

NBER WORKING PAPER SERIES

OPTIMAL POLICY WITH PARTIAL INFORMATION
IN A FORWARD-LOOKING MODEL:
CERTAINTY-EQUIVALENCE REDUX

Lars E. O. Svensson
Michael Woodford

Working Paper 9430
<http://www.nber.org/papers/w9430>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2003

We thank Kosuke Aoki and Ernst Schaumburg for comments on an earlier draft. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2003 by Lars E. O. Svensson and Michael Woodford. All rights reserved. Short sections of text not to exceed two paragraphs, may be quoted without explicit permission provided that full credit including, © notice, is given to the source.

Optimal Policy with Partial Information in a Forward-Looking Model:
Certainty-Equivalence Redux
Lars E. O. Svensson and Michael Woodford
NBER Working Paper No. 9430
January 2003
JEL No. E37, E47, E52, E58

ABSTRACT

This paper proves a certainty equivalence result for optimal policy under commitment with symmetric partial information about the state of the economy in a model with forward-looking variables. This result is used in our previous paper, "Indicator Variables for Optimal Policy," which synthesizes what is known about the case of symmetric partial information, and derives useful general formulas for computation of the optimal policy response coefficients and efficient estimates of the state of the economy in the context of a fairly general forward-looking rational-expectations model. In particular, our proof takes into account that, under commitment, the policymaker can affect the future evolution of the observable variables, and thereby potentially affect the future information available.

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

Michael Woodford
Department of Economics
111 Fisher Hall
Princeton University
Princeton, NJ 08544-1021
and NBER
woodford@princeton.edu

1 Introduction

Monetary policy is inevitably conducted under considerable uncertainty about the state of the economy and the nature of recent disturbances. Analyses of optimal policy that take no account of this are therefore of doubtful practical utility. However, in the case of purely backward-looking models of the kind exclusively used by central banks prior to the 1990s, powerful general principles for efficient estimation of the state of the economy and for determining the optimal use to make of such estimates have been well-understood since at least the 1970s. In the case of a linear economic model, a quadratic loss function for the policymaker, uncertainty only about the state of the economy (that is, the current values of specific additive terms in the economic model), and Gaussian disturbances, a principle of *certainty equivalence* applies: the optimal policy is the same as if the state of the economy were fully observable, except that one responds to an efficient estimate of the state of the economy rather than to its actual value. Moreover, a *separation principle* applies, according to which the determination of the optimal response coefficients to be applied to one's estimate of the state of the economy (the optimization problem) and the estimation of the current state of the economy (the estimation or signal-extraction problem) can be treated as separate problems. The optimal response coefficients are independent of the specification of the central bank's incomplete information; and the optimal weights to place on alternative indicators in estimating the state of the economy are independent of the central bank's objective function.¹

However, the presence of forward-looking variables in the system to be controlled — a common feature of modern macroeconomic models, including the econometric models now used by many central banks — complicates matters in a number of respects. For example, optimal policy under commitment ceases in general to coincide with the outcome of discretionary optimization, as demonstrated for the general linear model with quadratic objectives in Backus and Driffill [1] and Currie and Levine [3]. Optimal policy under commitment (even in the deterministic case) is no longer a function solely of the vector of predetermined variables that suffices to characterize the set of possible future paths for the economy from a given date onward; thus one cannot expect that in the case of partial information optimal policy can depend solely on the optimal estimate of such a vector of predetermined variables.

Moreover, in the presence of partial information, estimation of the current state of the system is no longer so simple. For currently observable variables will generally depend not only on the

¹ Important early treatments include Chow [2], Kalchbrenner and Tinsley [4], and Leroy and Waud [5].

current vector of predetermined variables and random observation error with known properties, but also upon forward-looking variables, the values of which will depend on the private sector's expectations about future policy. This makes it far from obvious that a separation principle should apply, even in a linear-quadratic Gaussian framework. Because the relation between the unobserved state of the economy and the observable variables depends on expected policy, one may not be able to solve the optimal filtering problem independently of the solution for the optimal policy response to the estimated state of the economy.

Nonetheless, analogs of the classical control-theoretic results have been obtained for certain special kinds of forward-looking models with partial information. With regard to the estimation problem, Pearlman, Currie and Levine [8] have shown in a linear (non-optimizing) model with forward-looking variables and partial information that the state of the economy can still be estimated using a Kalman filter, although the solution is much more complex than in the purely backward-looking case. Pearlman [7] has used this solution in an optimizing model to demonstrate that certainty equivalence applies under both discretion and commitment in the presence of forward-looking variables and symmetric partial information, that is, in the case that both the central bank and the private sector have access to the same partial information. In the case of commitment, "certainty equivalence" means that the optimal instrument settings are the same linear function of the current estimate of the predetermined variables describing the state of the economy *and specific Lagrange multipliers* (related to the value that alternative expectations would have had in the previous period's policy problem) as in the case of the corresponding optimal policy problem under certainty.

Our previous paper [9] synthesizes what is known about the case of symmetric partial information, and derives useful general formulas for computation of the optimal policy response coefficients and efficient estimates of the state of the economy in the context of a fairly general forward-looking (rational-expectations) model. We find that not only does certainty equivalence continue to characterize optimal policy, but that a separation of the problems of optimal policy and optimal estimation of the current state of the economy continues to be possible, in that the coefficients of the optimal Kalman filter are again independent of the central bank's objective function.

The proof of certainty equivalence under commitment was not included in [9], in order to save space. The present paper provides this proof. The proof is for a more general model than in Pearlman [7] and more intuitive. In particular, our proof explicitly takes into account that,

under commitment, the policymaker can affect the future evolution of the observable variables, and thereby potentially affect the future information available.

Section 2 lays out the model, section 3 demonstrates certainty equivalence for the case of full information, and section 4 proves certainty equivalence for partial information. Section 5 outlines the separation principle, and section 6 concludes. Appendix A contains technical details regarding the degree to which the policymaker can affect the information revealed by the observable variables.

2 The model

Consider the linear model

$$\begin{bmatrix} X_{t+1} \\ Qx_{t+1|t} \end{bmatrix} = A^1 \begin{bmatrix} X_t \\ x_t \end{bmatrix} + A^2 \begin{bmatrix} X_{t|t} \\ x_{t|t} \end{bmatrix} + Bi_t + \begin{bmatrix} u_{t+1} \\ 0 \end{bmatrix}, \quad (2.1)$$

where X_t is a vector of n_X predetermined variables, x_t is a vector of n_x forward-looking variables, i_t is a vector of the central bank's n_i policy instruments, u_t is a vector of n_X iid shocks with mean zero and covariance matrix Σ_{uu} , and A^1 , A^2 , B and Q are matrices of appropriate dimension. The $n_x \times n_x$ matrix Q may be singular (this is a slight generalization of usual formulations when Q is the identity matrix). For any variable z_t , $z_{\tau|t}$ denotes $E[z_\tau|I_t]$, the rational expectation (the best estimate) of z_τ given the information I_t , the information available in period t . The information is specified below. Let Y_t denote a vector of n_Y target variables given by

$$Y_t = C^1 \begin{bmatrix} X_t \\ x_t \end{bmatrix} + C^2 \begin{bmatrix} X_{t|t} \\ x_{t|t} \end{bmatrix} + C_i i_t, \quad (2.2)$$

and let

$$L_t = Y_t' W Y_t \quad (2.3)$$

be a period loss function, where W is a positive-semidefinite weight matrix.

Let the vector of n_Z observable variables, Z_t , be given by

$$Z_t = D^1 \begin{bmatrix} X_t \\ x_t \end{bmatrix} + D^2 \begin{bmatrix} X_{t|t} \\ x_{t|t} \end{bmatrix} + v_t, \quad (2.4)$$

where v_t , the vector of noise, is iid with mean zero and covariance matrix Σ_{vv} . The information I_t in period t is given by

$$I_t = \{Z_\tau, A^1, A^2, B, C^1, C^2, C_i, D^1, D^2, Q, W, \delta, \Sigma_{uu}, \Sigma_{vv}; \tau \leq t\}, \quad (2.5)$$

where δ ($0 < \delta < 1$) is a discount factor. This incorporates the case when some or all of the predetermined and forward-looking variables are observable.

Note that (2.1) assumes that the expectations $x_{t+1|t}$ in the second block of equations are conditional on the information I_t . The case when these expectations are replaced by a private sector expectations $E[x_{t+1}|I_t^p]$ where the private-sector information I_t^p differs from I_t is treated in Svensson and Woodford [10].

Suppose that the central bank commits itself in an initial *ex ante* state (prior to the realization of any period zero random variables) to a state-contingent plan for the indefinite future that minimizes the expected discounted losses

$$E \sum_{t=0}^{\infty} \delta^t L_t.$$

Here E indicates the expectation with respect to information in the initial state, in which the commitment is made. It is important to consider optimal commitment from such an *ex ante* perspective, because, in the case of partial information, the information that the central bank possesses in any given state depends upon the way that it has committed itself to behave in other states that might have occurred instead.

We begin by reviewing the form of the commitment equilibrium under full information, when Z_t includes all elements of X_t and x_t . We then turn to the consequences of partial information, when the information is given by (2.5).

3 Certainty equivalence under full information

Note that in the case of full information, $X_{t|t} = X_t, x_{t|t} = x_t$, as a result of which it is obvious that only the aggregated matrices $A \equiv A^1 + A^2$ and $C \equiv C^1 + C^2$ matter to the optimization problem.

The Lagrangian for the commitment problem can be written in the form

$$\begin{aligned} \mathcal{L} = E & \left[\sum_{t=0}^{\infty} \delta^t L_t - \sum_{t=0}^{\infty} \delta^t \xi'_{t+1} (X_{t+1} - A_{11}X_t - A_{12}x_t - B_1i_t - u_{t+1}) \right. \\ & \left. - \sum_{t=0}^{\infty} \delta^t \Xi'_t (QE_t x_{t+1} - A_{21}X_t - A_{22}x_t - B_2i_t) - \delta^{-1} \xi'_0 (X_0 - \bar{X}_0 - u_0), \right] \end{aligned}$$

where in each period t , ξ_t and Ξ_t are vectors of Lagrange multipliers conformable to X_t and x_t respectively, and where $E_t[\cdot] \equiv E[\cdot|I_t^f]$ denotes expectations conditional on the full information, I_t^f . The dating of the multipliers indicate the period information with which they are measurable

(that is, depend on). Thus, the constraint for X_{t+1} , the predetermined variables, depends on information available in period $t + 1$, I_{t+1}^f , whereas the constraint for $QE_t x_{t+1}$, the forward-looking variables, depends on the information available in period t , I_t^f . (That is, there is only one such latter constraint for each information I_t^f , that may be reached in period t , so there is only one vector of multipliers Ξ_t for each I_t^f ; in other words, Ξ_t depends only on the information available in period t .) The final term on the right-hand side corresponds to the constraint imposed by the vector of initial conditions on X_0 ,

$$X_0 = \bar{X}_0 + u_0, \quad (3.1)$$

where \bar{X}_0 is known at the time of commitment.

Using the law of iterated expectations ($EE_t x_{t+1} = Ex_{t+1}$)², we may equivalently write the Lagrangian in the form

$$\mathcal{L} = E \left\{ \sum_{t=0}^{\infty} \delta^t \left[L_t + (\xi'_{t+1}, \Xi'_t)(Ay_t + Bi_t + [u'_{t+1} \ 0']') - \delta^{-1}(\xi'_t, \Xi'_{t-1})\hat{I}y_t \right] + \delta^{-1}\xi'_0(\bar{X}_0 + u_0) \right\}, \quad (3.2)$$

where

$$y_t \equiv \begin{bmatrix} X_t \\ x_t \end{bmatrix},$$

and

$$\hat{I} \equiv \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix}.$$

We have added a term $-\delta^{-1}\Xi'_{-1}Qx_0$ to the right-hand side, for the sake of symmetry in notation, but now correspondingly stipulate the initial condition

$$\Xi_{-1} = 0. \quad (3.3)$$

(Note that these Lagrange multipliers do not correspond to any actual constraint upon the planning problem.) Finally, note that equations (2.2)–(2.3) define a quadratic function $L_t = L(y_t, y_{t|t}, i_t)$. Because $y_{t|t} = y_t$ in the case of full information, we can here write $L_t = L(y_t, y_t, i_t)$. Thus the Lagrangian (3.2) is a quadratic function of the evolution of the vectors y_t and i_t .

Differentiation of (3.2) with respect to y_t and i_t then yields the first-order conditions

$$L_{y_t} + E_t(\xi'_{t+1}, \Xi'_t)A = \delta^{-1}(\xi'_t, \Xi'_{t-1})\hat{I}, \quad (3.4)$$

² More precisely, $E[\Xi_t E_t x_{t+1}] = EE_t[\Xi_t x_{t+1}] = E[\Xi_t x_{t+1}]$ (where the first equality follows since Ξ_t is measurable with respect to I_t^f).

$$L_{it} + \mathbf{E}_t(\xi'_{t+1}, \Xi'_t)B = 0, \quad (3.5)$$

respectively, where for each of the two arguments $z = y, i$, $L_{zt} \equiv \partial L(y_t, y_t, i_t)/\partial z_t$. Recalling the form of the quadratic function L , we have

$$L(y_t, y_t, i_t) = \begin{bmatrix} y'_t & i'_t \end{bmatrix} \begin{bmatrix} C' \\ C'_i \end{bmatrix} W \begin{bmatrix} C & C_i \end{bmatrix} \begin{bmatrix} y_t \\ i_t \end{bmatrix} \equiv \frac{1}{2} \begin{bmatrix} y'_t & i'_t \end{bmatrix} \begin{bmatrix} L_{yy} & L_{yi} \\ L_{iy} & L_{ii} \end{bmatrix} \begin{bmatrix} y_t \\ i_t \end{bmatrix}, \quad (3.6)$$

so that we can write

$$\begin{bmatrix} L'_{yt} \\ L'_{it} \end{bmatrix} = \begin{bmatrix} L_{yy} & L_{yi} \\ L_{iy} & L_{ii} \end{bmatrix} \begin{bmatrix} y_t \\ i_t \end{bmatrix},$$

where the L_{jk} are matrices of constant coefficients (corresponding to the second partial derivatives of L), that depend only upon the matrices C, C_i , and W as above. Using this notation, we can equivalently write the first-order conditions (3.4) – (3.5) as

$$L_{yy}y_t + L_{yi}i_t + A'\mathbf{E}_t \begin{bmatrix} \xi_{t+1} \\ \Xi_t \end{bmatrix} = \delta^{-1}\hat{I}' \begin{bmatrix} \xi_t \\ \Xi_{t-1} \end{bmatrix}, \quad (3.7)$$

$$L_{iy}y_t + L_{ii}i_t + B'\mathbf{E}_t \begin{bmatrix} \xi_{t+1} \\ \Xi_t \end{bmatrix} = 0. \quad (3.8)$$

Assuming that L_{ii} is of full rank (see Svensson and Woodford [10, appendix B] for a the case when L_{ii} is not of full rank), we can solve (3.8) for i_t , obtaining

$$i_t = -L_{ii}^{-1}L_{iy}y_t - L_{ii}^{-1}B'\mathbf{E}_t \begin{bmatrix} \xi_{t+1} \\ \Xi_t \end{bmatrix}. \quad (3.9)$$

Substituting (3.9) into (2.1) and (3.7) to eliminate i_t , we then obtain a system of equations for the evolution of y_t and $(\xi'_{t+1}, \Xi'_t)'$, that can be written in the form

$$\begin{bmatrix} 0 & R' \\ \hat{I} & U \end{bmatrix} \begin{bmatrix} \mathbf{E}_t y_{t+1} \\ \mathbf{E}_t \begin{bmatrix} \xi_{t+1} \\ \Xi_t \end{bmatrix} \end{bmatrix} = \begin{bmatrix} V & \delta^{-1}\hat{I}' \\ R & 0 \end{bmatrix} \begin{bmatrix} y_t \\ \begin{bmatrix} \xi_t \\ \Xi_{t-1} \end{bmatrix} \end{bmatrix}, \quad (3.10)$$

where

$$\begin{aligned} R &\equiv A - BL_{ii}^{-1}L_{iy}, \\ U &\equiv BL_{ii}^{-1}B', \\ V &\equiv -L_{yy} + L_{yi}L_{ii}^{-1}L_{iy}. \end{aligned}$$

Here it is worth noting that U and V are symmetric matrices.

Let us assume furthermore that the square matrix on the left-hand side of (3.10) is of full rank.³ Then we can invert this matrix, to obtain a system of the form

$$\begin{bmatrix} E_t y_{t+1} \\ E_t \xi_{t+1} \\ \Xi_t \end{bmatrix} = \mathcal{M} \begin{bmatrix} y_t \\ \xi_t \\ \Xi_{t-1} \end{bmatrix}. \quad (3.11)$$

We then wish to consider solutions to (3.11) that are consistent with given initial values for X_0 and Ξ_{-1} according to (3.1) and (3.3). We note that the number of variables in (3.11) is $2(n_X + n_x)$, where n_X and n_x is the dimension of X_t and x_t , respectively, and that there are $n_X + n_x$ initial conditions ((3.1) and (3.3)). We shall restrict our attention to bounded solutions, by which we mean solutions in which for any t , $E_t y_{t+\tau}$, $E_t \xi_{t+\tau}$ and $E_t \Xi_{t+\tau-1}$ satisfy a uniform bound for all τ . Such solutions necessarily satisfy the transversality condition for an optimal plan, and since our equations (2.1)–(2.4) will usually represent only a local approximation to the true structural equations and true loss function, unbounded solutions need not correspond at all closely to solutions to the true equations.

As usual (and ignoring non-generic cases), there is a unique bounded solution to (3.11) consistent with the initial conditions if the number of eigenvalues of \mathcal{M} inside the unit circle (that is, with modulus less than one) is exactly equal to the number of initial conditions, $n_X + n_x$. The eigenvalues λ of \mathcal{M} are the roots of the characteristic equation

$$\text{Det} \begin{bmatrix} V & \delta^{-1} \hat{I}' - \lambda R' \\ R - \lambda \hat{I} & -\lambda U \end{bmatrix} = 0. \quad (3.12)$$

Multiplication of the right blocks of this matrix by $-\lambda^{-1}$, then multiplication of the lower blocks by $-\lambda^{-1} \delta^{-1}$, and finally transposition of the matrix does not change the sign of its determinant. Thus we may equivalently write

$$\text{Det} \begin{bmatrix} V & \delta^{-1} \hat{I}' - \lambda^{-1} \delta^{-1} R' \\ R - \lambda^{-1} \delta^{-1} \hat{I} & -\lambda^{-1} \delta^{-1} U \end{bmatrix} = 0.$$

Comparison of this with (3.12) shows that if λ is a root, $\lambda^{-1} \delta^{-1}$ must also be. It follows that \mathcal{M} has as many eigenvalues with $|\lambda| > \frac{1}{\sqrt{\delta}}$ as with $|\lambda| < \frac{1}{\sqrt{\delta}}$. Thus, since $\frac{1}{\sqrt{\delta}} > 1$, *at most* half of the

³ Even when Q is singular, so that this matrix also is, our conclusions below remain essentially valid. Equation (3.12) is still the relevant characteristic equation, and again there is a unique bounded solution, in the generic case, if and only if exactly $n_X + n_x$ roots are inside the unit circle. Furthermore, it is again true that there are necessarily no more than this number of such roots, and that for δ close enough to 1, the condition assumed here almost inevitably holds. However, we omit the algebra for the more general case.

eigenvalues (that is, at most $n_X + n_x$) are inside the unit circle (that is, with $|\lambda| < 1$), so there is no possibility of multiple stationary solutions to (3.11). If δ is close to 1 (as will often be the case), there are likely to be exactly half inside the unit circle. We shall assume this condition from now on. Thus (3.11) has a unique bounded solution in which $E_t y_{t+\tau}$ and $E_t(\xi'_{t+1}, \Xi'_t)'$ can be expressed as linear functions of the initial conditions X_0 and (3.3), for arbitrary $\tau \geq 0$.

In particular, the optimal equilibrium involves evolution of the instrument according to a relation of the form

$$i_t = FX_t + \Phi \Xi_{t-1}, \quad (3.13)$$

the *optimal reaction function in state-space form*, where F and Φ are matrices of constant coefficients. We have just argued that y_0 and $E_0 \xi_1$ can be expressed as linear functions of X_0 and $\Xi_{-1} \equiv 0$; substitutions of these solutions into (3.9) then yields (3.13) for $t = 0$. However, exactly the same reasoning can be applied to solve equations (3.11) for all $\tau \geq t$, given initial values X_t and Ξ_{t-1} , and the unique bounded solution will be linear in the initial values, with exactly the same coefficients in period $t = 0$. Thus (3.13) must hold for all t .

Similarly, the forward-looking variables evolve according to a relation of the form

$$x_t = GX_t + \Gamma \Xi_{t-1}, \quad (3.14)$$

while the Lagrange multipliers associated with the forward-looking variables evolve according to

$$\Xi_t = SX_t + \Sigma \Xi_{t-1}, \quad (3.15)$$

starting from the initial condition (3.3). Substitution of these equations into (2.1) then implies that the predetermined state variables evolve according to

$$X_{t+1} = (A_{11} + A_{12}G + B_1F)X_t + (A_{12}\Gamma + B_1\Phi)\Xi_{t-1} + u_{t+1}, \quad (3.16)$$

starting from the initial condition X_0 and (3.3). Note that (3.15) can be integrated to yield

$$\Xi_t = \sum_{\tau=0}^t \Sigma^\tau S X_{t-\tau}.$$

Thus, we have

$$i_t = \sum_{\tau=0}^t F_\tau X_{t-\tau}, \quad (3.17)$$

the *optimal reaction function in integrative form*, where $F_0 \equiv F$ and $F_\tau \equiv \Phi \Sigma^{\tau-1} S$, $\tau \geq 1$. Thus the most fundamental difference with respect to the discretion case is that under the optimal

commitment, i_t (and x_t) are no longer a linear function of the current state X_t alone, but instead depends upon past state vectors $X_{t-\tau}$ as well. The inertial character of optimal policy that this can result in is illustrated in Woodford [11].

Equations (3.13)–(3.16) then completely describe the evolution of the predetermined variables, the forward-looking variables, and the policy settings i_t , as a function of the sequence of realizations of the disturbances u_t (and the initial conditions (3.1) and (3.3)). Note that (3.15) implies that the Lagrange multipliers Ξ_t are predetermined variables.

Note also that the matrices $F, G, S, \Phi, \Gamma, \Sigma$ depend on A, B, Q, C, C_i, W and δ , but that they are independent of Σ_{uu} . Thus these coefficients are the same as in the optimal plan under *certainty*. This is the *certainty equivalence* result for the case of full information.

4 Certainty equivalence under partial information

Now suppose instead that both the private sector and the central bank observe only the variables Z_t in period t , that is, have the information I_t rather than I_t^f . In this case the Lagrangian takes the form

$$\begin{aligned} \mathcal{L} = \text{E} & \left[\sum_{t=0}^{\infty} \delta^t L_t - \sum_{t=0}^{\infty} \delta^t \xi'_{t+1} \left(X_{t+1} - A_{11}^1 X_t - A_{11}^2 X_{t|t} - A_{12}^1 x_t - A_{12}^2 x_{t|t} - B_1 i_t - u_{t+1} \right) \right. \\ & \left. - \sum_{t=0}^{\infty} \delta^t \psi'_t (x_{t+1|t} - A_{21}^1 X_t - A_{21}^2 X_{t|t} - A_{22}^1 x_t - A_{22}^2 x_{t|t} - B_2 i_t) - \delta^{-1} \xi'_0 (X_0 - \bar{X}_0 - u_0) \right]. \end{aligned}$$

We now distinguish between $z_{t+1|t} \equiv \text{E}[z_{t+1} | I_t]$ and $\text{E}_t z_{t+1} \equiv \text{E}[z_{t+1} | I_t^f]$, the expectation (of any variable z_{t+1}) conditional upon all X_τ, x_τ and Z_τ for all $\tau \leq t$. Note also that now the distinction between the two components A^1 and A^2 is relevant for the problem's constraints. Now the multipliers ψ_t are measurable with respect to the full period t information, I_t^f , as the term in brackets represents a constraint that applies in period t . However, they are *not* necessarily measurable with respect to I_t (that is, they do not necessarily depend only on I_t) as there is *not* a single constraint for each information I_t . Thus, $\psi_t \neq \psi_{t|t}$.

Note also that we do not write explicitly, in the Lagrangian, the constraints indicating the way in which the choice of stochastic processes for X_t, x_t , and i_t affects the conditioning information in the various conditional expectations $z_{\tau|t}$, as a result of (2.4). As shown in Appendix A, we obtain the correct first-order conditions even when the additional constraints (discussed there) are omitted.

This expression can be simplified, if we note that

$$\mathbb{E} \left[\psi'_t x_{t+1|t} \right] = \mathbb{E} \left\{ \mathbb{E} \left[\psi'_t x_{t+1|t} \mid I_t \right] \right\} = \mathbb{E} \left[\psi'_{t|t} x_{t+1|t} \right] = \mathbb{E} \left\{ \mathbb{E} \left[\psi'_{t|t} x_{t+1} \mid I_t \right] \right\} = \mathbb{E} [\Xi_t x_{t+1}],$$

using the law of iterated expectations and introducing $\Xi_t \equiv \psi_{t|t}$. We can then equivalently express the Lagrangian as

$$\mathcal{L} = \mathbb{E} \left\{ \sum_{t=0}^{\infty} \delta^t \left[L_t + (\xi'_{t+1}, \Xi'_t)(A^1 y_t + B i_t) + (\xi'_{t+1|t}, \Xi'_t) A^2 y_t - \delta^{-1} (\xi'_t, \Xi'_{t-1}) \hat{I} y_t \right] + \delta^{-1} \xi'_0 (\bar{X}_0 + u_0) \right\}, \quad (4.1)$$

once again stipulating the initial condition (3.3). It should also be noted that now i_t must be measurable with respect to the information I_t . Note also that $\Xi_t \equiv \psi_{t|t}$ is measurable with respect to I_t , even though ψ_t is not.

Differentiation of (4.1) then yields the first-order conditions

$$L_{1t} + L_{2t|t} + \mathbb{E}_t(\xi'_{t+1}, \Xi'_t) A^1 + (\xi'_{t+1|t}, \Xi'_t) A^2 = \delta^{-1} (\xi'_t, \Xi'_{t-1}) \hat{I}, \quad (4.2)$$

$$L_{it|t} + (\xi'_{t+1|t}, \Xi'_t) B = 0, \quad (4.3)$$

where for $j = 1, 2, i$, L_{jt} denotes the partial derivative of $L(y_t, y_{t|t}, i_t)$ with respect to its first, second, or third argument respectively. Here we have used result (A.14) from appendix A to differentiate functions of $y_{t|t}$ with respect to y_t . More precisely, condition (4.2) should also contain a term that is proportional to μ'_t , the vector of Lagrange multipliers associated with the constraint on how changes in the evolution of y_t affect the information content of the observables Z_τ in (2.4), as shown in appendix A. However, as is established there in (A.19), $\mu_{t|t} = 0$, as a result of which the neglected term has no consequences for condition (4.4) below, which is all that matters for our subsequent analysis.

Note that the expectations in (4.2) and (4.3) are *not* conditioned upon the same information, because y_t may take a different value in each state of the world in period t , while i_t must have the same value in each state of the world that corresponds to the same information I_t . (The consequences of the latter constraint are also treated in the Appendix.) Finally, note that in the case of full information, conditions (4.2)–(4.3) are identical to (3.7)–(3.8).

As it turns out, only the conditional expectations of these first-order conditions with respect to public information I_t matter for determination of the optimal evolution of y_t and i_t . Each term in (4.3) is already conditional upon I_t . However, taking the conditional expectation of (4.2) with respect to I_t , we obtain the simpler expression

$$L_{1t|t} + L_{2t|t} + (\xi'_{t+1|t}, \Xi'_t) A = \delta^{-1} (\xi'_t, \Xi'_{t-1}) \hat{I}. \quad (4.4)$$

Furthermore, a calculation similar to (3.6) implies that

$$\begin{bmatrix} L'_{1t|t} + L'_{2t|t} \\ L'_{it|t} \end{bmatrix} = \begin{bmatrix} L_{yy} & L_{yi} \\ L_{iy} & L_{ii} \end{bmatrix} \begin{bmatrix} y_{t|t} \\ i_{t|t} \end{bmatrix},$$

where the matrices L_{jk} are exactly the same as in (3.6). Thus conditions (4.3) and (4.4) can alternatively be written

$$L_{iy}y_{t|t} + L_{ii}i_{t|t} + B' \begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = 0, \quad (4.5)$$

$$L_{yy}y_{t|t} + L_{yi}i_{t|t} + A' \begin{bmatrix} \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = \delta^{-1} \hat{I}' \begin{bmatrix} \xi_{t|t} \\ \Xi_{t-1} \end{bmatrix}. \quad (4.6)$$

The pair of difference equations (4.5) and (4.6) will be observed to be of exactly the same form as (3.7) and (3.8) in the full-information case, except that conditional expectations are now with respect to I_t rather than I_t^f . Thus, exactly in the same way as above, we can obtain a system of equations

$$\begin{bmatrix} y_{t+1|t} \\ \xi_{t+1|t} \\ \Xi_t \end{bmatrix} = \mathcal{M} \begin{bmatrix} y_{t|t} \\ \xi_{t|t} \\ \Xi_{t-1} \end{bmatrix}, \quad (4.7)$$

where the matrix \mathcal{M} is the same as in (3.11).

As above (under our assumption about the eigenvalues of \mathcal{M}), there is a unique bounded solution for $y_{\tau|t}$, $\xi_{\tau|t}$ and $\Xi_{\tau-1|t}$ ($\tau \geq t$), given any initial values $X_{t|t}$ and Ξ_{t-1} . Solving (4.5) for i_t , we can associate with any such solution to (4.7) a unique bounded solution for $i_{\tau|t}$ as well. This solution satisfies

$$i_t = FX_{t|t} + \Phi\Xi_{t-1}, \quad (4.8)$$

$$x_{t|t} = GX_{t|t} + \Gamma\Xi_{t-1}, \quad (4.9)$$

$$\Xi_t = SX_{t|t} + \Sigma\Xi_{t-1}, \quad (4.10)$$

where the matrices $F, G, S, \Phi, \Gamma, \Sigma$ are the same as in (3.13)–(3.15) for the full-information case. Equation (4.10) can be integrated to yield Ξ_t as a distributed lag of past values of $X_{t-\tau|t-\tau}$; thus, using this with (4.8), the optimal reaction function in state-space form, results in

$$i_t = \sum_{\tau=0}^t F_{\tau} X_{t-\tau|t-\tau}, \quad (4.11)$$

the optimal reaction function in integrated form, where F_τ , $\tau \geq 0$, are the same as for the full-information case. Again, a certain amount of inertia is introduced into the way that $X_{t|t}$ determines i_t (and x_t).

Note that equations (4.8)–(4.10) and (4.11) take exactly the same form as (3.13)–(3.15) and (3.17), once expectations conditional upon I_t^f are replaced by expectations conditional upon I_t . This represents an extension of certainty equivalence to the partial-information case.

5 The separation principle

The second row of (2.1) implies that

$$A_{21}^1(X_t - X_{t|t}) + A_{22}^1(x_t - x_{t|t}) = 0. \quad (5.1)$$

Assuming that A_{22}^1 is non-singular, this can be solved for x_t . Substituting (4.9) for $x_{t|t}$, one obtains

$$x_t = G^1 X_t + G^2 X_{t|t} + \Gamma \Xi_{t-1}, \quad (5.2)$$

where again

$$G^1 \equiv -(A_{22}^1)^{-1} A_{21}^1,$$

$$G^2 \equiv G - G^1.$$

Note that the matrices G^1 and G^2 , like the others, are independent of the specifications of D , Σ_{uu} and Σ_{vv} .

Substitution of (4.8), (4.9) and (5.2) into the first row of (2.1) furthermore yields

$$X_{t+1} = H X_t + J X_{t|t} + \Psi \Xi_{t-1} + u_{t+1}, \quad (5.3)$$

where H and J are given by

$$H \equiv A_{11}^1 + A_{12}^1 G^1, \quad (5.4)$$

$$J \equiv B_1 F + A_{12}^1 G^2 + A_{11}^2 + A_{12}^2 G, \quad (5.5)$$

and

$$\Psi \equiv A_{12} \Gamma + B_1 \Phi. \quad (5.6)$$

Equations (4.10) and (5.2)–(5.3) then describe the evolution of the state variables y_t in equilibrium, once we determine the evolution of the estimates $X_{t|t}$ of the predetermined variables.

Substituting (5.2) into (2.4), we obtain

$$Z_t = LX_t + MX_{t|t} + \Lambda \Xi_{t-1} + v_t, \quad (5.7)$$

where L and M are given by

$$L \equiv D_1^1 + D_2^1 G^1, \quad (5.8)$$

$$M \equiv D_2^1 G^2 + D_1^2 + D_2^2 G, \quad (5.9)$$

and

$$\Lambda \equiv D_2 \Gamma. \quad (5.10)$$

Equations (5.3) and (5.7) are then the transition and measurement equations for an optimal filtering problem. Again the transformation into a problem without forward-looking variables allows us to derive the estimation equations in a manner that is simpler than that used in Pearlman, Currie and Levine [8].

As demonstrated in more detail in Svensson and Woodford [9], the optimal prediction of $X_{t|t}$ is then given by a Kalman filter,

$$X_{t|t} = X_{t|t-1} + K(Z_t - LX_{t|t-1} - MX_{t|t} - \Lambda \Xi_{t-1}), \quad (5.11)$$

We can rationalize (5.11) by observing that $Z_t - MX_{t|t} - \Lambda \Xi_{t-1} = LX_t + v_t$, hence,

$$Z_t - LX_{t|t-1} - MX_{t|t} = L(X_t - X_{t|t-1}) + v_t,$$

so (5.11) can be written in the conventional form

$$X_{t|t} = X_{t|t-1} + K[L(X_t - X_{t|t-1}) + v_t], \quad (5.12)$$

which allows us to identify K as (one form of) the Kalman gain matrix. From (5.3) we get

$$X_{t+1|t} = (H + J)X_{t|t} + \Psi \Xi_{t-1}, \quad (5.13)$$

and the dynamics of the model are given by (4.10), (5.2), (5.3), (5.11) and (5.13).

The Kalman gain matrix is given by

$$K = PL'(LPL' + \Sigma_{vv})^{-1}, \quad (5.14)$$

where the matrix $P \equiv \text{Cov}[X_t - X_{t|t-1}]$ is the covariance matrix for the prediction errors $X_t - X_{t|t-1}$ and fulfills

$$P = H[P - PL'(LPL' + \Sigma_{vv})^{-1}LP]H' + \Sigma_{uu}. \quad (5.15)$$

Thus P can be solved from (5.15), either numerically or analytically, depending upon the complexity of the matrices H , L and Σ_{uu} . Then K is given by (5.14). Note that this implies that the Kalman gain K depends only upon the matrices A , Σ_{uu} , D , and Σ_{vv} .

6 Conclusions

The above proof demonstrates the certainty-equivalence result that the optimal policy under commitment given an estimate of the state of the economy is independent of the degree of uncertainty and hence the same policy as under full information. Furthermore, a separation principle holds, in that the problem of finding the optimal policy and the problem of optimally estimating the current state of the economy can be treated as separate problems (in particular, the optimal estimation does not depend on the loss function or the reaction function).

These results hold under symmetric partial information about the economy. As demonstrated in Svensson and Woodford [10], in the asymmetric case in which the policymaker has partial information and the private sector has full information, the certainty-equivalence result holds only for the reaction function in state-space form, (4.8), but not for the reaction function in integrative form, (4.11). Furthermore, the separation principle does not hold, since the optimal estimation is no longer independent of the loss-function parameters and the reaction function.

A Differentiation results for conditional expectations

Here we address some technical issues that arise in the characterization of the optimal commitment problem in the case of partial information. These relate to the fact that the policymaker should recognize that his or her pattern of action affects the statistical relation between the observables and underlying (exogenous) shocks, and thus might affect the information that is publicly available about those shocks. This problem can be ignored in the case of discretion because an independent optimization problem is solved in each state, and we may suppose that behavior in any single state is of only negligible importance for the correlations that determine the optimal linear estimates of unobserved states. But we must consider the matter more carefully in the case of commitment. In fact, we show here that the results presented in the text, derived without taking into account the effects of policy upon the content of public information, continue to be correct.

We wish to consider the problem of minimizing a discounted sum of expected losses

$$\mathbb{E} \sum_{t=0}^{\infty} \delta^t L(y_t, y_{t|t}, i_t) \quad (\text{A.1})$$

subject to a series of constraints of the form

$$M_0 y_{t+1} + M_1 y_{t+1|t} = M_2 y_t + M_3 y_{t|t} + M_4 i_t + u_{t+1} \quad (\text{A.2})$$

for each period $t \geq 0$, where M_0, \dots, M_4 are matrices of appropriate dimension. Here i_t is a vector of control variables chosen in period t , y_t is a vector of endogenous state variables determined by structural equations (A.2), and u_{t+1} of random disturbances in period $t + 1$, assumed independently distributed over time. (For present purposes it is not necessary to distinguish between the predetermined and non-predetermined elements of y_t .) There is also a set of initial conditions specifying

$$M_0 y_0 = u_0, \quad (\text{A.3})$$

where u_0 is another random vector, distributed independently of the u_{t+1} vectors for all later periods.

For any random variable z_T , the conditional expectation $z_{T|t}$ denotes

$$z_{T|t} \equiv \mathbb{E}[z_T | Z_t, Z_{t-1}, \dots, Z_0], \quad (\text{A.4})$$

where Z_t is a vector of observables in period t , implicitly defined by

$$Z_t = D^1 y_t + D^2 y_{t|t} + v_t, \quad (\text{A.5})$$

where v_t is a vector of additional iid random disturbances. We introduce the notation $\bar{u}_t \equiv (u'_t, v'_t)'$ for the complete vector of exogenous random disturbances in period t . The controls must be chosen on the basis of public information, so that

$$i_t = i_{t|t} \quad (\text{A.6})$$

is also a constraint.

We may incorporate constraint (A.6) by instead writing the objective (A.1) as

$$\mathbb{E} \sum_{t=0}^{\infty} \delta^t L(y_t, y_{t|t}, i_{t|t}), \quad (\text{A.7})$$

and making a similar substitution into (A.2). Then substituting (A.4) into (A.7) for the conditional expectations, we obtain an objective that depends upon the state-contingent paths of

the variables y_t, i_t and Z_t . Our problem is then to choose y_t, i_t and, in particular, Z_t in each period $t \geq 0$, as functions of the history of realizations, the state $s_t \equiv (\bar{u}'_t, \bar{u}'_{t-1}, \dots, \bar{u}'_0)'$, subject to constraints (A.2), (A.3), and (A.5), so as to minimize (A.7). This will only determine the path of $i_{t|t}$, but by then imposing (A.6) as well we can determine the complete evolution of the control i_t .

The first-order conditions for such a problem can, as usual, be obtained by writing a Lagrangian

$$\begin{aligned} \mathcal{L} = & \sum_{t=0}^{\infty} \delta^t \mathbf{E} L(y_t, y_{t|t}, i_{t|t}) \\ & - \sum_{t=0}^{\infty} \delta^t \mathbf{E} \varphi'_{t+1} (M_0 y_{t+1} + M_1 y_{t+1|t} - M_2 y_t - M_3 y_{t|t} - M_4 i_{t|t} - u_{t+1}) \end{aligned} \quad (\text{A.8})$$

$$- \delta^{-1} \mathbf{E} \varphi'_0 (M_0 y_0 - u_0) - \sum_{t=0}^{\infty} \delta^t \mathbf{E} \mu'_t (Z_t - D^1 y_t - D^2 y_{t|t} - v_t), \quad (\text{A.9})$$

where φ_{t+1} , φ_0 , and μ_t are the Lagrange multipliers associated with constraints (A.2), (A.3), and (A.5) respectively. (Thus, φ_t corresponds to (ξ'_t, ψ'_{t-1}) in the text. There is a separate vector of multipliers φ_{t+1} for each possible state $s_{t+1} \equiv (\bar{u}'_{t+1}, \bar{u}'_t, \dots, \bar{u}'_0)'$, a separate φ_0 for each possible $s_0 \equiv \bar{u}_0$, and a separate μ_t for each possible state s_t . The expectation operator $\mathbf{E}[\cdot]$ indicates unconditional expectations (expectations in the *ex ante* state in which the optimal commitment is chosen). The commitment is chosen prior to the realization of any period 0 states, as the systematic pattern of behavior committed to in period 0 affects the information revealed in alternative states in that period, as in others. The aim of this appendix is to explain the calculation of the partial derivatives of such a Lagrangian with respect to the random paths specified for y_t, i_t and Z_t .

A.1 Properties of the conditional expectation $z_{T|t}$

We first note, that the conditional expectation (A.4) is a linear operator of the form

$$z_{T|t}(s_t) \equiv \mathbf{E}_{s_T} [P_{t,T}(s_t, s_T) z_T(s_T)], \quad (\text{A.10})$$

where $P_{t,T}(s_t, s_T)$ is a (scalar) kernel. Here $s_T \equiv (\bar{u}'_T, \dots, \bar{u}'_0)'$ is an arbitrary state in period T , s_t is an arbitrary state in period $t \leq T$, and \mathbf{E}_{s_T} denotes the expectation over the possible states s_T , under the *ex ante* or unconditional probability measure. (Note that $\mathbf{E}_{s_T}[\cdot]$ should not be confused with $\mathbf{E}_t[\cdot] \equiv \mathbf{E}[\cdot | I_t^f]$, the expectation conditional on the full information in period t , I_t^f .) Under our assumption of normally distributed disturbances, the conditional expectation

is simply a linear projection upon the observables. This means that the stochastic kernel takes the form

$$P_{t,T}(s_t, s_T) = P_{t,t}(s_t, s_{T,t}),$$

where $s_{T,t}$ is the *predecessor* of state s_T in period t (that is, the part of the history that has occurred as of period t , that is, the unique state $s_{T,t}$ in period t that fulfills $s_T \equiv (\bar{u}'_T, \dots, \bar{u}'_{t+1}, s'_{T,t})'$). Furthermore, $P_{t,t}$ can be expressed in the standard form for linear projections,

$$P_{t,t}(s_t^1, s_t^2) = \hat{Z}'_t(s_t^2) E[\hat{Z}_t \hat{Z}'_t]^{-1} \hat{Z}_t(s_t^1), \quad (\text{A.11})$$

where s_t^1 and s_t^2 are two states in period t and $\hat{Z}_t(s_t)$ is a vector of observables in period t and earlier (including a constant) that results under s_t and spans the entire public information space in period t , but that includes no redundant variables (so that the matrix $E[\hat{Z}_t \hat{Z}'_t]$ is non-singular). (Note that if s_t^1 and s_t^2 correspond to the same information I_t , we have $\hat{Z}_t(s_t^1) = \hat{Z}_t(s_t^2)$. Also, if there are no redundant observables, $\hat{Z}_t = (Z'_t, Z'_{t-1}, \dots, Z'_0)'$.)

The choice of a particular representation \hat{Z}_t need not concern us here. We do, however, wish to observe certain consequences of the general form (A.11) for the stochastic kernel. One is that the kernel is symmetric, that is,

$$P_{t,t}(s_t^1, s_t^2) = P_{t,t}(s_t^2, s_t^1). \quad (\text{A.12})$$

Another is that, for each fixed state s_t^2 , the random variable $P_{t,t}(\cdot, s_t^2)$ is a linear combination of the vector of observables as of period t , so that $P_{t,t}(\cdot, s_t^2)$ is in the linear space spanned by the vector of observables in period t . More generally, for any period $T \geq t$ and any state s_T in period T , $P_{t,T}(\cdot, s_T)$ is a random variable that is measurable with respect to the information I_t . These properties of the kernel suffice for the calculations that we need to do here.

A.2 Differentiation of expressions involving conditional expectations

We now consider differentiation of expressions involving conditional expectations, of the sort that appear in the Lagrangian (A.9). We first consider the effect upon an expression of the form $z_{T|t}$ of varying the state-contingent path of random variable z_T , holding fixed the state-contingent paths of the variables (Z_t, \dots, Z_0) , and hence keeping the stochastic kernel $P_{t,T}$ unchanged. Suppose we let

$$z_T = \bar{z}_T + \phi \nu_T, \quad (\text{A.13})$$

where \bar{z}_T is the value of the random variable at which we wish to evaluate the partial derivative, ϕ is a scalar constant, and ν_T is another random variable (that is, another function of the state in period T , s_T) indicating the direction in which we wish to vary z_T . Then we define the partial derivative of a functional $g(z_T)$ with respect to perturbations of the random variable z_T as the function $[\partial g(z_T)/\partial z_T](s_T)$ with the property that

$$\frac{\partial g(\bar{z}_T)}{\partial \phi} = \mathbb{E}_{s_T} \left[\frac{\partial g(\bar{z}_T)}{\partial z_T}(s_T) \nu_T(s_T) \right]$$

in the case of any perturbation of the form (A.13), that is, for any random variable ν_T in period T .

From (A.10), we then can observe that

$$\frac{\partial z_{T|t}(s_t)}{\partial z_T}(s_T) = P_{t,T}(s_t, s_T), \quad (\text{A.14})$$

where s_t is any state in period $t \leq T$. By substituting (A.13) into (A.10), differentiating with respect to ϕ and using (A.14), it follows that, for any process ν_T ,

$$\mathbb{E}_{s_T} \left[\frac{\partial z_{T|t}(s_t)}{\partial z_T}(s_T) \nu_T(s_T) \right] = \mathbb{E}_{s_T} [P_{t,T}(s_t, s_T) \nu_T(s_T)] = \nu_{T|t}(s_t). \quad (\text{A.15})$$

Note that (A.15) is just the conditional expectation with respect to I_t of the expression that would be obtained if one were differentiating z_T instead of $z_{T|t}$. This is intuitive since, by the law of iterated expectations,

$$\mathbb{E}[z_{T|t} \nu_T] = \mathbb{E}[z_{T|t} \nu_{T|t}] = \mathbb{E}[z_T \nu_{T|t}].$$

Thus the partial derivatives of each of these equivalent expressions with respect to z_T should be the same; but the partial derivative of the final expression is obviously $\nu_{T|t}$.

A.3 The effect of variation in the dependence of Z_t on the state s_t

We turn next to the effect upon conditional expectations of variation in the way that the variables Z_t depend upon the state $s_t \equiv (\bar{u}'_t, \dots, \bar{u}'_0)'$. We consider the effect of variation in a particular (scalar) random observable Z_{jt} (for a given j , $1 \leq j \leq n_Z$) in the information I_τ for some $\tau \geq t$. By analogy with (A.13), we consider perturbations of the form

$$Z_{jt} = \bar{Z}_{jt} + \phi \zeta_t$$

around the random observable \bar{Z}_{jt} for any random variable ζ_t . We wish to consider the effect of variation in ϕ on a conditional expectation $z_{T|\tau}$. We shall assume that our baseline pattern of

variation in the observables \bar{Z}_t is such that no small perturbation changes the *dimension* of the linear space spanned by the observables; that is, we assume that we are not in a degenerate case in which \bar{Z}_{jt} is an exact linear combination of other observables, but would cease to be under an infinitesimal perturbation. This is necessary in order for a partial derivative of a conditional expectation with respect to Z_{jt} to exist. Our assumption means that the first-order conditions that we derive here are necessary conditions for an optimal commitment *under the assumption* that optimal policy does not involve a degeneracy of this kind. (It does not seem to us likely that it should, in general; but here we must simply note that our methods apply only to the case in which it does not.)

Now, under the above observation, one observes that for any ζ_t in the linear space spanned by the observables \bar{Z}_t in period t , the linear space spanned by \bar{Z}_{lt} , $l \neq j$ and $Z_{jt} = \bar{Z}_{jt} + \phi\zeta_t$ coincides with the linear space spanned by \bar{Z}_t , regardless of the value of ϕ . Thus, in such a case,

$$\frac{\partial z_{T|\tau}(s_\tau)}{\partial \phi} = E_{\tilde{s}_t} \left[\frac{\partial z_{T|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t)\zeta_t(\tilde{s}_t) \right] = 0.$$

This is true in particular for ζ_t equal to the kernel $P_{t,t}(s_t, \tilde{s}_t)$, where s_t is a fixed arbitrary state in period t , so that

$$E_{\tilde{s}_t} \left[\frac{\partial z_{T|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t)P_{t,t}(s_t, \tilde{s}_t) \right] = 0. \quad (\text{A.16})$$

This property of the partial derivative with respect to Z_{jt} suffices for our purposes. It means that as long as we are only interested in the expectation of our first-order conditions conditional upon public information, we can ignore the effects upon conditional expectations of variations in the way that observables depend upon the history of realizations. This in turn makes constraint (A.5) irrelevant to the optimization problem. We now show this explicitly by turning to the first-order conditions associated with our problem.

Differentiating the Lagrangian (A.9) with respect to $i_t(s_t)$, and using (A.14) together with (A.12), we obtain the first-order condition

$$E_{\tilde{s}_t} P_{t,t}(s_t, \tilde{s}_t)[L_{it}(\tilde{s}_t) + \varphi'_{t+1}(\tilde{s}_t)M_4] = 0,$$

where $L_{it}(s_t) \equiv L_i[y_t(s_t), y_{t|t}(s_t), i_{t|t}(s_t)]$. This can equivalently be written

$$L_{it|t} + \varphi'_{t+1|t}M_4 = 0. \quad (\text{A.17})$$

Differentiating with respect to $y_t(s_t)$, we obtain the corresponding first-order condition

$$L_{1t} + L_{2t|t} + E_t \varphi'_{t+1}M_2 + \varphi'_{t+1|t}M_3 - \delta^{-1}\varphi'_t M_0 - \delta^{-1}\varphi'_{t|t-1}M_1 + \mu'_t D^1 + \mu'_{t|t}D^2 = 0. \quad (\text{A.18})$$

Condition (A.18) holds for any possible history in any period $t \geq 0$, but for $t = 0$ we set $\varphi_{0|-1} = 0$, so the term $\delta^{-1}\varphi'_{0|-1}M_1$ is deleted.

Finally, differentiating with respect to $Z_{jt}(\tilde{s}_t)$, we obtain,

$$\begin{aligned} \mu_{jt}(\tilde{s}_t) = & \sum_{\tau=t}^{\infty} \delta^{\tau-t} \mathbf{E}_{s_\tau} \left\{ [L_{2\tau}(s_\tau) + \mathbf{E}_\tau \varphi'_{\tau+1}(s_\tau) M_3 + \mu'_\tau(s_\tau) D^2] \frac{\partial y_{\tau|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t) \right. \\ & \left. + [L_{i\tau}(s_\tau) + \mathbf{E}_\tau \varphi'_{\tau+1}(s_\tau) M_4] \frac{\partial i_{\tau|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t) - \mathbf{E}_\tau \varphi'_{\tau+1}(s_\tau) M_1 \frac{\partial y_{\tau+1|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t) \right\}, \end{aligned}$$

where for each period $\tau \geq t$, s_τ indexes an arbitrary state in that period. But taking the conditional expectation of this with respect to I_t , we obtain, for a state s_t consistent with the information I_t ,

$$\begin{aligned} \mu_{jt|t}(s_t) = & \sum_{\tau=t}^{\infty} \delta^{\tau-t} \mathbf{E}_{s_\tau} \left\{ [L_{2\tau}(s_\tau) + \mathbf{E}_\tau \varphi'_{\tau+1}(s_\tau) M_3 + \mu'_\tau(s_\tau) D^2] \mathbf{E}_{\tilde{s}_t} \left[\frac{\partial y_{\tau|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t) P_{t,t}(s_t, \tilde{s}_t) \right] \right. \\ & + [L_{i\tau}(s_\tau) + \mathbf{E}_\tau \varphi'_{\tau+1}(s_\tau) M_4] \mathbf{E}_{\tilde{s}_t} \left[\frac{\partial i_{\tau|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t) P_{t,t}(s_t, \tilde{s}_t) \right] \\ & \left. - \mathbf{E}_\tau \varphi'_{\tau+1}(s_\tau) M_1 \mathbf{E}_{\tilde{s}_t} \left[\frac{\partial y_{\tau+1|\tau}(s_\tau)}{\partial Z_{jt}}(\tilde{s}_t) P_{t,t}(s_t, \tilde{s}_t) \right] \right\} \\ = & 0, \end{aligned} \tag{A.19}$$

where we use (A.16) to evaluate each term in large square brackets. Then the conditional expectation of (A.18) with respect to I_t is given simply by

$$L_{1t|t} + L_{2t|t} + \varphi'_{t+1|t}(M_2 + M_3) - \delta^{-1}\varphi'_{t|t}M_0 - \delta^{-1}\varphi'_{t|t-1}M_1 = 0, \tag{A.20}$$

for $t \geq 0$, using (A.19). (Here, we let $\varphi_{0|-1} = 0$.)

In fact, as discussed in the text (identifying φ'_t with $(\xi'_t, \psi'_{t-1})'$ and $\varphi'_{t|t}$ with $(\xi'_{t|t}, \Xi'_{t-1})'$), the first-order conditions (A.17) and (A.20) suffice, along with the structural equations (A.2), to completely determine the expected dynamics $y_{\tau|t}$, $\varphi_{\tau|t}$ for all $\tau \geq t$, given initial conditions for period t . These results suffice, in turn, to completely determine the optimal evolution of the state vector y_t and the optimal actions i_t , as a function of the history of realizations of the exogenous disturbances. Hence, only $\mu_{t|t}$ (which is always zero) matters for our purposes, rather than μ_t itself. Consequently, the relevant first-order conditions, (A.17) and (A.20), are the same as if we simply neglected the constraints implied by (A.5) in computing the optimal plan.

References

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

- [2] Chow, Gregory C. (1975), *Analysis and Control of Dynamic Economic Systems*, John Wiley & Sons, New York.
- [3] Currie, David, and Paul Levine (1993), *Rules, Reputation and Macroeconomic Policy Coordination*, Cambridge University Press, Cambridge.
- [4] Kalchbrenner, J.H., and Peter A. Tinsley (1975), "On the Use of Optimal Control in the Design of Monetary Policy," Special Studies Paper No. 76, Federal Reserve Board.
- [5] LeRoy, Stephen F., and Roger N. Waud (1977), "Applications of the Kalman Filter in Short-run Monetary Control," *International Economic Review* 18, 195–207
- [6] Pearlman, Joseph G. (1986), "Diverse Information and Rational Expectations Models," *Journal of Economic Dynamics and Control* 10, 333–338.
- [7] Pearlman, Joseph G. (1992), "Reputational and Nonreputational Policies under Partial Information," *Journal of Economic Dynamics and Control* 16, 339–357.
- [8] Pearlman, Joseph G., David Currie and Paul Levine (1986), "Rational Expectations Models with Partial Information," *Economic Modelling* 3, 90–105.
- [9] Svensson, Lars E.O., and Michael Woodford (2002a), "Indicator Variables for Optimal Policy," working paper.
- [10] Svensson, Lars E.O., and Michael Woodford (2002b), "Indicator Variables for Optimal Policy under Asymmetric Information," working paper.
- [11] Woodford, Michael, "Optimal Monetary Policy Inertia," NBER working paper no. 7261, July 1999.