in (A.3), I can write the first-order conditions more compactly as (A.2) for $t \geq 0$ and (A.4).

2

The system of difference equations (A.2) has $n_X + n_x + n_i$ equations. The first n_X equations can be associated with the Lagrange multipliers ξ_t. Indeed, $-\xi_t/\delta$ 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: Lagrange multipliers for predetermined variables are always forward-looking, whereas the Lagrange multipliers for the (equations for the) forward-looking variables are predetermined. The middle n_x equations are associated with the Lagrange multipliers Ξ_t. Indeed, $\Xi_t A_{22}$ can be interpreted as the total marginal losses in period t of the forward-looking variables, x_t. Also, $\Xi_t C$ can be seen as the marginal loss in period t of expectations $\mathrm{E}_t x_{t+1}$ of the forward-looking variables. 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 of $D'WD$ is of full rank, the instruments can be solved in terms of the other variables and eliminated from (A.2), 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

$$\tilde{A}' \begin{bmatrix} \mathrm{E}_t \xi_{t+1} \\ \Xi_t \end{bmatrix} = D'WD \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} + \frac{1}{\delta}\tilde{C}' \begin{bmatrix} \xi_t \\ \Xi_{t-1} \end{bmatrix}. \tag{A.5}$$

They can be combined with the model equations (A.1) to get a system of $2(n_X + n_x) + n_i$ difference equations for $t \geq 0$,

$$\begin{bmatrix} \tilde{C} & 0 \\ 0 & \tilde{A}' \end{bmatrix} \begin{bmatrix} X_{t+1} \\ \mathrm{E}_t x_{t+1} \\ \mathrm{E}_t i_{t+1} \\ \mathrm{E}_t \xi_{t+1} \\ \Xi_t \end{bmatrix} = \begin{bmatrix} \tilde{A} & 0 \\ D'WD & \frac{1}{\delta}\tilde{C}' \end{bmatrix} \begin{bmatrix} X_t \\ x_t \\ i_t \\ \xi_t \\ \Xi_{t-1} \end{bmatrix} + \begin{bmatrix} z_{t+1} \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{A.6}$$

Here, X_t and Ξ_t are predetermined variables, and x_t, i_t, and ξ_t are non-predetermined variables.

This can be rearranged as the system

$$\mathcal{C} \begin{bmatrix} y_{1,t+1} \\ \mathrm{E}_t y_{2,t+1} \end{bmatrix} = \mathcal{M} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} z_{t+1} \\ 0 \\ 0 \end{bmatrix},$$

where

$$\mathcal{C} \equiv \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & 0 & C & 0 & 0 \\ 0 & A'_{21} & 0 & 0 & A'_{11} \\ 0 & A'_{22} & 0 & 0 & A'_{12} \\ 0 & B'_2 & 0 & 0 & B'_1 \end{bmatrix}, \tag{A.7}$$

$$y_{1t} \equiv \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix}, \qquad y_{2t} \equiv \begin{bmatrix} x_t \\ i_t \\ \xi_t \end{bmatrix}.$$

3

Thus, y_{1t} is a vector of $m_1 \equiv n_X + n_x$ predetermined variables, and y_{2t} is a vector of $m_2 \equiv n_x + n_i + n_X$ non-predetermined variables.

Under suitable assumptions (see appendix B), such a system has a unique solution, which can be written

$$y_{2t} = F_1 y_{1t} + Z_t \tag{A.8}$$

$$y_{1,t+1} = M_1 y_{1t} + N\mathrm{E}_t Z_{t+1} + P\mathrm{E}_t z_{t+1} + \begin{bmatrix} z_{t+1} - \mathrm{E}_t z_{t+1} \\ 0 \end{bmatrix}, \tag{A.9}$$

where Z_t is an m_2-dimensional stochastic process given by

$$Z_t \equiv \sum_{\tau=0}^{\infty} R_\tau \mathrm{E}_t z_{t+1+\tau} \equiv Rz^t, \tag{A.10}$$

where I can interpret R as a linear operator on $z^t \equiv \mathrm{E}_t(z'_{t+1}, z'_{t+2}, ...)'$.

In terms of the original variables, the solution for $t \geq 0$, given X_0 and $\Xi_{-1} = 0$, can be written

$$
\begin{bmatrix} x_t \\ i_t \\ \xi_t \end{bmatrix} = F_1 \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix} + Rz^t
$$

$$
\equiv F \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1} \end{bmatrix}, \tag{A.11}
$$

$$
\begin{bmatrix} X_{t+1} \\ \Xi_t \end{bmatrix} = M_1 \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix} + NR\mathrm{E}_t z^{t+1} + P\mathrm{E}_t z_{t+1} + \begin{bmatrix} z_{t+1} - \mathrm{E}_t z_{t+1} \\ 0 \end{bmatrix}
$$

$$
\equiv M \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1} \end{bmatrix} + \begin{bmatrix} z_{t+1} - \mathrm{E}_t z_{t+1} \\ 0 \end{bmatrix}, \tag{A.12}
$$

where F and M are linear operators. The details of the solution are derived in appendix B. The matrices F_1, M_1, N, P, and $\{R_\tau\}_{\tau=0}^{\infty}$—and thereby the linear operators M and F—depend on A, B, C, D, W, and δ, but are independent of the second and higher moments of the exogenous stochastic process $\{z_t\}_{t=1}^{\infty}$. This demonstrates the certainty equivalence of the commitment solution.[37]

If the commitment is once and for all and starts in period 0, $\Xi_{-1} = 0$. Commitment in a timeless perspective can be seen as corresponding to a situation where the lower block of (A.12) is restricted to apply also for previous periods. Then, Ξ_{t-1} is determined by

$$
\Xi_{t-1} = M_{121} X_{t-1} + M_{122} \Xi_{t-2} + N_2 \mathrm{E}_{t-1} Z_t + P_2 \mathrm{E}_{t-1} z_t
$$

$$
= \sum_{\tau=0}^{\infty} M_{122}{}^\tau (M_{121} X_{t-1-\tau} + N_2 \mathrm{E}_{t-1-\tau} Z_{t-\tau} + P_2 \mathrm{E}_{t-1-\tau} z_{t-\tau}),
$$

[37] The middle block of (A.11) is the optimal *explicit instrument rule* for this problem, the instrument written as a function of predetermined and exogenous variables. Eliminating the Lagrange multipliers from (A.2) results in the optimal *targeting rule* for this problem, a consolidated optimal first-order condition for the target variables. See Svensson [25] on instrument and targeting rules, as well as the lecture notes Svensson [A7].

where M_1, N, and P are partitioned conformably with X_t and Ξ_{t-1}.

Alternatively, the commitment in a timeless perspective can be generated as optimization under commitment or discretion with a term added to the intertemporal loss function in period 0,

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \delta^t L_t + \Xi_{-1} \frac{1}{\delta} C x_0,$$

where Ξ_{-1} is the Lagrange multiplier for the block of forward-looking equations from the optimization in period -1 (see Svensson and Woodford [30] and Svensson [25]).

In the standard case, when z_t is a vector of iid zero-mean shocks, I have $\mathrm{E}_t z_{t+1} \equiv 0$, $Z_t \equiv \mathrm{E}_t Z_{t+1} \equiv 0$, and $z^t \equiv 0$. Thus, the terms involving Z_t in (A.11) and (A.12) vanish.[38] Consequently, the effect of z_t being an arbitrary exogenous stochastic process shows up only in the addition of the terms involving Z_t and the corresponding matrices N, P, and $\{R_\tau\}_{\tau=0}^{\infty}$. Then, I can set $M \equiv M_1$ and $F \equiv F_1$, and

$$y_{1,t+1} = M y_{1t} + z_{t+1}.$$

Let Σ denote the variance-covariance matrix of the iid shocks z_{t+1}. Define the matrices $\bar{D}$ and $\bar{W}$ according to

$$Y_t = D \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} = D \begin{bmatrix} I & 0 \\ F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} y_{1t} \equiv \bar{D} y_{1t},$$

$$L_t = \frac{1}{2} Y_t' W Y_t = \frac{1}{2} y_{1t}' \bar{D}' W \bar{D} y_{1t} \equiv \frac{1}{2} y_{1t}' \bar{W} y_{1t},$$

where $\bar{W}$ is symmetric and positive semidefinite. Then twice the minimum loss in period t will satisfy

$$
\begin{aligned}
y_{1t}' V y_{1t} + w \;=\;& \mathrm{E}_t \sum_{\tau=0}^{\infty} \delta^\tau y_{1,t+\tau}' \bar{W} y_{1,t+\tau} \\
=\;& y_{1t}' \bar{W} y_{1t} + \mathrm{E}_t \sum_{\tau=1}^{\infty} \delta^\tau y_{1,t+\tau}' \bar{W} y_{1,t+\tau} \\
=\;& y_{1t}' \bar{W} y_{1t} + \delta \mathrm{E}_t \mathrm{E}_{t+1} \sum_{\tau=0}^{\infty} \delta^\tau y_{1,t+1+\tau}' \bar{W} y_{1,t+1+\tau} \\
=\;& y_{1t}' \bar{W} y_{1t} + \delta \mathrm{E}_t (y_{1,t+1}' V y_{1,t+1} + w) \\
=\;& y_{1t}' \bar{W} y_{1t} + \delta (y_{1t}' M' V M y_{1t} + \mathrm{E}_t z_{1,t+1}' V z_{1,t+1} + w) \\
=\;& y_{1t}' \bar{W} y_{1t} + \delta y_{1t}' M' V M y_{1t} + \delta \mathrm{trace}(V \Sigma) + \delta w.
\end{aligned}
$$

---

[38] In the case when $\{z_t\}$ is an autoregressive process and can be written $z_{t+1} = \Psi z_t + \varepsilon_{t+1}$, where Ψ is a matrix and ε_t an iid zero-mean process, z_t can simply be included among the predetermined variable.

It follows that

$$w = \frac{\delta}{1 - \delta} \text{trace}(V\Sigma),$$

and that the matrix V satisfies the Lyapunov equation

$$V = \bar{W} + \delta M'VM. \tag{A.13}$$

It follows that when $\text{trace}(V\Sigma)$ is nonzero, I must have $\delta < 1$ for the existence of an finite w.

I can use the relations $\text{vec}(A + B) = \text{vec}(A) + \text{vec}(B)$ and $\text{vec}(ABC) = (C' \otimes A)\,\text{vec}(B)$ on (A.13) (where $\text{vec}(A)$ denotes the vector of stacked column vectors of the matrix A, and $\otimes$ denotes the Kronecker product) which results in

$$\begin{aligned} \text{vec}(V) &= \text{vec}(\bar{W}) + \delta \text{vec}(M'VM) \\ &= \text{vec}(\bar{W}) + \delta (M' \otimes M')\,\text{vec}(V). \end{aligned}$$

Solving for $\text{vec}(V)$ gives

$$\text{vec}(V) = \left[I - \delta (M' \otimes M')\right]^{-1} \text{vec}(\bar{W}). \tag{A.14}$$

A.1. No forward-looking variables

If there are no forward-looking variables, so $n_x = 0$, I have

$$\tilde{C} \begin{bmatrix} X_{t+1} \\ \mathrm{E}_t i_{t+1} \end{bmatrix} = \tilde{A} \begin{bmatrix} X_t \\ i_t \end{bmatrix} + z_{t+1}, \tag{A.15}$$

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

$$\tilde{A} \equiv \begin{bmatrix} A & B \end{bmatrix}, \quad \tilde{C} \equiv \begin{bmatrix} I & 0 \end{bmatrix}.$$

The period loss function is

$$L_t = \frac{1}{2} Y_t' W Y_t \equiv \frac{1}{2} \begin{bmatrix} X_t' & i_t' \end{bmatrix} D'WD \begin{bmatrix} X_t \\ i_t \end{bmatrix}.$$

The $n_X + n_i$ first-order conditions can be written

$$\tilde{A}' \mathrm{E}_t \xi_{t+1} = D'WD \begin{bmatrix} X_t \\ i_t \end{bmatrix} + \frac{1}{\delta} \tilde{C}' \xi_t. \tag{A.16}$$

Combined with the model equations, I get a system of $2n_X + n_i$ difference equations for $t \geq 0$,

$$\begin{bmatrix} \tilde{C} & 0 \\ 0 & \tilde{A}' \end{bmatrix} \begin{bmatrix} X_{t+1} \\ \mathrm{E}_t i_{t+1} \\ \mathrm{E}_t \xi_{t+1} \end{bmatrix} = \begin{bmatrix} \tilde{A} & 0 \\ D'WD & \frac{1}{\delta}\tilde{C}' \end{bmatrix} \begin{bmatrix} X_t \\ i_t \\ \xi_t \end{bmatrix} + \begin{bmatrix} z_{t+1} \\ 0 \end{bmatrix}.$$

Here, X_t are predetermined variables, and i_t and ξ_t are non-predetermined variables.

Under suitable assumptions, this system will have a unique solution for $t \geq 0$, given X_0, which can be written

$$
\begin{bmatrix} i_t \\ \xi_t \end{bmatrix} = F_1 X_t + R z^t,
$$
$$
X_{t+1} = M_1 X_t + N_0 R z^t + z_{t+1}.
$$

When there are no forward-looking variables, X_{t+1} is directly determined by X_t, i_t, and z_{t+1} according to (2.1), so M_1 and N_0 are determined by A, B, and F_1 as

$$
M_1 \equiv A + B F_i,
$$
$$
N_0 \equiv [B \ 0],
$$

where

$$
F_1 \equiv \begin{bmatrix} F_i \\ F_\xi \end{bmatrix}
$$

is partitioned conformably with i_t and ξ_t. In comparison with the general solution of (A.9), for the backward-looking case,

$$
N_0 R z^t \equiv N R \mathrm{E}_t z^{t+1} + (P - I) \mathrm{E}_t z_{t+1}.
$$

B. The solution of a system of difference equations with the deviation

In order to understand the term in the solution (A.10) and (A.11) that corresponds to the deviation, consider the system

$$
\mathcal{C} \begin{bmatrix} y_{1,t+1} \\ \mathrm{E}_t y_{2,t+1} \end{bmatrix} = \mathcal{M} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} \theta_{t+1} \\ 0 \end{bmatrix} \tag{B.1}
$$

for $t \geq 0$; where y_{1t} is a vector of m_1 predetermined variables ($y_{1t} \equiv (X_t', \Xi_{t-1}')'$ and $m_1 = n_X + n_x$ in the previous section); y_{2t} is a vector of m_2 non-predetermined variables ($y_{2t} \equiv (x_t', i_t', \xi_t')'$ and $m_2 = n_x + n_i + n_X$ in the previous section); θ_t is an m_1-vector of stochastic processes ($\theta_t \equiv (z_t', 0')'$ in the previous section); and y_{10} is given.

By defining the m_2-vector of endogenous expectation errors, η_t, as

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

(B.1) can be written in the form used in Sims [A6],

$$
\Gamma_0 y_t = \Gamma_1 y_{t-1} + \Psi \theta_t + \Pi \eta_t,
$$

7

where $y_t \equiv (y'_{1t}, y'_{2t})'$. Sims shows that, under suitable assumptions, this system has a unique solution of the form

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

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. These matrices can be calculated by his Matlab program Gensys, available at www.princeton.edu/~sims. An advantage with Sims's approach is that one need not keep track of what variables are predetermined or nonpredetermined. An arguable disadvantage is that the determination of the expectational errors is somewhat complex.

Here, I prefer to keep close track of what variables are predetermined and nonpredetermined and therefore choose to derive the solution to (B.1) following a route closer to Klein [A4] than Sims [A6], but going beyond Klein in, as Sims, explicitly treating the case of θ_t being an arbitrary stochastic process rather than an autoregressive process. The solution will then be of the form

$$
\begin{aligned}
y_{2t} &= F_1 y_{1t} + Z_t, \\
y_{1,t+1} &= M_1 y_{1t} + N \mathrm{E}_t Z_{t+1} + P \mathrm{E}_t \theta_{t+1} + (\theta_{t+1} - \mathrm{E}_t \theta_{t+1}), \\
Z_t &\equiv \sum_{\tau=0}^{\infty} R_\tau \mathrm{E}_t \theta_{t+1+\tau},
\end{aligned}
$$

where F_1, M_1, N, P, and R_τ are real matrices to be determined.

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

$$\mathcal{C} \begin{bmatrix} \mathrm{E}_t y_{1,t+1} \\ \mathrm{E}_t y_{2,t+1} \end{bmatrix} = \mathcal{M} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} \mathrm{E}_t \theta_{t+1} \\ 0 \end{bmatrix}. \tag{B.2}$$

Following Klein [A4], Sims [A6], and Söderlind [20], I use the generalized Schur decomposition. This decomposition results in the square complex matrices Q, S, T, and Z such that

$$\mathcal{C} = Q'SZ', \tag{B.3}$$

$$\mathcal{M} = Q'TZ', \tag{B.4}$$

where Z' for a complex matrix denotes the complex conjugate transpose of Z (the transpose of the complex conjugate of Z).[39] The matrices Q and Z are unitary ($Q'Q = Z'Z = I$), and S and T are upper triangular (see Golub and van Loan [A2]). The decomposition is furthermore ordered so the block consisting of the stable generalized eigenvalues (the jth diagonal element of T divided

---

[39] Let the elements of the complex matrix Z be denoted $z_{jk} \equiv \mathrm{Re}(z_{jk}) + i\,\mathrm{Im}(z_{jk})$. Then the complex conjugate of the matrix Z is the matrix of elements $\bar{z}_{jk} \equiv \mathrm{Re}(z_{jk}) - i\,\mathrm{Im}(z_{jk})$.

by the jth diagonal element of S, $\lambda_j \equiv t_{jj}/s_{jj}$) comes first and the block of unstable generalized eigenvalues comes last.[40]

More precisely, I assume the *saddle-point property* emphasized by Blanchard and Kahn [A1]: The number of eigenvalues with modulus larger than unity equals the number of nonpredetermined variables. Thus, I assume that $|\lambda_j| > 1$ for $m_1 + 1 \le j \le m_1 + m_2$ and $|\lambda_j| < 1$ for $1 \le j \le m_1$ (for an exogenous predetermined variable with a unit root, I can actually allow $|\lambda_j| = 1$ for some $1 \le j \le m_1$).

Define

$$\left[\begin{array}{c} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{array} \right] \equiv Z' \left[\begin{array}{c} y_{1t} \\ y_{2t} \end{array} \right]. \tag{B.5}$$

I can interpret $\tilde{y}_{1t}$ as a complex vector of m_1 transformed predetermined variables and $\tilde{y}_{2t}$ as a complex vector of m_2 transformed non-predetermined variables. Premultiply the system (B.2) by Q and use (B.3)-(B.5) to write it as

$$\left[\begin{array}{cc} S_{11} & S_{12} \\ 0 & S_{22} \end{array} \right] \left[\begin{array}{c} \mathrm{E}_t \tilde{y}_{1,t+1} \\ \mathrm{E}_t \tilde{y}_{2,t+1} \end{array} \right] = \left[\begin{array}{cc} T_{11} & T_{12} \\ 0 & T_{22} \end{array} \right] \left[\begin{array}{c} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{array} \right] + \left[\begin{array}{c} Q_{11} \\ Q_{21} \end{array} \right] \mathrm{E}_t \theta_{t+1}, \tag{B.6}$$

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

Consider the lower block of (B.6),

$$S_{22} \, \mathrm{E}_t \tilde{y}_{2,t+1} = T_{22} \, \tilde{y}_{2t} + Q_{21} \mathrm{E}_t \theta_{t+1}. \tag{B.7}$$

Since the diagonal terms of S_{22} and T_{22} (s_{jj} and t_{jj} for $m_1 + 1 \le j \le m_1 + m_2$) 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

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

$$= \sum_{\tau=0}^{\infty} J^\tau K \mathrm{E}_t \theta_{t+1+\tau} \tag{B.9}$$

for $t \ge 0$, where the complex matrices J and K ($m_2 \times m_2$ and $m_2 \times m_1$, respectively) are given by

$$J \equiv T_{22}^{-1} S_{22}, \tag{B.10}$$

$$K \equiv -T_{22}^{-1} Q_{21}. \tag{B.11}$$

Here, I have exploited that the modulus of the diagonal terms of $T_{22}^{-1} S_{22}$ is less than one. I also assume that $\mathrm{E}_t \tilde{y}_{2,t+\tau}$ and $\mathrm{E}_t \theta_{t+\tau}$ are sufficiently bounded. Then $J^\tau \mathrm{E}_t \tilde{y}_{2,t+\tau} \to 0$ when $\tau \to \infty$, and

---

[40] The sorting of the eigenvalues is often done by two programs written by Sims and available at www.princeton.edu/~sims, Qzdiv and Qzswitch.

the infinite sum on the right side converges. Note that J may not be invertible, since S_{22} may not be invertible.

I have, by (B.5),

$$y_{1t} = Z_{11}\tilde{y}_{1t} + Z_{12}\tilde{y}_{2t}, \tag{B.12}$$

$$y_{2t} = Z_{21}\tilde{y}_{1t} + Z_{22}\tilde{y}_{2t}, \tag{B.13}$$

where

$$Z \equiv \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \tag{B.14}$$

is partitioned conformably with y_{1t} and y_{2t}. Under the assumption of the saddle-point property, Z_{11} is square. I furthermore assume that Z_{11} is *invertible*. Then I can solve for $\tilde{y}_{1t}$ in (B.12),

$$\tilde{y}_{1t} = Z_{11}^{-1} y_{1t} - Z_{11}^{-1} Z_{12} \tilde{y}_{2t}, \tag{B.15}$$

and use this in (B.13) to get

$$y_{2t} = F_1 y_{1t} + H \tilde{y}_{2t}, \tag{B.16}$$

where the real $m_2 \times m_1$ matrix F_1 and the complex $m_2 \times m_2$ matrix H are given by

$$F_1 \equiv Z_{21} Z_{11}^{-1}, \tag{B.17}$$

$$H \equiv Z_{22} - Z_{21} Z_{11}^{-1} Z_{12}. \tag{B.18}$$

I will show below that H is invertible.

By (B.9) and (B.16), I can then write the solution of y_{2t} as

$$y_{2t} = F_1 y_{1t} + Z_t, \tag{B.19}$$

where Z_t is a real exogenous m_2-vector stochastic process (not to be confused with the unitary matrix Z in the Schur decomposition) given by

$$Z_t \equiv H\tilde{y}_{2t} \equiv \sum_{\tau=0}^{\infty} R_\tau \mathrm{E}_t \theta_{t+1+\tau}, \tag{B.20}$$

$$R_\tau \equiv H J^\tau K \quad (\tau \geq 0), \tag{B.21}$$

where the matrices R_τ are real.

I note that the complex conjugate transpose of Z, Z', satisfies

$$Z' \equiv \begin{bmatrix} Z'_{11} & Z'_{21} \\ Z'_{12} & Z'_{22} \end{bmatrix}, \tag{B.22}$$

10

where the submatrices are given by (B.14). Since $Z'Z = ZZ' = I$, I have

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} Z'_{11} & Z'_{21} \\ Z'_{12} & Z'_{22} \end{bmatrix} = \begin{bmatrix} Z_{11}Z'_{11} + Z_{12}Z'_{12} & Z_{11}Z'_{21} + Z_{12}Z'_{22} \\ Z_{21}Z'_{11} + Z_{22}Z'_{12} & Z_{21}Z'_{21} + Z_{22}Z'_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \qquad \text{(B.23)}$$

By (B.22), I can write

$$\tilde{y}_{2t} = Z'_{12}y_{1t} + Z'_{22}y_{2t}.$$

Using this in (B.16) gives

$$\begin{aligned} y_{2t} &= F_1 y_{1t} + H(Z'_{12}y_{1t} + Z'_{22}y_{2t}) \\ &= (F_1 + HZ'_{12})y_{1t} + HZ'_{22}y_{2t}. \end{aligned}$$

It follows that

$$F_1 + HZ'_{12} = 0, \qquad \text{(B.24)}$$

$$HZ'_{22} = I. \qquad \text{(B.25)}$$

I can also show (B.24) by using (B.23),

$$\begin{aligned} Z_{21}Z_{11}^{-1} + (Z_{22} - Z_{21}Z_{11}^{-1}Z_{12})Z'_{12} &= Z_{21}Z_{11}^{-1} + Z_{22}Z'_{12} - Z_{21}Z_{11}^{-1}Z_{12}Z'_{12} \\ &= Z_{21}Z_{11}^{-1} + Z_{22}Z'_{12} - Z_{21}Z_{11}^{-1}(I - Z_{11}Z'_{11}) \\ &= Z_{21}Z_{11}^{-1} + Z_{22}Z'_{12} - Z_{21}Z_{11}^{-1} + Z_{21}Z'_{11} \\ &= 0. \end{aligned}$$

Similarly, I can show (B.25) by

$$\begin{aligned} (Z_{22} - Z_{21}Z_{11}^{-1}Z_{12})Z'_{22} &= Z_{22}Z'_{22} - Z_{21}Z_{11}^{-1}Z_{12}Z'_{22} \\ &= Z_{22}Z'_{22} - Z_{21}Z_{11}^{-1}(-Z_{11}Z'_{21}) \\ &= Z_{22}Z'_{22} + Z_{21}Z'_{21} \\ &= I. \end{aligned}$$

It follows from (B.25) that H is invertible and that its inverse is given by

$$H^{-1} = Z'_{22}. \qquad \text{(B.26)}$$

It remains to find a solution for $y_{1,t+1}$. The upper block of (B.6) is

$$S_{11}\mathrm{E}_t\tilde{y}_{1,t+1} + S_{12}\mathrm{E}_t\tilde{y}_{2,t+1} = T_{11}\tilde{y}_{1t} + T_{12}\tilde{y}_{2t} + Q_{11}\mathrm{E}_t\theta_{t+1}.$$

11

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

$$\mathrm{E}_t\tilde{y}_{1,t+1} = S_{11}^{-1}(T_{11}\tilde{y}_{1t} + T_{12}\tilde{y}_{2t}) - S_{11}^{-1}S_{12}\mathrm{E}_t\tilde{y}_{2,t+1} + S_{11}^{-1}Q_{11}\mathrm{E}_t\theta_{t+1}.$$

By (B.12),

$$
\begin{aligned}
\mathrm{E}_t y_{1,t+1} &= Z_{11}\mathrm{E}_t\tilde{y}_{1,t+1} + Z_{12}\mathrm{E}_t\tilde{y}_{2,t+1}\\
&= Z_{11}[S_{11}^{-1}(T_{11}\tilde{y}_{1t} + T_{12}\tilde{y}_{2t}) - S_{11}^{-1}S_{12}\mathrm{E}_t\tilde{y}_{2,t+1} + S_{11}^{-1}Q_{11}\mathrm{E}_t\theta_{t+1}] + Z_{12}\mathrm{E}_t\tilde{y}_{2,t+1}\\
&= Z_{11}S_{11}^{-1}T_{11}\tilde{y}_{1t} + Z_{11}S_{11}^{-1}T_{12}\tilde{y}_{2t} + (Z_{12} - Z_{11}S_{11}^{-1}S_{12})\mathrm{E}_t\tilde{y}_{2,t+1} + Z_{11}S_{11}^{-1}Q_{11}\mathrm{E}_t\theta_{t+1}\\
&= Z_{11}S_{11}^{-1}T_{11}(Z_{11}^{-1}y_{1t} - Z_{11}^{-1}Z_{12}\tilde{y}_{2t}) + Z_{11}S_{11}^{-1}T_{12}\tilde{y}_{2t} + (Z_{12} - Z_{11}S_{11}^{-1}S_{12})\mathrm{E}_t\tilde{y}_{2,t+1}\\
&\quad + Z_{11}S_{11}^{-1}Q_{11}\mathrm{E}_t\theta_{t+1}\\
&= Z_{11}S_{11}^{-1}T_{11}Z_{11}^{-1}y_{1t} + Z_{11}S_{11}^{-1}(T_{12} - T_{11}Z_{11}^{-1}Z_{12})\tilde{y}_{2t}\\
&\quad + (Z_{12} - Z_{11}S_{11}^{-1}S_{12})\mathrm{E}_t\tilde{y}_{2,t+1} + Z_{11}S_{11}^{-1}Q_{11}\mathrm{E}_t\theta_{t+1}\\
&= Z_{11}S_{11}^{-1}T_{11}Z_{11}^{-1}y_{1t} + Z_{11}S_{11}^{-1}(T_{12} - T_{11}Z_{11}^{-1}Z_{12})(J\mathrm{E}_t\tilde{y}_{2,t+1} + K\mathrm{E}_t\theta_{t+1})\\
&\quad + (Z_{12} - Z_{11}S_{11}^{-1}S_{12})\mathrm{E}_t\tilde{y}_{2,t+1} + Z_{11}S_{11}^{-1}Q_{11}\mathrm{E}_t\theta_{t+1}\\
&= Z_{11}S_{11}^{-1}T_{11}Z_{11}^{-1}y_{1t}\\
&\quad + [Z_{11}S_{11}^{-1}(T_{12} - T_{11}Z_{11}^{-1}Z_{12})J + (Z_{12} - Z_{11}S_{11}^{-1}S_{12})]\mathrm{E}_t\tilde{y}_{2,t+1}\\
&\quad + Z_{11}S_{11}^{-1}[Q_{11} + (T_{12} - T_{11}Z_{11}^{-1}Z_{12})K]\mathrm{E}_t\theta_{t+1}, \tag{B.27}
\end{aligned}
$$

where I have used (B.15) and (B.8).

It follows that I can use (B.27), (B.20), and (B.26) and write the solution as

$$y_{1,t+1} = My_{1t} + N\mathrm{E}_t Z_{t+1} + P\mathrm{E}_t\theta_{t+1} + (\theta_{t+1} - \mathrm{E}_t\theta_{t+1}), \tag{B.28}$$

where the real matrices M, N, and P are given by

$$
\begin{aligned}
M &\equiv Z_{11}S_{11}^{-1}T_{11}Z_{11}^{-1}, \tag{B.29}\\
N &\equiv [Z_{11}S_{11}^{-1}(T_{12} - T_{11}Z_{11}^{-1}Z_{12})J + (Z_{12} - Z_{11}S_{11}^{-1}S_{12})]Z_{22}', \tag{B.30}\\
P &\equiv Z_{11}S_{11}^{-1}[Q_{11} + (T_{12} - T_{11}Z_{11}^{-1}Z_{12})K]. \tag{B.31}
\end{aligned}
$$

Thus, the solution to the system (B.1) is given by (B.19) and (B.28) for $t \geq 0$. This results in the solution (A.11)-(A.12) above, where the matrix P in (A.12) is the submatrix of the first n_X rows of the matrix P in (B.31) (since $\theta_{t+1} \equiv (z_{t+1}', 0')'$).

12

C. The model when judgment is a finite-order moving average

When the deviation is a finite-order moving-average process and the dynamics of the deviation and judgment is described by (2.16), the model can be written as

$$
\begin{bmatrix} X_{t+1} \\ z^{t+1} \\ Cx_{t+1|t} \end{bmatrix} = \bar{A} \begin{bmatrix} X_t \\ z^t \\ x_t \end{bmatrix} + \bar{B}i_t + \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon^{t+1} \\ 0 \end{bmatrix},
\tag{C.1}
$$

where the matrices $\bar{A}$ and $\bar{B}$ are given by

$$
\bar{A} \equiv \begin{bmatrix} A_{11} & A_{z12} & A_{12} \\ 0 & A_{z22} & 0 \\ A_{21} & 0 & A_{22} \end{bmatrix}, \qquad \bar{B} \equiv \begin{bmatrix} B_1 \\ 0 \\ B_2 \end{bmatrix},
$$

the matrix A_z is partitioned conformably with z_t and z^t as

$$
A_z \equiv \begin{bmatrix} 0 & A_{z12} \\ 0 & A_{z22} \end{bmatrix},
$$

and $\tilde{\varepsilon}_t \equiv (\varepsilon_t', \varepsilon^{t\prime})'$ is zero-mean and iid. Thus, this results in the standard forward-looking linear-quadratic model, with the predetermined variables being X_t and z^t. The optimal policy projection can then be described as (2.17) and (2.18), where F and M are finite-dimensional matrices. The intertemporal loss for the optimal policy projection can then be written as

$$
\frac{1}{2} \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix}' V \begin{bmatrix} X_t \\ z^t \\ \Xi_{t-1,t-1} \end{bmatrix},
$$

where the matrix V is the solution to the Lyapunov equation,

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

the symmetric and positive semidefinite matrix $\bar{W}$ is defined by

$$
\bar{W} = \begin{bmatrix} I & 0 & 0 \\ & F_x & \\ & F_i & \end{bmatrix}' D'WD \begin{bmatrix} I & 0 & 0 \\ & F_x & \\ & F_i & \end{bmatrix},
$$

and the matrix F is partitioned conformably with x_t and i_t as

$$
F \equiv \begin{bmatrix} F_x \\ F_i \end{bmatrix}.
$$

D. The Marcet-Marimon method to solve the linear-quadratic optimization problem with forward-looking variables

Let $\bar{X}_t \equiv (X_t, z^t)$ and write the model (C.1) as

$$\bar{X}_{t+1} = \bar{A}_{11}\bar{X}_t + \bar{A}_{12}x_t + \bar{B}_1 i_t + \tilde{\varepsilon}_{t+1}, \tag{D.1}$$

$$C\mathrm{E}_t x_{t+1} = \bar{A}_{21}\bar{X}_t + \bar{A}_{22}x_t + \bar{B}_2 i_t. \tag{D.2}$$

Write the period loss function as

$$L_t = \frac{1}{2} \begin{bmatrix} \bar{X}_t \\ x_t \\ i_t \end{bmatrix}' W^0 \begin{bmatrix} \bar{X}_t \\ x_t \\ i_t \end{bmatrix}, \tag{D.3}$$

where the symmetric positive semidefinite matrix W^0 is defined by

$$\begin{bmatrix} \bar{X}_t \\ x_t \\ i_t \end{bmatrix}' W^0 \begin{bmatrix} \bar{X}_t \\ x_t \\ i_t \end{bmatrix} \equiv \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}' D'WD \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}.$$

Consider the problem in period 0,

$$\min_{\{i_t\}_{t\geq 0}} \mathrm{E}_0 \sum_{t=0}^{\infty} \delta^t L_t, \tag{D.4}$$

subject to (D.1), (D.2) and X_0 given. The minimization is taken to be under commitment.

Marcet and Marimon [14] show that this problem can be reformulated as a recursive saddlepoint problem,

$$\max_{\{\gamma_t\}_{t\geq 0}} \min_{\{x_t, i_t\}_{t\geq 0}} \mathrm{E}_0 \sum_{t=0}^{\infty} \delta^t \tilde{L}_t, \tag{D.5}$$

where the modified period loss function satisfies

$$\begin{aligned} \tilde{L}_t &\equiv \tilde{L}(\bar{X}_t, \Xi_{t-1}; x_t, i_t, \gamma_t) \\ &\equiv L_t + L_t^1 \\ &\equiv L_t + \gamma_t'(-\bar{A}_{21}\bar{X}_t - \bar{A}_{22}x_t - \bar{B}_2 i_t) + \frac{1}{\delta}\Xi_{t-1}'Cx_t, \end{aligned}$$

and the optimization is subject to (D.1), to

$$\Xi_t = \gamma_t, \tag{D.6}$$

and to X_0 and $\Xi_{-1} = 0$ given. The value function for the saddlepoint problem, starting in any period t, satisfies

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

14

subject to (D.1) and (D.6).

Define

$$\tilde{X}_t \equiv \left[\begin{array}{c} \bar{X}_t \\ \Xi_{t-1} \end{array} \right], \qquad \tilde{\imath}_t \equiv \left[\begin{array}{c} x_t \\ i_t \\ \gamma_t \end{array} \right],$$

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

$$\tilde{L}_t \equiv \frac{1}{2} \left[\begin{array}{c} \tilde{X}_t \\ \tilde{\imath}_t \end{array} \right]' \bar{W} \left[\begin{array}{c} \tilde{X}_t \\ \tilde{\imath}_t \end{array} \right], \tag{D.7}$$

$$\tilde{X}_{t+1} = \tilde{A}\tilde{X}_t + \tilde{B}\tilde{\imath}_t + \tilde{C}\tilde{\varepsilon}_{t+1}. \tag{D.8}$$

The problem (D.5) subject to (D.8) and given $\tilde{X}_t$ is isomorphic to a standard backward-looking linear-quadratic problem, 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, (D.4) subject to (D.1) and (D.2).

The value function for the saddlepoint problem is quadratic,

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

where $\tilde{V}$ solves the Riccati equation,

$$\tilde{V} = Q + \delta\tilde{A}'\tilde{V}\tilde{A} - (\delta\tilde{B}'\tilde{V}\tilde{A} + N')'(\delta\tilde{B}'\tilde{V}\tilde{B} + R)^{-1}(\delta\tilde{B}'\tilde{V}\tilde{A} + N'),$$

where

$$\bar{W} \equiv \left[\begin{array}{cc} Q & N \\ N' & R \end{array} \right],$$

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

The optimal reaction function for the saddlepoint problem is linear,

$$\tilde{\imath}_t = F\tilde{X}_t \equiv \left[\begin{array}{c} F_x \\ F_i \\ F_\gamma \end{array} \right] \tilde{X}_t,$$

where F is partitioned conformably with x_t, i_t, and γ_t and satisfies

$$F \equiv -(\delta\tilde{B}'\tilde{V}\tilde{B} + R)^{-1}(\delta\tilde{B}'\tilde{V}\tilde{A} + N').$$

This reaction function function is the optimal reaction function function for the original problem. Optimization in a timeless perspective in period t corresponds to taking Ξ_{t-1} from the previous period's decision problem as given, also in period 0.

The equilibrium dynamics will be given by

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

where

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

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

The value function for the saddlepoint problem can be decomposed according to

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

where

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

is the value function for the original problem starting in period t with $\tilde{X}_t \equiv (X_t', \Xi_{t-1}')'$ given. The matrix V will satisfy the Lyapunov equation,

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

and, when $\delta < 1$, the constant w will satisfy

$$
w = \frac{\delta}{1-\delta}\mathrm{tr}(\tilde{C}'V\tilde{C}\Sigma_{\tilde{\varepsilon}\tilde{\varepsilon}}),
$$

where $\Sigma_{\tilde{\varepsilon}\tilde{\varepsilon}}$ is the covariance matrix for $\tilde{\varepsilon}_t$.

E. An alternative finite-horizon numerical procedure for forward-looking models

In the finite-horizon model in section 3.1, there is an obvious alternative numerical procedure that will provide a projection arbitrarily close to the optimal policy projection without requiring such a long horizon that $X_{t+T,t}$ and $\Xi_{t+T-1,t}$ are close to their steady-state levels. It requires iterations, though.

Assume that iteration $j-1$ has resulted in $\Xi_{t+T-1,t}^{(j-1)}$. Start iteration j by using (2.17) and (2.18) to replace (3.3) by

$$x_{t+T+1,t} = F_x M_1 \begin{bmatrix} X_{t+T,t} \\ \Xi_{t+T-1,t}^{(j-1)} \end{bmatrix},$$

where the matrices F_1 and M_1 are defined by

$$F \begin{bmatrix} X_t \\ 0 \\ \Xi_{t-1} \end{bmatrix} \equiv F_1 \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix}, \qquad M \begin{bmatrix} X_t \\ 0 \\ \Xi_{t-1} \end{bmatrix} \equiv M_1 \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix},$$

and F_1 is partitioned conformably with x_t and i_t as

$$F_1 \equiv \begin{bmatrix} F_x \\ F_i \end{bmatrix}.$$

Consequently, replace (3.4) by

$$-A_{21} X_{t+T,t} - A_{22} x_{t+T,t} - B_2 i_{t+T,t} + CF_x M_1 \begin{bmatrix} X_{t+T,t} \\ \Xi_{t+T-1,t}^{(j-1)} \end{bmatrix} = 0. \tag{E.1}$$

Use (3.1), (3.2), and (E.1) to construct G and g^t (the left submatrix of the matrix $CF_x M_1$ will enter the last block of G and the product of the right submatrix and $\Xi_{t+T-1,t}^{(j-1)}$ will enter the last block of g^t). Furthermore, add the term (3.7) with $\Xi_{t+T-1,t} = \Xi_{t+T-1,t}^{(j-1)}$ to the loss function (that is, modify the diagonal block of Ω that corresponds to $X_{t+T,t}$ and add a linear term that corresponds to the cross products of $X_{t+T,t}$ and $\Xi_{t+T-1,t}^{(j-1)}$). Find the optimal policy projection $\hat{s}^{t(j)}$ and Lagrange multiplier $\Lambda^{t(j)}$ via the analogue of (3.12). This ends iteration j and results in $\Xi_{t+T-1,t}^{(j)}$. Continue until $\Xi_{t+T-1,t}^{(j)}$ has converged.

Obviously this alternative procedure does not require that $X_{t+T,t}$ and $\Xi_{t+T-1,t}$ are close to their steady-state levels. Which procedure is fastest will depend on the number of variables in the problem and the rate of convergence towards the steady state of the optimal policy projection.

F. The feasible set of projections of the states of the economy, the feasible set of projections of the target variables, and the optimal targeting rule

In the finite-horizon projection model in section 3.1, the feasible set of projections in period t of the states of the economy, $\mathcal{S}_t$, is the set of projections s^t that satisfy (3.5), repeated here as

$$Gs^t = g^t. \tag{F.1}$$

That is, $\mathcal{S}_t$ is the set of solutions to (F.1) for given G and g^t. Define $n \equiv (T+1)(n_X + n_x + n_i)$, $m \equiv (T+1)(n_X + n_x) < n$, and $p \equiv (T+1)n_i \equiv n - m$. Note that G is $m \times n$, s^t is $n \times 1$, and g^t is $m \times 1$. Assume that G is of rank m.

Since G is of rank m, the set of solutions to (F.1) is a linear manifold of R^n of dimension $p \equiv n - m$.[41] It can be written as the set of projections s^t that satisfy

$$s^t = G^+ g^t + (I - G^+ G)\xi \qquad \text{(F.2)}$$

for any $\xi \in R^n$ (see Harville [A2, chapters 11 and 20]). Here, the $n \times m$ matrix G^+ is the Moore-Penrose inverse of G. When G is $m \times n$ and of rank m, the Moore-Penrose inverse is given by

$$G^+ = G'(GG')^{-1}$$

(note that GG' is $m \times m$, of rank m, and hence invertible). Then, $G^+ G = G'(GG')^{-1}G$ is a projection matrix that projects vectors in R^n on the m-dimensional column space of the $n \times m$ matrix G', the transpose of G.[42] Denote the column space of G' by $\mathcal{C}(G')$. For any ξ in R^n, the vector $G^+ G\xi$ lies in $\mathcal{C}(G')$. Then $I - G^+ G$ is a projection matrix that projects vectors in R^n *off* the column space of G', that is, on the p-dimensional subspace of R^n orthogonal to $\mathcal{C}(G')$, the orthogonal complement of $\mathcal{C}(G')$ (relative to R^n), denoted $\mathcal{C}^\perp(G')$. Hence, the solution set $\mathcal{S}_t$ consist of $\mathcal{C}^\perp(G')$ shifted away from the origin by the vector $G^+ g^t$,

$$\mathcal{S}_t = \{G^+ g^t\} + \mathcal{C}^\perp(G').$$

Furthermore, the vector $G^+ g^t$ is the s^t of minimum norm that satisfies (F.1). Then, $G^+ g^t$ is orthogonal to the solution set $\mathcal{S}_t$ and lies in the column space of G, $\mathcal{C}(G')$.[43]

Figure F.1 provides an illustration of the above, when $n = 2$ and $m = p = 1$. The linear manifold $\mathcal{S}_t$, the set of feasible projections of the states of the economy, s^t, is shown as the negatively sloped line through the point $s^t = G^+ g^t$. The column space $\mathcal{C}(G')$ is the positively sloped line through the origin. The linear manifold $\mathcal{S}_t$ is orthogonal to the column space. The orthogonal complement of the column space, $\mathcal{C}^\perp(G')$, is the negatively sloped line through the origin. The linear manifold is the orthogonal complement shifted away from the origin to the point $G^+ g^t$. Furthermore, the point $G^+ g^t$ is the point in the linear manifold with the shortest distance to the origin.
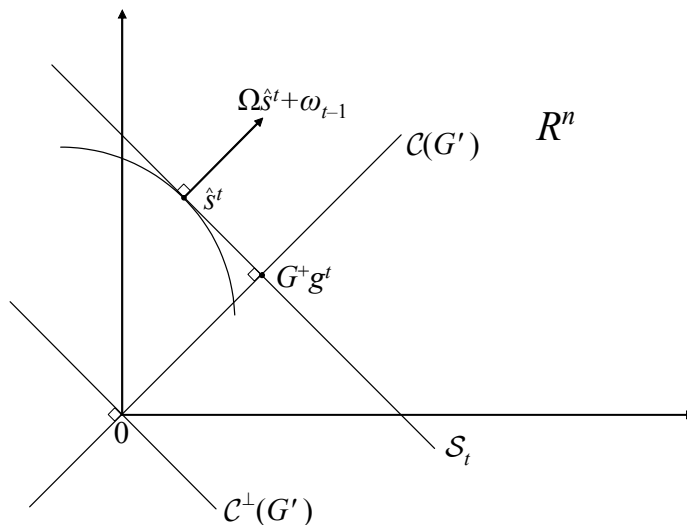
Let $G^\perp$ denote a $p \times n$ matrix with p linearly independent rows, each of which is orthogonal to the m rows of G. Then $\mathcal{C}^\perp(G') = \mathcal{C}(G^{\perp\prime})$, where the latter expression denotes the column space of

---

[41] Let V be a linear space. A subset S of V is a linear manifold of V (also called a linear variety of V), if there is a v in V such that the set $S - \{v\} \equiv \{s - v | s \in S\}$ is a subspace of V. The dimension of S is the dimension of $S - \{v\}$. Hence, a linear manifold is a subspace that has possibly been shifted away from the origin (in the above case by the vector v).

[42] In this section, the word "projection" is used not only to refer to mean forecasts but also, depending on the context, to refer to mathematical projections in linear space.

[43] A vector is orthogonal to a linear manifold if it is orthogonal to the corresponding subspace.

Figure F.1: The set of feasible projections of the state of the economy, $\mathcal{S}_t$

$G^{\perp\prime}$, and $\mathcal{S}_t$ can be written as the set of projections s^t that satisfy

$$s^t = G^+ g^t + G^{\perp\prime}\xi$$

for any $\xi \in R^n$.

The projection of the target variables, Y^t, is a linear function of the projection of the states of the economy according to (3.6), repeated here as
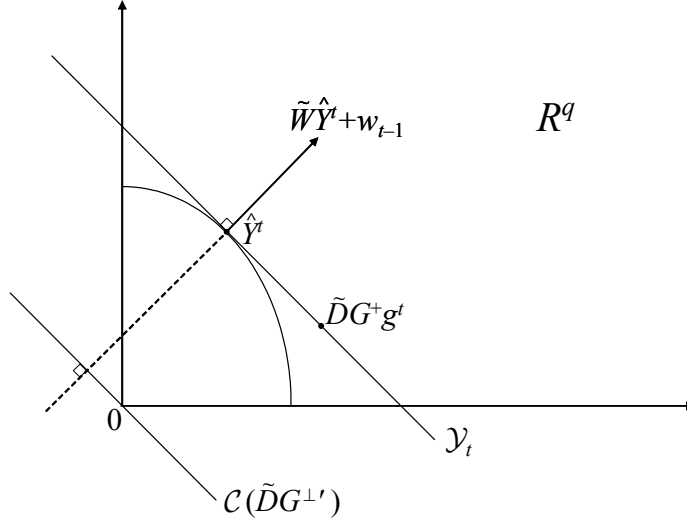
$$Y^t = \tilde{D}s^t. \tag{F.3}$$

Let $q \equiv (T+1)n_Y \leq n$, note that Y^t is $q \times 1$ and $\tilde{D}$ is $q \times n$, and take $\tilde{D}$ to be of rank q. It follows that the set of feasible projections of the target variables, $\mathcal{Y}_t$, consists of the set of projections Y^t that satisfy

$$Y^t = \tilde{D}G^+ g^t + \tilde{D}G^{\perp\prime}\xi$$

for any ξ in R^n. This is a linear manifold of R^q of dimension at most $\min(p, q)$. If I take as the normal case that the number of target variables is at least as large as the number of instruments, $n_Y \geq n_i$ (typically, there are at least two target variables, inflation and the output gap, but only one instrument, the instrument rate), I have $q \geq p$, and the set of feasible projections of the target variables, $\mathcal{Y}_t$, is a linear manifold of R^q of dimension at most $p \leq q$. The matrix $\tilde{D}$ simply maps the p-dimensional linear manifold $\mathcal{S}_t$ of R^n into the at most p-dimensional linear manifold $\mathcal{Y}_t$ of R^q.

19

Figure F.2: The set of feasible projections of the target variables, $\mathcal{Y}_t$

It follows that $\mathcal{Y}_t$ is the at most p-dimensional column space $\mathcal{C}(\tilde{D}G^{\perp\prime})$ in R^q shifted away from the origin by the vector $\tilde{D}G^+g^t$,

$$\mathcal{Y}_t = \{\tilde{D}G^+g^t\} + \mathcal{C}(\tilde{D}G^{\perp\prime}).$$

Figure F.2 provides an illustration of the above, when $q = 2$ and $p = 1$. The linear manifold $\mathcal{Y}_t$, the set of feasible projections of the target variables, Y^t, is shown as the negatively sloped line through the point $Y^t = \tilde{D}G^+g^t$. The column space of the matrix $\tilde{D}G^{\perp\prime}$, $\mathcal{C}(\tilde{D}G^{\perp\prime})$, is shown as the negative sloped line through the origin. The linear manifold $\mathcal{Y}_t$ is this column space shifted away from the origin to the point $\tilde{D}G^+g^t$.

F.1. An optimal targeting rule for the forward-looking model

Consider the first-order condition for optimal policy under commitment in a timeless perspective in the forward-looking model, (3.10), rewritten here as

$$\Omega s^t + \omega_{t-1} + G'\Lambda^t = 0 \tag{F.4}$$

The *optimal targeting rule* is the first-order condition in terms of Y^t when the Lagrange multiplier has been eliminated.

Let me interpret the first-order condition in terms of s^t, eliminate the Lagrange multiplier, and interpret the resulting targeting rule. Note that Ω is $n \times n$, s^t and ω_{t-1} are $n \times 1$, G' is $n \times m$ and of rank m, and Λ^t is $m \times 1$.

Write the first-order condition as

$$\Omega s^t + \omega_{t-1} = G'(-\Lambda^t). \tag{F.5}$$

The term $\Omega s^t + \omega_{t-1}$ on the left side is the gradient of the loss function with respect to s^t, a vector in R^n. The condition (F.5) can be interpreted as stating that the gradient of the loss function is an element of the m-dimensional column space of the $n \times m$ matrix G', $\mathcal{C}(G')$, with $-\Lambda^t$ providing the coefficients of the corresponding linear combination of the column vectors of G'. This is equivalent to the tangency of the loss function's iso-loss surface in R^n with the feasible set of projections, $\mathcal{S}_t$. The gradient of the loss function is orthogonal to the iso-loss surface. Tangency of the iso-loss surface with $\mathcal{S}_t$ is then equivalent to the gradient being orthogonal to $\mathcal{S}_t$. The subspace orthogonal to $\mathcal{S}_t$ is $\mathcal{C}(G')$, as noted above.

This is illustrated in figure F.1 when $n = 2$ and $m = p = 1$. The curve shows part of the iso-loss surface of the loss function that is tangential to the linear manifold $\mathcal{S}_t$. The tangency occurs at the optimal policy projection, $\hat{s}^t$. The gradient of the loss function at that point, $\Omega \hat{s}^t + \omega_{t-1}$, is shown as the vector pointing northeast from that point. Tangency between the iso-loss surface and the linear manifold is equivalent to the gradient being orthogonal to the linear manifold, or the gradient being an element in the column space, $\mathcal{C}(G')$.

In order to eliminate the Lagrange multipliers, premultiply (F.5) by G,[44]

$$G(\Omega s^t + \omega_{t-1}) = GG'(-\Lambda^t). \tag{F.6}$$

Exploit that GG' is $m \times m$, of rank m, and hence invertible, and solve for $-\Lambda^t$,

$$-\Lambda^t = (GG')^{-1}G(\Omega s^t + \omega_{t-1}). \tag{F.7}$$

(The matrix $(GG')^{-1}G$ is actually the Moore-Penrose inverse of G', G'^{+}, where G' is $n \times m$ with rank m.) Substitution of Λ^t in (F.4) gives

$$M(\Omega s^t + \omega_{t-1}) = 0, \tag{F.8}$$

where M is the $n \times n$ matrix (not to be confused with the matrix denoted M in other sections of this paper)

$$M \equiv I - G'(GG')^{-1}G = I - G^{+}G.$$

---

[44] One might ask why multipliying with the matrix G with rank $m < n$ rather than a matrix with full rank n does not loose any information of (F.5). More formally, let $G^{\perp}$ be a $p \times n$ matrix whose p rows are linearly independent and orthogonal to the m rows of G. That is, the column space of $G^{\perp\prime}$ is the space in R^n orthogonal to the column space of G'. Then the $n \times n$ matrix $\begin{bmatrix} G \\ G^{\perp} \end{bmatrix}$ is of full rank. Multiplying (F.5) by this matrix leads to the m equations of (F.6) and p additional trivial equations of zero equals zero, since we know that the left and right sides of (F.5) lie in the column space of G'.

As noted above, M is the projection matrix that projects vectors in R^n on the p-dimensional orthogonal complement of the column space of G', $\mathcal{C}^\perp(G')$. Hence, (F.8) states that the projection on $\mathcal{C}^\perp(G')$ of the gradient of the loss function is zero. Of course, this follows directly from the observation above that the gradient lies in $\mathcal{C}(G')$.

In any case, the optimal targeting rule in terms of s^t is equivalent to the statement that the gradient is orthogonal to the feasible set of projections of the states of the economy, $\mathcal{S}_t$, which can be expressed algebraically as (F.8).

However, (F.8) involves n equations, but only p independent equations. It is hence desirable to condense (F.8) to only p equations. The projection matrix M is a symmetric idempotent matrix of rank p. Then its spectrum consists of p eigenvalues equal to one and m eigenvalues equal to zero, and it can be decomposed as

$$ M = Q \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} Q' \equiv \begin{bmatrix} Q_p & Q_m \end{bmatrix} \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q'_p \\ Q'_m \end{bmatrix} \equiv Q_p Q'_p. $$

Here Q is the orthonormal $n \times n$ matrix whose columns are the eigenvectors of M, I_p is the $p \times p$ identity matrix, and Q_p is the $n \times p$ matrix whose columns are the p eigenvectors corresponding to the p nonzero eigenvalues. Then, pre-multiplying (F.8) by Q' gives the p nontrivial equations,

$$ Q'_p(\Omega s^t + \omega_{t-1}) = 0, \tag{F.9} $$

and m trivial equations of zero equals zero.

Furthermore, (F.9) is expressed in terms of the projection of the states of the economy, s^t. In order to express it in terms of the projection of the target variables, Y^t, note that, by the definition of Ω for the forward-looking model in section 3.1,

$$ \Omega s^t \equiv \tilde{D}' \tilde{W} \tilde{D} s^t \equiv \tilde{D}' \tilde{W} Y^t, $$

where $\tilde{W}$ is a symmetric positive semidefinite block-diagonal $(T+1)n_Y$ matrix with the $(\tau+1)$-th diagonal block being $\delta^\tau W$ for $0 \leq \tau \leq T$. Hence, I can write (F.9) as involving only the target variables and, through the vector ω_{t-1}, the Lagrange multiplier $\Xi_{t-1,t-1}$ from the optimization in period $t-1$,

$$ Q'_p[\tilde{D}' \tilde{W} Y^t + \omega_{t-1}] = 0. \tag{F.10} $$

This is one concise form of the targeting rule. The history-dependence of the optimal policy under commitment in a timeless perspective enters via $\Xi_{t-1,t-1}$.

By combining (F.9) with (3.5), I get

$$\begin{bmatrix} G \\ Q_p'\Omega \end{bmatrix} s^t = \begin{bmatrix} g^t \\ -Q_p'\omega_{t-1} \end{bmatrix},$$

and

$$\hat{s}^t = \begin{bmatrix} G \\ Q_p'\Omega \end{bmatrix}^{-1} \begin{bmatrix} g^t \\ -Q_p'\omega_{t-1} \end{bmatrix} \equiv H \begin{bmatrix} X_t \\ \Xi_{t-1,t-1} \\ z^t \end{bmatrix}, \tag{F.11}$$

$$\hat{Y}^t = \tilde{D}\hat{s}^t = \tilde{D}H \begin{bmatrix} X_t \\ \Xi_{t-1,t-1} \\ z^t \end{bmatrix}.$$

From (F.7) and (F.11), I can extract

$$\Xi_{t,t} = H_\Xi \begin{bmatrix} X_t \\ \Xi_{t-1,t-1} \\ z^t \end{bmatrix},$$

to be used in the intertemporal loss function for the decision problem in period $t+1$.

If the forward-looking variables, x_t, are target variables—elements of Y_t—the intertemporal loss function with the added term can be written

$$\frac{1}{2}Y^{t\prime}\tilde{W}Y^t + w'_{t-1}Y^t,$$

where w_{t-1} is a q-vector whose only nonzero elements contain the vector $(\Xi_{t-1,t-1}\frac{1}{\delta}C)'$ such that $w'_{t-1}Y^t \equiv \Xi_{t-1,t-1}\frac{1}{\delta}Cx_{t,t}$. Then, the optimal targeting rule can be expressed as the gradient, $\tilde{W}Y^t + w_{t-1}$, being orthogonal to the linear manifold $\mathcal{Y}_t$. Suppose $\mathcal{Y}_t$ is of dimension p, and let $F \equiv \tilde{D}G^{\perp\prime}$ (not to be confused with the matrix denoted F in other sections of the paper). The projection matrix that projects vectors in R^q on the p-dimensional subspace $\mathcal{Y}_t - \{\tilde{D}G^+g^t\}$ is then $F(F'F)^{-1}F'$, so the condition that the gradient is orthogonal to the linear manifold $\mathcal{Y}^t$ can be written as the p equations.

$$F(F'F)^{-1}F'\tilde{W}(\Omega Y^t + w_{t-1}) = 0.$$

This is the optimal targeting rule for this case.

This case is illustrated in figure F.2. The curve in the figure shows a part of the iso-loss surface of the loss function that is tangential to the linear manifold $\mathcal{Y}_t$. The tangency point is the optimal policy projection of the target variables, $\hat{Y}^t$. The gradient of the loss function at that point, $\tilde{W}\hat{Y}^t + w_{t-1}$, is shown as the vector at that point that points northeast. It is orthogonal to the linear manifold.

Svensson [25] interprets optimal targeting rules in terms of the equality between the marginal rates of transformation and marginal rates of substitution between the target variables. A vector

of marginal rates of transformation between the target variables is a vector in the column space $\mathcal{C}(\tilde{D}G^{\perp\prime})$, the subspace associated with $\mathcal{Y}_t$. A vector of marginal rates of substitution between the target variables is a vector in the tangent space of the intertemporal loss function, the subspace orthogonal to the gradient of the loss function. Equality between the marginal rates of transformation and substitution is equivalent to the gradient being orthogonal to $\mathcal{Y}_t$, that is, the iso-loss surface being tangential to $\mathcal{Y}_t$.

G. An optimal restricted instrument rule

Add to the model (2.1) an explicit instrument rules of the form

$$i_t = f_X X_t, \tag{G.1}$$

where the $n_i \times n_X$ matrix f_X is restricted to be an element $f_X \in \mathcal{F}$ of a given class $\mathcal{F}$ of instrument rules. Assume that the deviation z_t is an iid zero-mean process with variance-covariance matrix Σ. Let the loss function in period t be

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

where L_t is given by (2.7). By appendix A, for a given instrument rule f_X, the conditional loss in period t is, for a given δ ($0 < \delta < 1$), given by

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

where $V(f_X,\delta)$ is a symmetric positive semidefinite $n_X \times n_X$ matrix that depends on A, B, C, D, W, f_X, and δ. It follows that

$$\mathrm{E}[L_t] = \frac{1}{2}\mathrm{trace}[V(f_X,1)\Sigma].$$

The optimal restricted instrument rule, $\hat{f}_X$, is then given by

$$\hat{f}_X = \arg\min_{f_X \in \mathcal{F}} \frac{1}{2}\mathrm{trace}[V(f_X,1)\Sigma].$$

It depends on the class $\mathcal{F}$ and the variance-covariance matrix Σ, in addition to A, B, C, D, and W.

Note that there is little point in considering implicit instrument rules here,

$$i_t = f_X X_t + f_x x_t. \tag{G.2}$$

For any such implicit instrument rule $f \equiv [f_X \ f_x]$ for which a unique equilibrium exists,

$$x_t = g(f)X_t,$$

where the matrix $g(f)$ depends on f. Then,

$$i_t = [f_X + f_x g(f)]X_t \equiv \tilde{f}_X(f)X_t.$$

That is, for each implicit instrument rule f for which there is a unique equilibrium, there is a unique explicit instrument rule $\tilde{f}_X(f)$ consistent with that equilibrium. Furthermore, for any explicit instrument rule f_X in (G.1) for which there is a unique equilibrium, there is a *continuum* of implicit instrument rules consistent with that equilibrium. For any given instrument rule f_X for which there exists a unique equilibrium,$x_t = g(f_X)X_t$, where the matrix $g(f_X)$ depends on f_X. For any arbitrary $n_i \times n_x$ matrix f_x, I can then write

$$i_t = f_X X_t + f_x[x_t - g(f_X)X_t] = [f_X - f_x g(f_X)]X_t + f_x x_t \equiv \tilde{f}_X(f_X, f_x)X_t + f_x x_t.$$

The only reason for considering implicit instrument rules rather than an explicit instrument rule in this context (when the deviation is an iid zero-mean shock) is when an explicit instrument rule has a determinacy problem—multiple equilibria—in which case one may be able to find a corresponding implicit instrument rule for which there is a unique equilibrium. Svensson and Woodford [30] examine such issues further.

H. An empirical backward-looking model

The two equations of the model of Rudebusch and Svensson [18] are

$$\pi_{t+1} = \alpha_{\pi 1}\pi_t + \alpha_{\pi 2}\pi_{t-1} + \alpha_{\pi 3}\pi_{t-2} + \alpha_{\pi 4}\pi_{t-3} + \alpha_y y_t + z_{\pi,t+1} \tag{H.1}$$

$$y_{t+1} = \beta_{y1}y_t + \beta_{y2}y_{t-1} - \beta_r \left(\frac{1}{4}\Sigma_{j=0}^3 i_{t-j} - \frac{1}{4}\Sigma_{j=0}^3 \pi_{t-j} \right) + z_{y,t+1}, \tag{H.2}$$

where π_t is quarterly inflation in the GDP chain-weighted price index (p_t) in percentage points at an annual rate, i.e., $400(\ln p - \ln p_{t-1})$; i_t is the quarterly average federal funds rate in percentage points at an annual rate; y_t is the relative gap between actual real GDP (q_t) and potential GDP (q_t^*) in percentage points, i.e., $100(q_t - q_t^*)/q_t^*$. These five variables were de-meaned prior to estimation, so no constants appear in the equations.

The estimated parameters, using the sample period 1961:1 to 1996:2, are shown in table H.1.

Table H.1

$\alpha_{\pi 1}$	$\alpha_{\pi 2}$	$\alpha_{\pi 3}$	$\alpha_{\pi 4}$	α_y	β_{y1}	β_{y2}	β_r
0.70	-0.10	0.28	0.12	0.14	1.16	-0.25	0.10
(0.08)	(0.10)	(0.10)	(0.08)	(0.03)	(0.08)	(0.08)	(0.03)

The hypothesis that the sum of the lag coefficients of inflation equals one has a p-value of .16, so this restriction was imposed in the estimation.[45]

The state-space form can be written

$$
\begin{bmatrix} \pi_{t+1} \\ \pi_t \\ \pi_{t-1} \\ \pi_{t-2} \\ y_{t+1} \\ y_t \\ i_t \\ i_{t-1} \\ i_{t-2} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{4} \alpha_{\pi j} e_j + \alpha_y e_5 \\ e_1 \\ e_2 \\ e_3 \\ \beta_r e_{1:4} + \beta_{y1} e_5 + \beta_{y2} e_6 - \beta_r e_{7:9} \\ e_5 \\ e_0 \\ e_7 \\ e_8 \end{bmatrix} \begin{bmatrix} \pi_t \\ \pi_{t-1} \\ \pi_{t-2} \\ \pi_{t-3} \\ y_t \\ y_{t-1} \\ i_{t-1} \\ i_{t-2} \\ i_{t-3} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\frac{\beta_r}{4} \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} i_t + \begin{bmatrix} z_{\pi,t+1} \\ 0 \\ 0 \\ 0 \\ z_{y,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},
$$

where e_j $(j = 0, 1, ..., 9)$ denotes a 1×9 row vector, for $j = 0$ with all elements equal to zero, for $j = 1, ..., 9$ with element j equal to unity and all other elements equal to zero; and where $e_{j:k}$ $(j < k)$ denotes a 1×9 row vector with elements $j, j+1, ..., k$ equal to $\frac{1}{4}$ and all other elements equal to zero. The predetermined variables are π_t, π_{t-1}, π_{t-2}, π_{t-3}, y_t, y_{t-1}, i_{t-1}, i_{t-2}, i_{t-2}, and i_{t-3}. There are no forward-looking variables.

For a loss function (5.3) with $\delta = 1$, $\lambda = 1$, and $\nu = 0.2$, and the case where z_t is an iid zero-mean shock; the optimal reaction function (2.21) is (the coefficients are rounded to two decimal points),

$$i_t = 1.22\,\pi_t + 0.43\,\pi_{t-1} + 0.53\,\pi_{t-2} + 0.18\,\pi_{t-3} + 1.93\,y_t - 0.49\,y_{t-1} + 0.36\,i_{t-1} - 0.09\,i_{t-2} - 0.05\,i_{t-3}.$$

I. An empirical forward-looking model

An empirical New Keynesian model estimated by Lindé [13] is

$$\pi_t = \omega_f \pi_{t+1|t} + (1 - \omega_f)\pi_{t-1} + \gamma y_t + z_{\pi t},$$

$$y_t = \beta_f y_{t+1|t} + (1 - \beta_f)(\beta_{y1} y_{t-1} + \beta_{y2} y_{t-2} + \beta_{y3} y_{t-3} + \beta_{y4} y_{t-4}) - \beta_r(i_t - \pi_{t+1|t}) + z_{yt},$$

where the restriction $\sum_{j=1}^{4} \beta_{yj} = 1$ is imposed. The estimated coefficients are (Table 6a in Lindé [13], non-farm business output) are shown in table I.1.

---

[45] This p-value was obtained by simulating the above inflation equation 1000 times and ranking the sum of coefficients from the unrestricted Phillips curve estimated from the actual data (i.e., 0.969) in the set of unrestricted sums estimated from the simulated data. This is in the spirit of Rudebusch [A5]. For comparison, the simple t-test gives a p-value of 0.42.

<div align="center">

Table I.1

</div>

ω_f	γ	β_f	β_r	β_{y1}	β_{y2}	β_{y3}
0.457	0.048	0.425	0.156	1.310	-0.229	-0.011
(0.065)	(0.007)	(0.027)	(0.016)	(0.174)	(0.279)	(0.037)

For simplicity, I set $\beta_{y1} = 1$, $\beta_{y2} = \beta_{y3} = \beta_{y4} = 0$. Then the state-space form can be written as

$$
\begin{bmatrix}
\pi_t \\
y_t \\
i_t \\
z_{\pi,t+1} \\
z_{y,t+1} \\
\omega_f \pi_{t+1|t} \\
\beta_r \pi_{t+1|t} + \beta_f y_{t+1|t}
\end{bmatrix}
=
$$

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-(1-\omega_f) & 0 & 0 & -1 & 0 & 1 & -\gamma \\
0 & -(1-\beta_f) & 0 & 0 & -1 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\pi_{t-1} \\
y_{t-1} \\
i_{t-1} \\
z_{\pi t} \\
z_{yt} \\
\pi_t \\
y_t
\end{bmatrix}
+
\begin{bmatrix}
0 \\
0 \\
1 \\
0 \\
0 \\
0 \\
\beta_r
\end{bmatrix}
i_t +
\begin{bmatrix}
0 \\
0 \\
0 \\
z_{\pi,t+1} \\
z_{y,t+1} \\
0 \\
0
\end{bmatrix}.
$$

The predetermined variables are π_{t-1}, y_{t-1}, i_{t-1}, $z_{\pi t}$, and z_{yt}, and the forward-looking variables are π_t and y_t.

For a loss function (5.3) with $\delta = 1$, $\lambda = 1$, and $\nu = 0.2$, and the case where z_t is an iid zero-mean shock; the optimal reaction function (2.21) is (the coefficients are rounded to two decimal points),

$$
i_t = 0.58\,\pi_{t-1} + 0.80\,y_{t-1} + 0.41\,i_{t-1} + 1.06\,z_{\pi t} + 1.38\,z_{yt} + 0.02\,\Xi_{\pi,t-1,t-1} + 0.20\,\Xi_{y,t-1,t-1},
$$

where $\Xi_{\pi,t-1,t-1}$ and $\Xi_{y,t-1,t-1}$ are the Lagrange multipliers for the two equations for the forward-looking variables in the decision problem in period $t-1$.

References

[A1] Blanchard, Olivier J., and Charles M. Kahn (1980), "The Solution of Linear Difference Models under Rational Expectations," *Econometrica* 48, 1305–1311.

[A2] Golub, Gene H., and Charles F. van Loan (1989), *Matrix Computation*, 2nd ed., Johns Hopkins University Press, Baltimore, MD.

[A2] Harville, David A. (1997), *Matrix Algebra From a Statistician's Perspective*, Springer-Verlag, New York.

[A4] Klein, Paul (2000), "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control* 24, 1405–1423.

[A5] Rudebusch, Glenn D. (1992), "Trends and Random Walks in Macroeconomic Time Series: A Re-Examination," *International Economic Review* 33, 661–680.

[A6] Sims, Christopher A. (2000), "Solving Linear Rational Expectations Models," working paper, www.princeton.edu/∼sims.

[A7] Svensson, Lars E.O. (2005), "Optimization under Commitment and Discretion, and Targeting Rules and Instrument Rules: Lecture Notes," www.princeton.edu/∼svensson.