SWAS208.tex

Indicator Variables for Optimal Policy under Asymmetric Information*†

Lars E.O. Svensson‡ and Michael Woodford§

First draft: January 1999
This version: August 2002

Abstract

The optimal weights on indicators in models with partial information about the state of the economy and forward-looking variables are derived and interpreted, both for equilibria under discretion and under commitment. (32)The private sector is assumed to have more information than the policymaker. *Certainty equivalence* holds for the optimal reaction function in state-space form, but not for the reaction function in integrative form. (65) Furthermore, the usual *separation principle* does not hold, as estimation of the state of the economy is not independent of optimization. We provide a general characterization of optimal filtering and control for these settings.

JEL Classification: E37, E47, E52, E58

Keywords: Partial information, Kalman filter, monetary policy, commitment, certainty equivalence.

*We thank two referees and the editors for helpful comments. Any errors and expressed views are our own. Woodford thanks the National Science Foundation for research support through a grant to the NBER.

†Corresponding Author: Michael Woodford, Department of Economics, Princeton University, Princeton, NJ 08544. Telephone: 609-258-4003; Fax: 609-258-6413; e-mail: woodford@princeton.edu. *September 1 - December 31, 2002 only:* Department of Economics, Harvard University, Littauer Center, Cambridge, MA 02138. Telephone: 617-495-6675; Fax: 617-495-8570.

‡Princeton University, CEPR and NBER; svensson@princeton.edu; www.princeton.edu/∼svensson.

§Princeton University and NBER; woodford@princeton.edu; www.princeton.edu/∼woodford.

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 (1986) and Currie and Levine (1993). 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

[1] Important early treatments include Chow (1975), Kalchbrenner and Tinsley (1975), and Leroy and Waud (1977).

is no longer so simple. For current observable variable will generally depend not only on the 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 (1986) 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 (1992) 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.[2]

Our own previous paper (Svensson and Woodford, 2002a) 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.

In the present paper, we instead consider the additional complications that arise in the

[2] See equation (4) below, and further discussion there.

case of a particular type of asymmetry between the information available to the central bank and that available to the private sector. The general problem of optimal policy in the case of asymmetric information appears to be quite difficult, so we here restrict our analysis to a particular, relatively simple case, that is nonetheless of considerable interest.[3] This is the case in which the private sector has complete information about the current state of the economy, while the central bank observes only a particular set of indicators, which are also contaminated by "noise" that is unrelated to the fundamental determinants of the state variables that matter for its optimization problem.

One may or may not believe that central banks typically possess less information about the state of the economy than does the private sector. However, there is at least one important argument for the appeal of this assumption. This is that it is the only case in which it is intellectually coherent to assume a *common information set* for all members of the private sector, so that the model's equations can be expressed in terms of aggregate equations that refer to only a single "private sector information set," while at the same time these model equations are treated as structural, and hence invariant under the alternative policies that are considered in the central bank's optimization problem. It does not make sense that any variables should matter for the determination of economically relevant quantities (that is, relevant to the central bank's objectives), if they are not known to *anyone* in the private sector. But if all private agents are to have a common information set, they must then have full information about the relevant variables.[4] It does not follow from this reasoning, of course, that it is more accurate to assume that all private agents have superior information to that of the central bank; it follows only that this case is one in which the complications resulting from partial information are especially tractable. The development of methods for characterizing optimal policy when different private agents have different information sets remains an important topic for further research.[5]

In this paper, we characterize both (Markov-perfect) equilibria in which the central bank

[3] Examples of recent applications of the kind of general analysis offered here to particular problems include Aoki (2000a, 2000b) and Dotsey and Hornstein (2000). The appendix to the latter paper independently derives some of the results presented in section 2 below for the case of discretionary policy, building upon the analysis in Svensson and Woodford (2002a) as we do. Here we also treat the more difficult case of optimal policy under commitment.

[4] The kind of optimal policy problem treated in Svensson and Woodford (2002a) is therefore not one for which rigorous microfoundations can be supplied. Relations such as the Lucas (1973) aggregate supply equation, which appear to be of the form considered in that paper, are actually not fully structural; for (when written as relations between aggregate variables) they involve coefficients that are not invariant to changes in the policy regime, as Lucas stresses.

[5] Important early studies that develop methods for characterizing rational expectations equilibria in which different private agents have different information sets include Townsend (1983), Pearlman (1986), and Sargent (1991).

optimizes under discretion and the optimal policy under commitment, giving particular attention to the optimal filtering problem of the central bank in each case. Some of our main results are parallel in the two cases. In each case, we establish a *certainty-equivalence* principle once again: optimal instrument settings are given by a *reaction function in state-space form* in which the coefficients — describing optimal central-bank responses to the bank's estimates of the predetermined variables characterizing the state of the economy and, in the case of commitment, of the bank's estimates of the shadow prices, Lagrange multipliers, of the equations for the forward-looking variables — are independent of the nature of the bank's partial information (though the estimates that the bank responds to, of course, depend upon this).[6]

At the same time, it is worth noting that, in the case of *commitment in a forward-looking system*, not all representations of optimal policy that would be valid under certainty or under full information continue to be correct in our setting, when the actual values of all variables are replaced by their optimal estimates. In the case of full information, optimal policy under commitment can be described by a reaction function in state-space form,

$$i_t = FX_t + \Phi\Xi_{t-1}, \tag{1}$$

where i_t is the vector of instruments, X_t is the vector of predetermined variables mentioned above, Ξ_{t-1} is the vector of Lagrange multipliers mentioned above (a function of information at date $t-1$), and F and Φ are matrices. But the period-t vector of Lagrange multipliers Ξ_t is a linear function of the same variables as is i_t in (1), and this law of motion can be used to express the Lagrange multipliers as a function of the history of the vector of predetermined variables,

$$\Xi_t = SX_t + \Sigma\Xi_{t-1} = \sum_{j=0}^{t} \Sigma^j S X_{t-j}, \tag{2}$$

where S and Σ are matrices and the commitment is made in period 0. Consequently, optimal policy under commitment can also be described by a reaction function *in integrative form*,

$$i_t = \sum_{j=0}^{t} F_j X_{t-j}, \tag{3}$$

where the matrices $\{F_j\}_{j=1}^{\infty}$ are given by $F_0 = F$, $F_j = \Phi\Sigma^{j-1}S$, $j = 1, 2, ...$[7]

[6] This statement of the principle of certainty equivalence in the context of a forward-looking system follows Pearlman (1992).

[7] Note that in the case of a backward-looking system, or *discretionary* policy for a forward-looking system, $\Phi \equiv 0$ in (1), so that the instrument vector is a function only of the predetermined variables. Thus the integrative form of the reaction function is only relevant for the case of commitment in a forward-looking system.

In the case of symmetric partial information, Pearlman (1992) shows that optimal policy under commitment can still be described by a reaction function in state-space form,

$$i_t = FX_{t|t} + \Phi\Xi_{t-1}, \tag{4}$$

where $X_{t|t}$ is now the central bank's optimal estimate of X_t given its partial information at date t. Here the coefficients F and Φ are the same as in (1). In Svensson and Woodford (2002a) we show that, in the case of symmetric partial information, a principle of certainty equivalence applies to representation (3) as well. Optimal policy under commitment can also be represented by a reaction function in integrative form,

$$i_t = \sum_{j=0}^{t} F_j X_{t-j|t-j}. \tag{5}$$

where once again the matrices $\{F_j\}$ are the same as in the case of full information.

Under the form of asymmetric information considered here, we find that certainty-equivalence still applies to the reaction function in state-space form. Optimal policy under commitment is described by

$$i_t = FX_{t|t} + \Phi\Xi_{t-1|t}, \tag{6}$$

generalizing (4), where once again, the coefficients F and Φ are the same as in the case of certainty (or full information). The only difference is that now the true shadow prices Ξ_{t-1} must be replaced by the central bank's estimates given its information in period t.[8] But the true value of the Lagrange multipliers of the equations for the forward-looking variables need no longer satisfy an equation analogous to (2), nor do the central bank's optimal estimates of the multipliers, $\Xi_{t-1|t}$, generally depend solely on the history of central-bank estimates of the predetermined variables, $\{X_{t-j|t-j}\}$, through coefficients that are the same as in the case of full information. Hence we can no longer express optimal policy solely as a function of the history of estimates of the predetermined variables, as in (6). Thus, under asymmetric information and commitment, certainty equivalence does *not* hold for the reaction function in integrative form. We illustrate the possible failure of certainty equivalence in this sense in an example in section 4, drawn from the work of Aoki (2000b).

Furthermore, under discretion as well as under commitment, we no longer obtain a *separation principle* of the sort obtained for backward-looking systems, or in the case of symmetric partial

[8] When the central bank has less information than the private sector, it may not know the true value that changing the private sector's forecasts at date $t-1$ would have had.

information treated in Svensson and Woodford (2002a). Although the coefficients of the reaction function in state-space form (6) are independent of the parameters describing the central bank's partial information, the central bank's optimal estimates of the unobserved predetermined variables and of the Lagrange multipliers are in general not independent of the parameters of the objective function that it seeks to minimize. Furthermore, the complete equilibrium evolution of the state of the economy, given the path of the central bank's estimates, can no longer be derived independently of the specification of the bank's partial information and hence of the filter that it uses to derive its estimates.

In section 2, we set out the class of linear-quadratic models with which we are concerned, and then consider optimization under discretion. Section 3 presents corresponding results for the case of optimal policy with commitment. Finally, section 4 illustrates how our framework applies to the optimal policy problem considered by Aoki (2000a, 2000b).

2 Optimization under discretion

We begin with an exposition of the general linear-quadratic-Gaussian framework that we assume throughout this paper, and then characterize optimizing policy under discretion, assuming a Markov-perfect equilibrium.

2.1 The framework

We consider a linear model of an economy, in which a vector of predetermined and forward-looking variables are determined by a system of structural equations intended to represent the conditions for rational-expectations equilibrium given optimizing private-sector behavior. We do not explicitly describe the optimization problems of private agents that underlie these equations, but we note that such optimization results in the presence of forward-looking terms (that is, conditional expectations) in the structural equations, and we take care to specify the common information set of all agents in the private sector, with respect to which these conditional expectations are defined. The optimization problem of the central bank (or more generally, the policymaker) is instead described explicitly; in this problem the structural equations resulting from private-sector optimization appear as constraints. We also take care to specify the information set of the central bank, which differs from that of the private sector.

The structural equations are given by a system of the form

$$\begin{bmatrix} X_{t+1} \\ \tilde{E}\mathrm{E}_t x_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ x_t \end{bmatrix} + Bi_t + \begin{bmatrix} u_{t+1} \\ 0 \end{bmatrix}, \tag{7}$$

where X_t is a vector of n_X *predetermined variables* in period t, x_t is a vector of n_x *forward-looking variables*, i_t is a vector of the central bank's n_i policy *instrument*(*s*), u_t is a vector of n_X iid shocks with mean zero and covariance matrix Σ_{uu}, and A, B and $\tilde{E}$ are matrices of appropriate dimension.[9] The $n_x \times n_x$ matrix $\tilde{E}$ (which should not be confused with the expectations operator $\mathrm{E}[\cdot]$) may be singular (this is a slight generalization of common formulations in which $\tilde{E}$ is the identity matrix).

We let $\mathrm{E}_t z_\tau$ denote the rational expectation (the best estimate) of any variable z_τ in period τ, given private-sector information in period t. We will assume that the private-sector has full information, I_t^f, so $\mathrm{E}_t z_\tau \equiv \mathrm{E}[z_\tau | I_t^f]$. We let $z_{\tau|t}$ denote the rational expectation (the best estimate) of z_τ given the central-bank information in period t. We let I_t denote central-bank information in period t, so $z_{\tau|t} \equiv \mathrm{E}[z_\tau | I_t]$. The information is further specified below.

We let Y_t denote a vector of n_Y *target variables* given by

$$Y_t = C \begin{bmatrix} X_t \\ x_t \end{bmatrix} + C_i i_t, \tag{8}$$

where C and C_i are matrices of appropriate dimension. We then let the quadratic form

$$L_t = Y_t' W Y_t \tag{9}$$

be the central bank's period loss function, where W is a positive semidefinite weight matrix.

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

$$Z_t = D \begin{bmatrix} X_t \\ x_t \end{bmatrix} + v_t, \tag{10}$$

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

$$I_t = \{Z_\tau, \tau \le t;\ A, B, C, C_i, D, \tilde{E}, W, \delta, \Sigma_{uu}, \Sigma_{vv}\}, \tag{11}$$

[9] In Svensson and Woodford (2002a), we allow the structural equations (7) that determine X_t and x_t to also include general linear terms involving the estimates $X_{t|t}$ and $x_{t|t}$ of the state of the economy based on partial information (of a kind specified below). However, the presence of such terms makes less sense when (as here) we assume that only the central bank has partial information. Equations (7) indicate how the state variables are determined by private-sector behavior given the central bank's instrument settings i_t, and the beliefs of the central bank (as opposed to those of the private sector) are not obviously relevant. Similar methods to those expounded here can be employed even in the presence of the additional terms, but at the cost of additional algebraic complexity.

where δ $(0 < \delta < 1)$ is a discount factor (to be introduced below). This incorporates the case when some or all of the predetermined and forward-looking variables are observable by the central bank. The full information I_t^f in period t, the private-sector information set, is given by

$$I_t^f = \{X_\tau, x_\tau, i_\tau, \tau \leq t;\ A, B, C, C_i, D, \tilde{E}, W, \delta, \Sigma_{uu}, \Sigma_{vv}\}. \tag{12}$$

Thus, we here assume that the central bank has the same or less information than the private sector. The special case where both the central bank and the private sector have the same partial information is treated in Svensson and Woodford (2002a).[10]

2.2 Certainty equivalence

In the present section, we assume that there is no commitment mechanism, so the central bank acts under discretion, and seek to characterize Markov-perfect equilibrium. This means that (as is often the case in the literature on discretionary monetary policy) we consider only equilibria in which there is no dependence of equilibrium outcomes on "irrelevant" state variables, that is, state variables that have no consequences for the set of possible outcomes from the current date onward that would be consistent with the structural relations (7).

Assume that the central bank each period, conditional on the information I_t, minimizes the expected discounted current and future values of the intertemporal loss function,

$$\mathrm{E}[\sum_{\tau=0}^{\infty} \delta^\tau L_{t+\tau} | I_t]. \tag{13}$$

Svensson and Woodford (2002a, Appendix A) show that *certainty equivalence* applies for the central bank's optimization when the central bank and the private sector has the same information. The same proof goes through in the present case of asymmetric information, as shown in appendix A.

To get some intuition for why the same proof goes through, we note that by taking the expectation of (7), conditional on the central banks information in period t , I_t, we get

$$\begin{bmatrix} X_{t+1|t} \\ \tilde{E}x_{t+1|t} \end{bmatrix} = A \begin{bmatrix} X_{t|t} \\ x_{t|t} \end{bmatrix} + Bi_t. \tag{14}$$

[10] Note that the predetermined and forward-looking variables can be interpreted as deviations from unconditional means and the target variables can be interpreted as deviations from constant target levels. More generally, constants, non-zero unconditional means and non-zero target levels can be incorporated by including unity among the predetermined variables, for instance, as the last element of X_t. The last row of the relevant matrices will then include the corresponding constants/means/target levels.

Thus the expectation of the structural equations look the same as they would under the symmetric-information case treated in Svensson and Woodford (2002a), so it is not surprising that the equilibrium policy and the estimate of the forward-looking variables will again be the same functions of the estimate of the predetermined variables.

Thus, the equilibrium under discretion will be characterized by the instrument being a linear function of the current estimate of the predetermined variables,

$$i_t = FX_{t|t}. \tag{15}$$

Furthermore, the estimate of the forward-looking variables will fulfill

$$x_{t|t} = GX_{t|t}, \tag{16}$$

where the matrix G by appendix A fulfills

$$G = (A_{22} - \tilde{E}GA_{12})^{-1}[-A_{21} + \tilde{E}GA_{11} + (\tilde{E}GB_1 - B_2)F], \tag{17}$$

where the matrices A and B are decomposed according to X_t and x_t,

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

and we assume that the matrix $A_{22} - \tilde{E}GA_{12}$ is invertible. The matrices F and G depend on A, B, C, C_i, $\tilde{E}$, W and δ, but (because of certainty equivalence) *not* on D, Σ_{uu} and Σ_{vv}.

2.3 Failure of the separation principle

The forward-looking variables will depend on both the true value and the central bank's estimate of the predetermined variables, and they can be written

$$x_t = G^1 X_t + (G - G^1)X_{t|t}, \tag{18}$$

where the matrix G^1 remains to be determined. In the case of symmetric information, Svensson and Woodford (2002a) shows that the matrix G^1 is given by

$$G^1 = -A_{22}{}^{-1}A_{21} \tag{19}$$

and hence depends only on A (A_{22} is assumed to be invertible).[11] We will see that in the present case of asymmetric information, G^1 is determined in a more complex way. In particular, it is no longer generally independent of the specification of the central bank's partial information.

[11] In that paper, we write (19) as $G^1 = -A_{22}^1{}^{-1}A_{21}^1$, because we allow there for a term corresponding to a matrix A^2 multiplying the central bank's estimate of the predetermined and forward-looking variables in (7), cf. footnote 9.

For a given G^1, it follows from (7), (10), (15) and (18) that the dynamics for X_t and Z_t follow

$$X_{t+1} = HX_t + JX_{t|t} + u_{t+1}, \tag{20}$$

$$Z_t = LX_t + MX_{t|t} + v_t, \tag{21}$$

where

$$H \equiv A_{11} + A_{12}G^1, \tag{22}$$

$$J \equiv B_1 F + A_{12}(G - G^1), \tag{23}$$

$$L \equiv D_1 + D_2 G^1, \tag{24}$$

$$M \equiv D_2(G - G^1), \tag{25}$$

and where $D \equiv [D_1 \quad D_2]$ is decomposed according to X_t and x_t. (Note that the matrix L in (24) should not be confused with the period loss function L_t in (9).)

We note that, as in the symmetric-information case, the problem of estimating the predetermined variables has been transformed to a problem without forward-looking variables, with the transition equation (20) and the measurement equation (21); once again, the only nonstandard feature of this filtering problem is the circularity implied by the appearance of $X_{t|t}$ on the right-hand side of the measurement equation. Thus, as discussed in Svensson and Woodford (2002a), with forward-looking variables among the observable variables, there is a simultaneity problem because the forward-looking variables depend on the current estimate of the predetermined variables and the latter depend on the observable variables.

Temporarily ignoring the simultaneity problem, the optimal estimate of X_{t+1} will be given by a Kalman filter updating equation,

$$X_{t+1|t+1} = X_{t+1|t} + K(Z_{t+1} - LX_{t+1|t} - MX_{t+1|t+1}), \tag{26}$$

where the Kalman gain matrix K remains to be determined. The updating equation can, given (21), be written

$$X_{t+1|t+1} = X_{t+1|t} + K[L(X_{t+1} - X_{t+1|t}) + v_{t+1}]. \tag{27}$$

The Kalman gain matrix is given by

$$K = PL'(LPL' + \Sigma_{vv})^{-1}, \tag{28}$$

where the matrix $P \equiv \mathrm{Cov}[X_{t+1} - X_{t+1|t}]$ is the covariance matrix for the prediction errors $X_{t+1} - X_{t+1|t}$ and satisfies

$$P = H[P - PL'(LPL' + \Sigma_{vv})^{-1}LP]H' + \Sigma_{uu}. \tag{29}$$

Thus for given G^1 and hence given H and L, P can be solved from (29), either numerically or analytically, depending upon the complexity of the matrices H, L and Σ_{uu}. Then K is given by (28).

The simultaneity in (26) means that the updating equation is not operational. Solving (26) for $X_{t+1|t+1}$ eliminates the simultaneity and results in

$$\begin{aligned} X_{t+1|t+1} &= (I + KM)^{-1}[(I - KL)X_{t+1|t} + KZ_{t+1}] \\ & \quad (I + KM)^{-1}[(I - KL)(H + J)X_{t|t} + KZ_{t+1}], \end{aligned} \tag{30}$$

where we have used the fact that, by (20), the prediction equation will be

$$X_{t+1|t} = (H + J)X_{t|t}. \tag{31}$$

Equation (30) is an operational recursive updating equation, which avoids the simultaneity problem and uses the current observable variables to update the previous estimate $X_{t|t}$.

It only remains to determine G^1. Assuming a solution of the form (18), we must have

$$\tilde{E}\mathrm{E}_t x_{t+1} = \tilde{E}G^1\mathrm{E}_t X_{t+1} + \tilde{E}(G - G^1)\mathrm{E}_t X_{t+1|t+1}. \tag{32}$$

We would like to express the right-hand side as a function of X_t, x_t and $X_{t|t}$. We can then equate this with the second block of (7),

$$\tilde{E}\mathrm{E}_t x_{t+1} = A_{21}X_t + A_{22}x_t + B_2 F X_{t|t} \tag{33}$$

(where we have used (16)), solve for x_t as a function of X_t and $X_{t|t}$, and identify G^1.

We note that, by (20) and (27), we have

$$\begin{aligned} \mathrm{E}_t X_{t+1} &= HX_t + JX_{t|t} \\ \mathrm{E}_t X_{t+1|t+1} &= X_{t+1|t} + KL(\mathrm{E}_t X_{t+1} - X_{t+1|t}) \\ &= (H + J)X_{t|t} + KLH(X_t - X_{t|t}). \end{aligned}$$

Substituting this into (32), we obtain

$$\begin{aligned} \tilde{E}\mathrm{E}_t x_{t+1} &= \tilde{E}G^1(HX_t + JX_{t|t}) + \tilde{E}(G - G^1)[(H + J)X_{t|t} + KLH(X_t - X_{t|t})] \\ &= \tilde{E}[G^1 + (G - G^1)KL]H(X_t - X_{t|t}) + \tilde{E}G(H + J)X_{t|t}. \end{aligned} \tag{34}$$

Equating this with (33) and solving for x_t, we obtain a solution of the form (18), where G^1 must satisfy

$$G^1 = A_{22}^{-1}\{-A_{21} + [G^1 + (G - G^1)KL]H\}. \tag{35}$$

Thus, in contrast to (19) for the case where the private sector and central bank have the same information, the matrix G^1 now depends not only on A but on KL as well. Thus G^1, H, L, K and P are all simultaneously determined, by the equations (35), (22), (24), (28) and (29). In particular, since K and P depend on Σ_{uu} and Σ_{vv}, so do the other matrices.

In the special case when $D_2 = 0$, that is, when none of the observable variables depend directly on any forward-looking variables, L is by (24) independent of G^1 and hence of K. The other matrices are then still simultaneously determined. Recall that F and G only depend on A, B, C, C_i, W and δ, and not on D, K or L, and hence not on Σ_{uu} and Σ_{vv}.

In the case when the private sector has the same information as the central bank, the terms that are multiplied with the factor $G^1 + (G - G^1)KL$ all vanish in (34) (since $\mathrm{E}_t X_t - X_{t|t} = X_{t|t} - X_{t|t} = 0$). As a consequence, the terms multiplying that same factor in (35) vanish, and (35) reduces to (19) as in Svensson and Woodford (2002a).

Thus, in the asymmetric-information case considered here, where the private-sector has full information and the central bank has partial information, *certainty equivalence* still applies for the central bank's optimization problem, in the sense that the implied reaction function and response F to its estimate of the predetermined variables is independent of the variance of the shocks, Σ_{uu}, and the information structure of the economy, D^1, D^2 and Σ_{vv}. However, the *separation principle* does not apply in the way that it does in the symmetric-information case. The estimation is now more complex and, in particular, it is not independent of the central bank's objective and its choice of F. (Note that G^1 by (35) depends on G, which depends on F and hence C, C_i, W and δ.)

3 Optimal policy with commitment

Again, we assume the model (7), in which the private sector has full information, but now suppose instead that the central bank commits itself to an optimal plan for the indefinite future at information set I_0 in period 0. As in the case of symmetric partial information treated in Svensson and Woodford (2002a), the optimal commitment can be derived using a Lagrangian approach to what is essentially a planning problem in which the structural equations (7) are constraints. In the present informational assumptions, the Lagrangian takes the form

$$\begin{aligned}\mathcal{L} &= \mathrm{E}\left\{\sum_{t=0}^{\infty}\delta^t L_t - \sum_{t=0}^{\infty}\delta^t\varphi'_{1,t+1}(X_{t+1} - A_{11}X_t - A_{12}x_t - B_1 i_t)\right. \\ &\left. - \sum_{t=0}^{\infty}\delta^t\varphi'_{2,t+1}(\tilde{E}\mathrm{E}_t x_{t+1} - A_{21}X_t - A_{22}x_t - B_2 i_t) - \delta^{-1}\varphi'_{10}X_0 \mid I_0\right\},\end{aligned}$$

where in each period t, φ_{1t} and φ_{2t} are vectors of Lagrange multipliers conformable to X_t and x_t respectively. The multipliers $\varphi_{1,t+1}$ and $\varphi_{2,t+1}$ indicate the value of relaxing the constraints represented by the first and second rows of (7) respectively; the term $-u_{t+1}$ has been suppressed inside the first set of square brackets as it is irrelevant for the first-order conditions derived below. Because the second row of (7) represents a set of equilibrium conditions for the determination of x_t as a function of information in I_t^f (rather than for the determination of x_{t+1} as a function of later information), the multipliers $\varphi_{2,t+1}$ are in fact measurable with respect to (that is, depend only on) the period t information set I_t^f, whereas the multipliers $\varphi_{1,t+1}$ are measurable with respect to I_{t+1}^f. (The notation $\varphi_{2,t+1}$ rather than φ_{2t} is nonetheless convenient in writing expressions such as (36) below.) The final term on the right-hand side corresponds to the constraint imposed by the vector of initial conditions X_0.[12]

3.1 Certainty equivalence of the reaction function in state-space form

The Lagrangian can be written more compactly[13] in the form

$$\mathcal{L} = \mathrm{E}\left\{\sum_{t=0}^{\infty}\delta^t\left[L_t + \varphi'_{t+1}(Ay_t + Bi_t) - \delta^{-1}\varphi'_t\tilde{I}y_t\right] \mid I_0\right\}, \tag{36}$$

where

$$y_t \equiv \begin{bmatrix} X_t \\ x_t \end{bmatrix}, \qquad \varphi_t \equiv \begin{bmatrix} \varphi_{1t} \\ \varphi_{2t} \end{bmatrix}, \qquad \tilde{I} \equiv \begin{bmatrix} I & 0 \\ 0 & \tilde{E} \end{bmatrix}.$$

We have added a term $-\delta^{-1}\varphi'_{20}x_0$ to the terms inside the square brackets, for the sake of symmetry in notation, but now correspondingly stipulate the initial condition

$$\varphi_{20} = 0.$$

(Note that these Lagrange multipliers do not correspond to any actual constraint upon the planning problem. They are added solely in order to allow us to write the first-order conditions (37)

[12] Note also that we don't 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 $z_{\tau|t}$ for different variables z_τ as a result of (10). As shown in Svensson and Woodford (2002b), we obtain the correct first-order conditions even when the additional constraints are omitted.

[13] Here we use the fact that $\mathrm{E}[\varphi_{2,t+1}\mathrm{E}_t x_{t+1} \mid I_0] = \mathrm{E}[\mathrm{E}_t\varphi_{2,t+1}x_{t+1} \mid I_0] = \mathrm{E}[\varphi_{2,t+1}x_{t+1} \mid I_0]$, by the law of iterated expectations.

below in a time-invariant form, despite the fact that the central bank's commitment regarding period zero does not affect expectations in any previous period, unlike all subsequent periods.) This should be interpreted as an initial condition for each possible state of economy in period zero consistent with the information set I_0, rather than a single initial condition for I_0.

Differentiation of (36) with respect to y_t and i_t then yields the first-order conditions

$$A'\mathrm{E}_t\varphi_{t+1} + L_{yy}y_t + L_{yi}i_t - \delta^{-1}\tilde{I}'\varphi_t = 0, \tag{37}$$

$$B'\varphi_{t+1|t} + L_{iy}y_{t|t} + L_{ii}i_t = 0, \tag{38}$$

where the matrices L_{jk} represent second partial derivatives of the period loss function. Note that equations (8) and (9) imply that L_t is a quadratic function of y_t and i_t,

$$L_t = \begin{bmatrix} y_t' & i_t \end{bmatrix} \begin{bmatrix} C' \\ C_i' \end{bmatrix} W \begin{bmatrix} C & C_i \end{bmatrix} \begin{bmatrix} y_t \\ i_t \end{bmatrix} \equiv \frac{1}{2} \begin{bmatrix} y_t' & i_t \end{bmatrix} \begin{bmatrix} L_{yy} & L_{yi} \\ L_{iy} & L_{ii} \end{bmatrix} \begin{bmatrix} y_t \\ i_t \end{bmatrix}, \tag{39}$$

so that the L_{jk} are matrices of constant coefficients, that depend only upon the elements of the matrices C, C_i, and W.

Assuming that L_{ii} is of full rank,[14] we can solve (38) for i_t, obtaining

$$i_t = -L_{ii}^{-1}L_{iy}y_{t|t} - L_{ii}^{-1}B'\varphi_{t+1|t}. \tag{40}$$

Substituting (40) into (7) and (37) to eliminate i_t, and taking the conditional expectation of both equations with respect to I_t, we then obtain a system of equations for the expected evolution of y_t and φ_t, that can be written in the form

$$\begin{bmatrix} 0 & R' \\ \tilde{I} & U \end{bmatrix} \begin{bmatrix} y_{t+1|t} \\ \varphi_{t+1|t} \end{bmatrix} = \begin{bmatrix} V & \delta^{-1}\tilde{I}' \\ R & 0 \end{bmatrix} \begin{bmatrix} y_{t|t} \\ \varphi_{t|t} \end{bmatrix}, \tag{41}$$

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.

[14] The case in which $C_i = 0$, with the consequence that $L_{ii} = 0$, can be treated using similar methods, as shown in appendix B.

Let us assume furthermore that the square matrix on the left-hand side of (41) is of full rank. Then we can invert this matrix, to obtain a system of the form

$$\begin{bmatrix} y_{t+1|t} \\ \varphi_{t+1|t} \end{bmatrix} = \mathcal{M} \begin{bmatrix} y_{t|t} \\ \varphi_{t|t} \end{bmatrix}. \tag{42}$$

We then wish to consider solutions to (42) that are consistent with given initial values for $X_{t|t}$ and $\varphi_{2,t|t}$. We note that the number of variables in (42) 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 ($X_{t|t}$ and $\varphi_{2,t|t}$). We shall restrict our attention to bounded solutions, by which we mean solutions in which for any t, $y_{t+j|t}$ and $\varphi_{t+j|t}$ satisfy a uniform bound for all j. Such solutions necessarily satisfy the transversality condition for an optimal plan, and since our equations (7)–(10) 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 (42) 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 μ of $\mathcal{M}$ are the roots of the characteristic equation

$$\mathrm{Det} \begin{bmatrix} V & \delta^{-1}\tilde{I}' - \mu R' \\ R - \mu \tilde{I} & -\mu U \end{bmatrix} = 0. \tag{43}$$

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

$$\mathrm{Det} \begin{bmatrix} V & \delta^{-1}\tilde{I}' - \mu^{-1}\delta^{-1} R' \\ R - \mu^{-1}\delta^{-1} \tilde{I} & -\mu^{-1}\delta^{-1} U \end{bmatrix} = 0.$$

Comparison of this with (43) shows that if μ is a root, $\mu^{-1}\delta^{-1}$ must also be. It follows that $\mathcal{M}$ has as many eigenvalues with $|\mu| > \frac{1}{\sqrt{\delta}}$ as with $|\mu| < \frac{1}{\sqrt{\delta}}$. Thus, since $\frac{1}{\sqrt{\delta}} > 1$, *at most* half of the eigenvalues (that is, at most $n_X + n_x$) are inside the unit circle (that is, with $|\mu| < 1$), so there is no possibility of multiple stationary solutions to (42). 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.[15] Then (42) has a unique bounded solution in which $y_{t+j|t}$ and $\varphi_{t+j|t}$ can be expressed as linear functions of the initial conditions $X_{t|t}$ and $\varphi_{2,t|t}$, for arbitrary $j \geq 0$.

[15] In the case of an exact linear-quadratic model, as opposed to a mere local approximation to a nonlinear

Substitution of this solution into (40) implies evolution of the instrument according to a relation of the form

$$i_t = FX_{t|t} + \Phi\Xi_{t-1|t}, \tag{44}$$

where F and Φ are matrices of constant coefficients, and we now introduce the notation $\Xi_{t-1} \equiv \varphi_{2t}$ for the sub-vector of Lagrange multipliers that are predetermined variables under our characterization of the optimal commitment, just as in the symmetric-information case treated in Svensson and Woodford (2002a). (In line with the notation in our earlier paper, we change the time subscript to emphasize that the elements of Ξ_{t-1} are determined at date $t-1$.) Thus optimal policy under commitment can once again be described by a reaction function *in state-space form,* as discussed in the introduction.

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

$$x_{t|t} = GX_{t|t} + \Gamma\Xi_{t-1|t}, \tag{45}$$

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

$$\Xi_{t|t} = SX_{t|t} + \Sigma\Xi_{t-1|t}, \tag{46}$$

where G, Γ, S, Σ are further matrices of constant coefficients. Substitution of these solutions into (7) then implies that the conditional expectations of the predetermined variables evolve according to

$$X_{t+1|t} = (H + J)X_{t|t} + \Psi\Xi_{t-1|t}, \tag{47}$$

where H and J are again defined as in (22)–(23), and

$$\Psi \equiv A_{12}\Gamma + B_1\Phi. \tag{48}$$

These equations completely define the conditional expectations of both the predetermined variables and the central bank's instrument settings at all future dates, as functions of the current estimates of the predetermined variables $X_{t|t}$ and $\Xi_{t-1|t}$. We note that all of the matrices F, Φ, G, Γ, S, Σ, $H+J$, and Ψ are exactly the same as in the case of the optimal plan with full information (and are independent of the matrices D and Σ_{vv} that define the partial information

model, this condition is not necessary in order for us to identify the unique relevant solution to (42), as the unique solution that does not explode fast enough to violate the transversality condition. But in practice we are not likely to deal with models that we can regard as exact, and instead assume that the optimal plan happens to be stationary. In this case, the eigenvalue condition assumed in the text must hold, as there would otherwise be no stationary solution.

of the central bank); they are furthermore the same as in the case where there is no uncertainty at all (and independent of the matrix Σ_{uu} defining the fundamental uncertainty). Hence a state-space representation of the consequences of optimal policy under commitment continues to conform to a principle of *certainty equivalence*.[16]

The system of equations (44)–(47) are also exactly the same as those that characterize the optimal commitment in the case of symmetric (though incomplete) information on the part of the central bank and the private sector, as shown in Svensson and Woodford (2002a). The only difference is that in the symmetric case, the Lagrange multipliers associated with the forward-looking variables satisfy $\Xi_{t-1|t} = \Xi_{t-1|t-1} = \Xi_{t-1}$, so that equations (44)–(47) can be written in terms of the multipliers themselves rather than their conditional expectations, while this is not generally true when the central bank has less information. (We show below that in general Ξ_{t-1} depends on information not possessed by the central bank at date $t-1$.)

As noted in the introduction, however, it is important to note that the principle of certainty equivalence does *not* apply to the reaction function in integrative form (5) in the present case. With asymmetric information, because $\Xi_{t|t} \neq \Xi_t$, we can no longer use equation (46) alone to express $\Xi_{t-1|t}$ as a function of the history of central-bank estimates of the predetermined variables, and thus obtain a representation of optimal policy of the form (5) that is independent of the information structure. Indeed, there is not generally a representation of optimal policy of this form at all. Further consideration of this question requires that we take up the central bank's optimal filtering problem.

3.2 Failure of the separation principle

In the case of symmetric information, we can also take the second row of (7), subtract each term from its expectation conditional upon I_t, and obtain a set of n_x linear relations between $x_t - x_{t|t}$ and $X_t - X_{t|t}$ that must hold at all times. (The derivation of these relations depends upon the fact that the $\mathrm{E}_t x_{t+1}$ term in the second row of (7) is instead replaced by $x_{t+1|t}$ when the private sector has the same incomplete information as the central bank.) These relations can be solved for x_t, yielding (when (45) is used as well) a solution for x_t as a linear function of $X_t, X_{t|t}$, and Ξ_{t-1}. Substituting this in turn into the first row of (7), one is able to solve for the complete dynamics of the predetermined variables (and not merely their expectations conditional upon the central bank's information) and of the Lagrange multipliers given the evolution of the estimates

[16] As noted in the introduction, this is the sense in which certainty equivalence is established by Pearlman (1992) for a forward-looking system with symmetric partial information.

$X_{t|t}$, with coefficients that are independent of the nature of the partial information. This is an aspect of the *separation principle* that can be shown to hold in the case of symmetric incomplete information, just as in the case of full information. However, that derivation (detailed in Svensson and Woodford, 2002b) cannot be applied here, given that $\mathrm{E}_t x_{t+1}$ must be distinguished from $x_{t+1|t}$.

However, the second row of (7) does imply that

$$\tilde{I}(\mathrm{E}_t y_{t+1} - y_{t+1|t}) = A\hat{y}_t, \tag{49}$$

where $\hat{y}_t \equiv y_t - y_{t|t}$. (Here we use the fact that i_t must be measurable with respect to I_t.) Similarly, (37) implies that

$$A'(\mathrm{E}_t \varphi_{t+1} - \varphi_{t+1|t}) = \delta^{-1}\, \tilde{I}'\hat{\varphi}_t - L_{yy}\hat{y}_t, \tag{50}$$

where $\hat{\varphi}_t \equiv \varphi_t - \varphi_{t|t}$. Let us write the system (49)–(50) as

$$\tilde{A}\left(\begin{bmatrix} \mathrm{E}_t y_{t+1} \\ \mathrm{E}_t \varphi_{t+1} \end{bmatrix} - \begin{bmatrix} y_{t+1|t} \\ \varphi_{t+1|t} \end{bmatrix}\right) = \tilde{B}\begin{bmatrix} \hat{y}_t \\ \hat{\varphi}_t \end{bmatrix}, \tag{51}$$

where

$$\tilde{A} \equiv \begin{bmatrix} \tilde{I} & 0 \\ 0 & A' \end{bmatrix}, \quad \tilde{B} \equiv \begin{bmatrix} A & 0 \\ -L_{yy} & \delta^{-1}\tilde{I}' \end{bmatrix}.$$

In order to solve (51) for the dynamics of $\hat{y}_t$ and $\hat{\varphi}_t$, we must specify how the central bank's conditional expectations are updated (for instance, how $y_{t+1|t+1}$ relates to $y_{t+1|t}$). Let us suppose that the Kalman filter for the bank's problem takes the form

$$\begin{bmatrix} y_{t+1|t+1} \\ \varphi_{t+1|t+1} \end{bmatrix} = \begin{bmatrix} y_{t+1|t} \\ \varphi_{t+1|t} \end{bmatrix} + \tilde{K}[\bar{L}(\bar{X}_{t+1} - \bar{X}_{t+1|t}) + v_{t+1}], \tag{52}$$

where

$$\bar{L} \equiv [L \ \ \Lambda^1] \ \text{ and } \ \bar{X}_t \equiv \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix},$$

where $\tilde{K}$ and $\bar{L}$ are matrices that have yet to be identified. (We show below that the Kalman filter does indeed take this form, and identify these matrices.)

Equation (52) implies

$$\begin{bmatrix} \mathrm{E}_t y_{t+1|t+1} \\ \mathrm{E}_t \varphi_{t+1|t+1} \end{bmatrix} = \begin{bmatrix} y_{t+1|t} \\ \varphi_{t+1|t} \end{bmatrix} + \tilde{K}\bar{L}\left(\mathrm{E}_t \bar{X}_{t+1} - \bar{X}_{t+1|t}\right). \tag{53}$$

By adding $(\mathrm{E}_t y'_{t+1}, \mathrm{E}_t \varphi'_{t+1})'$ to both sides and rearranging, we can write the equation as

$$\begin{bmatrix} \mathrm{E}_t y_{t+1} \\ \mathrm{E}_t \varphi_{t+1} \end{bmatrix} - \begin{bmatrix} y_{t+1|t} \\ \varphi_{t+1|t} \end{bmatrix} = \begin{bmatrix} \mathrm{E}_t \hat{y}_{t+1} \\ \mathrm{E}_t \hat{\varphi}_{t+1} \end{bmatrix} + \tilde{K}\bar{L}\left(\mathrm{E}_t \bar{X}_{t+1} - \bar{X}_{t+1|t}\right). \tag{54}$$

Let $\bar{I}$ be the submatrix of the identity matrix that selects the elements

$$\bar{X}_t \equiv \bar{I} \begin{bmatrix} y_t \\ \varphi_t \end{bmatrix}.$$

Then we can write (54) as

$$\begin{bmatrix} \mathrm{E}_t y_{t+1} \\ \mathrm{E}_t \varphi_{t+1} \end{bmatrix} - \begin{bmatrix} y_{t+1|t} \\ \varphi_{t+1|t} \end{bmatrix} = (I - \tilde{K}\bar{L}\bar{I})^{-1} \begin{bmatrix} \mathrm{E}_t \hat{y}_{t+1} \\ \mathrm{E}_t \hat{\varphi}_{t+1} \end{bmatrix}, \tag{55}$$

where we assume that $I - \tilde{K}\bar{L}\bar{I}$ is invertible. Substituting (55) into (51) results in

$$\tilde{A}(I - \tilde{K}\bar{L}\bar{I})^{-1} \begin{bmatrix} \mathrm{E}_t \hat{y}_{t+1} \\ \mathrm{E}_t \hat{\varphi}_{t+1} \end{bmatrix} = \tilde{B} \begin{bmatrix} \hat{y}_t \\ \hat{\varphi}_t \end{bmatrix}. \tag{56}$$

Again, we are interested in bounded solutions of (56) consistent with given initial values for $\hat{X}_t$ and $\hat{\Xi}_{t-1}$. We assume the appropriate configuration of eigenvalues of the system. There is then a unique bounded solution of the form

$$\hat{x}_t = \bar{G}^1(\bar{X}_t - \bar{X}_{t|t}), \tag{57}$$

$$\hat{\Xi}_t = \bar{S}^1(\bar{X}_t - \bar{X}_{t|t}), \tag{58}$$

where as usual the matrices $\bar{G}^1 \equiv [G^1 \;\; \Gamma^1]$ and $\bar{S}^1 \equiv [S^1 \;\; \Sigma^1]$ can be derived from the eigenvectors of the system (56).

Combining (45) with (57), we find that

$$x_t = \bar{G}^1\bar{X}_t + (\bar{G} - \bar{G}^1)\bar{X}_{t|t}, \tag{59}$$

where $\bar{G} \equiv [G \;\; \Gamma]$, and the matrices G and Γ are the ones that appear in (45). This equation is similar in form to the solution obtained in the case of symmetric partial information (discussed above), except that it includes separate terms in Ξ_{t-1} and $\Xi_{t-1|t}$ (which now differ), and that the matrix G^1 no longer takes the simple form (19). In particular, the matrix G^1 (and similarly Γ^1, which matters when $\Xi_{t-1} \neq \Xi_{t-1|t}$) depends in general upon the elements of the Kalman gain matrix $\bar{K}$, and so it is no longer independent of the nature of the central bank's partial information.

A complete solution for the dynamics of the predetermined variables also now requires that we solve for the dynamics of the multipliers Ξ_t, which are no longer fully characterized by (46). Combining (46) and (57), we find that

$$\Xi_t = \bar{S}^1 \bar{X}_t + (\bar{S} - \bar{S}^1)\bar{X}_{t|t}, \tag{60}$$

where $\bar{S} \equiv [S \quad \Sigma]$, and S and Σ are the matrices appearing in (46). It is the fact that $\bar{S}^1$ is generally non-zero in this derivation that implies that Ξ_t is no longer measurable with respect to I_t. Unless all elements of X_t are part of the central bank's information set at date $t+1$, Ξ_t will generally not be measurable with respect to I_{t+1}, either.

It remains to determine the evolution of the estimates $\bar{X}_{t|t}$. We note that, in the present case (unlike that of symmetric partial information), it is necessary to derive a Kalman filter for the evolution of the estimates $\Xi_{t-1|t}$ as well as for the estimates $X_{t|t}$. Substitution of (44) and (59) into the first row of (7), combination with (60), and substitution into (10) leads to the transition equation for $\bar{X}_t$,

$$\bar{X}_{t+1} = \bar{H}\bar{X}_t + \bar{J}\bar{X}_{t|t} + \bar{u}_{t+1}, \tag{61}$$

and to the measurement equation

$$Z_t = \bar{L}\bar{X}_t + \bar{M}\bar{X}_{t|t} + v_t. \tag{62}$$

Here

$$\bar{H} \equiv \begin{bmatrix} H & \Psi^1 \\ S^1 & \Sigma^1 \end{bmatrix}, \quad \bar{J} \equiv \begin{bmatrix} J & \Psi^2 \\ S - S^1 & \Sigma - \Sigma^1 \end{bmatrix}, \quad \bar{u}_t \equiv \begin{bmatrix} u_t \\ 0 \end{bmatrix}, \quad \bar{L} \equiv [L \quad \Lambda^1], \quad \bar{M} \equiv [M \quad \Lambda^2], \tag{63}$$

H, J, L, M are again defined as in (22)–(25), and

$$\Psi^1 = A_{12}\Gamma^1, \ \Psi^2 \equiv A_{12}(\Gamma - \Gamma^1) + B_1\Phi, \ \Lambda^1 \equiv D_2\Gamma^1, \ \Lambda^2 \equiv D_2(\Gamma - \Gamma^1).$$

Note that $\Psi \equiv \Psi^1 + \Psi^2$ is the same matrix as in (47), so that (61) is consistent with (though not implied by) equations (46) and (47) for the evolution of the conditional expectations.

The system of equations consisting of (61) then represents a system of transition equations with a form analogous to (20), in which the vector of predetermined variables X_t is replaced by the vector $\bar{X}_t$ that includes the elements of both X_t and Ξ_{t-1}. Equation (62) is similarly a measurement equation with a form analogous to (21). It follows that the optimal estimates of X_t and Ξ_{t-1} will be given by a Kalman filter analogous to (27). We thus obtain updating equations for X_t and Ξ_{t-1} of the form indicated by the corresponding rows of (52),

$$\bar{X}_{t+1|t+1} = \bar{X}_{t+1|t} + \bar{K}[\bar{L}(\bar{X}_{t+1} - \bar{X}_{t+1|t}) + v_{t+1}], \tag{64}$$

where $\bar{K} \equiv \bar{I}\tilde{K}$.

Using (61) to form the forecast $\bar{X}_{t+1|t}$, (64) may be written

$$\bar{X}_{t+1|t+1} = (I - \bar{K}\bar{L})(\bar{H} + \bar{J})\bar{X}_{t|t} + \bar{K}(\bar{L}\bar{X}_{t+1} + v_{t+1}). \tag{65}$$

For a given Kalman gain matrix $\bar{K}$, a complete system of equations for the evolution of the endogenous variables is then given by (59), (61) and (65), which apply in each period $t \geq 0$, starting from given initial values for X_0 and $X_{0|0}$, and from the initial values $\Xi_{-1} = \Xi_{-1|0} = 0$.

It remains to determine the Kalman gain matrix in (64). Making use of the analogy between equations (61)–(64) and (20), (21) and (27), we directly find that the Kalman gain matrix is given by the equation

$$\bar{K} = \bar{P}\bar{L}'(\bar{L}\bar{P}\bar{L}' + \Sigma_{vv})^{-1}, \tag{66}$$

analogous to (28), where the matrix $\bar{P} \equiv \mathrm{Cov}[\bar{X}_t - \bar{X}_{t|t-1}]$ is the covariance matrix for the prediction errors $\bar{X}_t - \bar{X}_{t|t-1}$. The matrix $\bar{P}$ in turn satisfies the Riccati equation

$$\bar{P} = \bar{H}[\bar{P} - \bar{P}\bar{L}'(\bar{L}\bar{P}\bar{L}' + \Sigma_{vv})^{-1}\bar{L}\bar{P}]\bar{H}' + \bar{\Sigma}_{uu} \tag{67}$$

analogous to (29), where

$$\bar{\Sigma}_{uu} \equiv \begin{bmatrix} \Sigma_{uu} & 0 \\ 0 & 0 \end{bmatrix}.$$

These equations are considerably more complicated to solve than in the case of symmetric partial information, treated in Svensson and Woodford (2002a). First of all, the matrices $\bar{K}$ and $\bar{P}$ are larger (of dimension $(n_X + n_x) \times (n_X + n_x)$ rather than simply $n_X \times n_X$). But more importantly, the elements of the matrices $\bar{H}$ and $\bar{L}$ that appear in (63) do not depend solely upon the elements of A and D, as in the case of symmetric information. For the matrix G^1 appearing in the definitions of H and L is no longer given by (19), but instead depends upon the matrices of the system (56) and hence on the Kalman gain matrix $\tilde{K}$ (though we show below that it only depends on $\bar{K}$). The same is true of the matrices Γ^1, S^1 and Σ^1 that enter into $\bar{H}$. Thus, one must solve a system of simultaneous equations for $\bar{P}$, $\bar{K}$, and the matrices that define the solution to (56); it is no longer possible to solve separately for the coefficients describing optimal policy on the one hand and optimal filtering on the other. This simultaneity is the same as we found in section 2, except that now the matrices involved are all of larger dimension.

The definitions of the matrices in (56) refer not only to $\bar{K}$ but also to the other elements of the Kalman gain matrix $\tilde{K}$. These, however, are easily expressed as functions of $\bar{K}$. Let us decompose $\tilde{K}$ with obvious notation,

$$\tilde{K} \equiv \begin{bmatrix} \tilde{K}_X \\ \tilde{K}_x \\ \tilde{K}_{\varphi 1} \\ \tilde{K}_{\varphi 2} \end{bmatrix},$$

so that

$$\bar{K} \equiv \bar{I}\tilde{K} \equiv \begin{bmatrix} \tilde{K}_X \\ \tilde{K}_{\varphi 2} \end{bmatrix}.$$

Then by (45) the Kalman gain matrix for the forward-looking variables, x_t, is given by

$$\tilde{K}_x = \bar{G}\bar{K}.$$

Similarly, the unique bounded solution to (42) allows us to write $\varphi_{1t|t}$ as a matrix times $\bar{X}_{t|t}$, where the matrix is the same as in the case of full information (or certainty). This equation implies that the Kalman gain matrix for the prediction equation for φ_{1t}, $\tilde{K}_{\varphi 1}$, will be the same matrix times $\bar{K}$. Thus the only simultaneity that must be dealt with is the dependence of the matrices G^1, Γ^1, S^1 and Σ^1 upon $\bar{K}$.

Finally, we note that our solution (64) for the evolution of the estimate $\Xi_{t-1|t}$ does not allow us to solve for this as a function of the history of estimates of the vector of predetermined variables alone, as in (2), let alone with coefficients that are independent of the information structure. Thus an assumption that optimal policy under commitment can be described by a reaction function in integrative form (5), with coefficients $\{F_j\}$ that are as the same as under full information, would represent an invalid application of the principle of certainty equivalence in this context. The application sketched in the next section gives an example of how certainty equivalence in this stronger sense may fail.

4 Application: Optimal monetary policy with real-time data

As an example of an application of our results, we sketch here the analysis of Aoki (2000a, 2000b) of the way in which optimal monetary policy should make use of the preliminary estimates of current macroeconomic conditions that are available in real time. As Orphanides (1998) in particular has stressed, policy must be made on the basis of preliminary estimates that are in fact revised substantially in subsequent months. Aoki models this by assuming that the available measures of current inflation and output are subject to measurement error, while the true values come to be known with a delay. The question that he considers is how the degree of uncertainty about the current state of the economy should affect the degree to which policy responds to the available real-time measures. Here we indicate how Aoki's analysis could be cast in terms of our framework.

Aoki assumes that output y_t and inflation π_t are determined by a model consisting of two structural equations,

$$y_t = E_t y_{t+1} - \sigma[i_t - E_t\pi_{t+1} - \rho_t], \tag{68}$$

$$\pi_t = \kappa(y_t - y_t^n) + \delta E_t\pi_{t+1}, \tag{69}$$

(where σ, κ and δ are positive and $\delta < 1$)which can be interpreted as an intertemporal "IS equation" and "aggregate supply equation" respectively. Here the short-term nominal interest rate i_t is assumed to be the central bank's policy instrument. The dependence of both relations upon expectations regarding future output and inflation introduces the forward-looking elements into the model's structural relation that complicates the central bank's optimal filtering and control problems.[17] Note that the expectations in these relations are conditional upon the full information set of the private sector.

The exogenous disturbances ρ_t and y_t^n are assumed to evolve as first-order autoregressive processes,

$$\rho_t = \eta\rho_{t-1} + e_{\rho t}, \tag{70}$$

$$y_t^n = \nu y_{t-1}^n + e_{yt}, \tag{71}$$

where $0 < \eta, \nu < 1$ and the disturbances $e_{\rho t}, e_{y_t}$ are iid mean-zero normal random variables, with variance s_ρ^2 and s_y^2 respectively. The central bank is assumed to wish to minimize a discounted loss function (13), where the period loss function is of the form

$$L_t = \pi_t^2 + \lambda(y - y_t^n)^2, \tag{72}$$

for a certain weight $\lambda > 0$. This is a standard specification of "flexible inflation targeting" (see

[17] See Woodford (2001) for the derivation of these relations from the underlying decision problems of households and firms, and for further discussion of the model.

Svensson, 1999), and can also be justified as a utility-based welfare criterion given the individual decision problems underlying relations (68)–(69) (see Woodford, 2002).

Because the "natural rate of output" around which one wishes to stabilize output according to (72) is exactly the shift factor in the aggregate supply relation (69), the aggregate supply relation itself implies no conflict between the goals of inflation and output-gap stabilization. It is possible in principle to completely stabilize both target variables, since $\pi_t = 0$ and $y_t = y_t^n$ at all times is consistent with equation (69), and also with equation (68) as long as the nominal interest rate satisfies

$$i_t = r_t^n \equiv \rho_t - \frac{1-\nu}{\sigma} y_t^n. \tag{73}$$

That is, complete stabilization of both target variables is a possible equilibrium, as long as the interest rate i_t perfectly tracks the exogenous variation in the Wicksellian "natural rate of interest" r_t^n (discussed further in Woodford, 2001). Under full information on the part of the central bank, this would represent optimal policy, as this equilibrium achieves the theoretical minimum value of the loss function (zero each period).[18] However, the central bank's real-time information set may not allow it to estimate the current natural rate of interest with complete accuracy. In this case, complete stabilization is not feasible, and the central bank faces a tradeoff between the two goals of inflation stabilization and output-gap stabilization.

Aoki assumes that the central bank's information set when setting i_t consists of complete information about all predetermined variables known to the private sector at date $t-1$, plus noisy preliminary estimates of current inflation and output. These preliminary estimates are assumed to be given by

$$\pi_t^o = \pi_t + \varepsilon_{\pi t}, \tag{74}$$

$$y_t^o = y_t + \varepsilon_{yt}, \tag{75}$$

where the measurement-error terms $\varepsilon_{\pi t}$ and ε_{yt} are iid mean-zero normal random variables, with variance σ_π^2 and σ_y^2 respectively. These errors in the central bank's preliminary estimates are assumed to be distributed independently of the "fundamental" disturbances $e_{\rho t}$ and e_{yt}.

[18] That is, an optimal policy would bring about an equilibrium in which the nominal interest rate would vary in this way. This does not necessarily mean that the optimal policy is for the central bank to set its instrument according to the function (73) of the exogenous disturbances; for while such a reaction function would be *consistent* with the optimal equilibrium, it would also allow an extremely large class of other rational-expectations equilibria as well, which equilibria are less desirable. Other reaction functions, that specify feedback from inflation and output-gap outcomes as under the "Taylor rule", are equally consistent with the optimal equilibrium and can render equilibrium determinate as well (see Woodford, 2001). There is no uniquely optimal reaction function of this form, since reaction functions that differ in the way that they specify out-of-equilibrium behavior may imply the same set of equilibria.

Aoki's model falls within the general framework set out in the previous sections. The structural equations (68)–(69), together with laws of motion (70)–(71) for the disturbances, comprise a model of the form (7). Here the predetermined variables consist of a 2-vector of exogenous disturbances X_t^* and the forward-looking endogenous variables consist of a 2-vector x_t, where

$$X_t^* \equiv \begin{bmatrix} \rho_t \\ y_t^n \end{bmatrix}, \qquad x_t \equiv \begin{bmatrix} \pi_t \\ y_t \end{bmatrix},$$

and it is possible to write equations (68)–(69) in a form such that $\tilde{E} = I$. While it would be possible to write the model's equations in the form (7) with $X_t = X_t^*$, this notation would not allow us to express the central bank's period-t observable variables as a function of X_t, x_t and measurement error alone, as assumed in (10). In order to directly apply our above expressions for the optimal filtering problem, it is necessary to augment the vector of predetermined variables to include lagged values of the fundamental disturbances, so that

$$X_t = \begin{bmatrix} X_t^* \\ X_{t-1}^* \end{bmatrix}.$$

The system (7) then becomes a system of six difference equations, including two identities.

Aoki's loss function (72) also falls under our framework. This is a loss function of the form (9), where

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C_i = 0, \qquad W = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}.$$

Because $C_i = 0$, we must apply the approach described in appendix B rather than the one followed in the text, but we are still able to derive a system of first-order conditions for optimal policy under commitment of the form (42).

Finally, the assumed information structure falls under the framework set out above, if we let

$$Z_t \equiv \begin{bmatrix} Z_t^* \\ X_{t-1}^* \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} X_t^* \\ X_{t-1}^* \\ x_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix},$$

where Z_t^* and ε_t are the 2-vectors of the observations, $[\pi_t^o \; y_t^o]'$, and measurement errors, $[\varepsilon_{\pi t} \; \varepsilon_{yt}]'$, respectively. Under this specification, the central bank's information set at date t includes the entire history $\{X_{t-j}^*\}$ for $j \geq 1$, and so a complete description of the state of the world as of date $t-1$, but only the information about date t fundamentals that is contained in the noisy measurements of current inflation and output (the two elements of Z_t^*). Aoki assumes, as we do, complete information on the part of the private sector.

Our results in sections 2 and 3 can then be directly applied. Discretionary optimization by the central bank (treated by Aoki, 2000a) results in the interest rate satisfying a relation of the form (15). Furthermore, the matrix F in this relation is the same as under full information. This allows us to write the relation in the simpler form

$$i_t = f' X_{t|t}^*,$$

since the vector X_{t-1}^* may be omitted from the vector X_t in the case of full information.[19] Furthermore, the corresponding relation in the full-information case is given by (73), allowing us to identify the elements of the vector f'. Optimal policy under discretion is thus characterized by the relation

$$i_t = r_{t|t}^n. \tag{76}$$

A Kalman filter can then be used to describe the evolution of the central bank's optimal estimate of the current natural rate of interest.

Under an optimal commitment (treated by Aoki, 2000b), instead, the interest rate satisfies a more complex relation of the form

$$i_t = f' X_{t|t}^* + \phi' \Xi_{t-1}. \tag{77}$$

(This corresponds to the optimal reaction function in state-space form (44), given that the elements X_{t-1}^* of the vector of predetermined variable X_t are irrelevant in the case of full information, and that $\Xi_{t-1|t} = \Xi_{t-1}$ under Aoki's information structure. More precisely, X_t^* provides a complete description of the state of the economy in period t, and the only role of X_{t-1}^* is to provide information in estimating X_t^*.) Once again, the vectors f' and ϕ' are the same as in the case of full information. Since the optimal equilibrium under full information satisfies (73),[20] the vector f' is the same as in the case of discretionary optimization, and (77) may equivalently be written

$$i_t = r_{t|t}^n + \phi' \Xi_{t-1}. \tag{78}$$

Note that it would *not* be correct to argue on the basis of certainty equivalence that since $i_t = r_t^n$ is the optimal commitment under full information, the optimal reaction function with

[19] That is, there exists an equilibrium of this form, which is the only equilibrium under partial information that corresponds via the principle of certainty equivalence to the Markov-perfect equilibrium under full information. Aoki does not consider whether additional equilibria may be possible under partial information.

[20] Note that in the case of full information, (73) describes the result of central-bank optimization under *either* discretion or commitment—the optimal commitment happens in this case to be time-consistent.

incomplete information will satisfy $i_t = r^n_{t|t}$. The reason that the optimal commitment (73) can be written without any feedback from a Lagrange multiplier in the case of full information is not because $\phi' = 0$ for this model, but rather because $\Xi_{t-1} = 0$ at all times.[21] This is because under full information, complete stabilization of the target variables is feasible, and so there is no gain from relaxing the constraints imposed by the model's structural equations. But one can show (by computing the relevant eigenvectors of the matrix $\mathcal{M}$ in (42)) that $\phi' \neq 0$, so that the second term on the right-hand side of (78) cannot be omitted. Nor is it true that Ξ_{t-1} is zero most of the time, in the case of imperfect information on the part of the central bank. This is because when information is imperfect, it will generally not be possible for the central bank to adjust its interest-rate instrument (which must depend only upon the central bank's information) so as to perfectly track variations in the natural rate of interest, and so perfect stabilization will be unattainable. In this case, there will generally be non-zero state-contingent Lagrange multipliers associated with perturbations of the constraints implied by the structural equations (68)–(69).

This does not contradict our certainty-equivalence results above, because these do *not* imply that the law of motion for the Lagrange multipliers must be the same as in the case of full information. It is true that (46) must hold, with the same matrices S and Σ as in the case of full information. In the case of full information, we know that $S = 0$, since Ξ_t will be zero regardless of the realization of the exogenous disturbances X_t. Thus (46) reduces in Aoki's model to

$$\Xi_{t|t} = \Sigma \Xi_{t-1|t} = \Sigma \Xi_{t-1}, \tag{79}$$

for a certain matrix Σ. However, Σ is not a zero matrix. In the case of full information, the corresponding law of motion

$$\Xi_t = \Sigma \Xi_{t-1}$$

implies that $\Xi_t = 0$ for all t as a result of the initial condition $\Xi_{-1} = 0$, despite the fact that $\Sigma \neq 0$. But with incomplete information, the initial condition $\Xi_{-1} = 0$ no longer implies that $\Xi_t = 0$ as well at all later dates. Instead, it implies only that $\Xi_{t|0} = 0$ for all $t \geq 0$. The actual value of the Lagrange multiplier will be given by the law of motion

$$\Xi_t = S^1[X_t - X_{t|t}] + \Sigma \Xi_{t-1},$$

[21] Alternatively, in the case of full information, optimal policy also has a representation in integrative form (5), and this representation also takes the simple form $i_t = r^n_t$. Here there are no missing Lagrange multiplier terms, because one expresses the instrument setting as a function of the history of the state $\{r^n_{t-j}\}$. But, as noted above, certainty equivalence does not hold under commitment and asymmetric information for the reaction function in integrative form.

where $S^1 \neq 0$. Thus (78) makes i_t a function not only of the central bank's period-t estimate of the current state of the economy — and indeed, not only of the entire history of central-bank estimates of that state — but also of what it has learned by period t about its *past errors* in estimation of the economy's state.[22]

Specifically, Aoki (2000b) shows that the optimal commitment involves making interest rates higher than the current estimate of the natural rate if the bank now knows that it underestimated the natural rate in the past. The reason is that the private sector is aware of the bank's misperception, and if it understands that the bank will later correct its error, then even when the *current* short rate is below the natural rate, the private sector's expectation of higher *future* short rates will dampen the high demand (and hence inflationary pressure) that would otherwise result. Of course, the cost of fulfilling such a commitment later is a short-term interest rate that deviates to a greater extent from the natural rate at that time; but some use of this mechanism can nonetheless improve the bank's overall stabilization objectives. Optimal policy is thus not purely forward-looking, in the sense discussed in Woodford (2000). Indeed, the response of the bank's interest-rate instrument to fluctuations in the natural rate of interest will be more inertial than is the natural rate itself, just as in the analysis of Woodford (1999). In the case of a positive innovation in the natural rate, the average immediate increase in the nominal interest rate will be smaller, because of the bank's inability to identify the increase immediately given its imperfect observation of current conditions. At the same time, the increase in interest rates will last longer, because in later periods the central bank will respond to its initial underestimate of the natural rate by keeping interest rates higher than the natural rate at those later times. Interestingly, Aoki obtains these results in the case of asymmetric information without any need for the assumption of an interest-rate stabilization objective in the central bank's loss function, relied upon in the full-information analysis of Woodford (1999).

5 Conclusions

We have presented a complete characterization of optimal filtering and optimal instrument choice for a fairly general class of linear-quadratic Gaussian policy problems, in which the evolution of the state of the economy depends on the private sector's (rational) expectations regarding future conditions, both under discretion and under commitment. A principle of certainty equivalence has been shown to hold in the case of discretionary optimization, and also for an optimal

[22] Thus optimal policy does not have a representation of the form (5).

commitment, when this is characterized by a reaction function in state-space form. However, we have shown that not all representations of optimal policy that would be correct in the case of full information continue to be correct under asymmetric partial information, if only actual states are replaced by the central bank's optimal estimates of those states. In particular, under commitment, a representation of optimal policy in integrative form — in which the current instrument settings are expressed as functions of the history of the state vector — is possible under either full information or symmetric partial information, but a corresponding expression is not correct in the case of asymmetric partial information of the kind considered here.

We have also shown that it is no longer possible, in general, to separate the problems of optimization and estimation, solving for the optimal central-bank estimates of the current state in a way that is independent of the quadratic loss function that the central bank seeks to minimize. Nonetheless, Kalman filtering techniques can be adapted to deal with the more complex optimal filtering problem that arises in this case.

The particular kind of asymmetric partial information assumed here, in which the information set of the central bank is a subset of that of the private sector, obviously represents only one special case, albeit one of some interest. The treatment of optimal policy under asymmetric partial information of other kinds, such as the case of non-nested information sets, is left for future work.

A Optimization under discretion and certainty equivalence

Consider the decision problem to choose i_t in period t to minimize (13) under discretion, that is, subject to (7)–(11) and

$$
\begin{aligned}
i_{t+1} &= F_{t+1} X_{t+1|t+1} && (80) \\
x_{t+1|t+1} &= G_{t+1} X_{t+1|t+1}, && (81)
\end{aligned}
$$

where F_{t+1} and G_{t+1} are determined by the decision problem in period $t+1$.

For the full information case, Oudiz and Sachs (1985) have derived an algorithm for the discretionary equilibrium, which is further discussed in Backus and Driffill (1986) and Currie and Levine (1993).[23] Following Pearlman (1992), but with a more explicit proof, this appendix shows that this algorithm, appropriately adapted, is valid also for our asymmetric-information case.

[23] See Söderlind (1999) for a detailed presentation.

First, using (81), taking expectations conditional on the central bank's information in period t, I_t, of the upper block of (7), we get

$$x_{t+1|t} = G_{t+1}X_{t+1|t} = G_{t+1}(A_{11}X_{t|t} + A_{12}x_{t|t} + B_1 i_t). \tag{82}$$

Taking the expectation in period t of the lower block of (7), we get

$$\tilde{E}x_{t+1|t} = A_{21}X_{t|t} + A_{22}x_{t|t} + B_2 i_t \tag{83}$$

(recall that $\tilde{E}$ is a matrix and not the expectations operator). Multiplying (82) by $\tilde{E}$, setting the result equal to (83) and solving for $x_{t|t}$ gives

$$x_{t|t} = \tilde{A}_t X_{t|t} + \tilde{B}_t i_t, \tag{84}$$

where

$$\begin{aligned}
\tilde{A}_t &\equiv (A_{22} - \tilde{E}G_{t+1}A_{12})^{-1}(\tilde{E}G_{t+1}A_{11} - A_{21}), \\
\tilde{B}_t &\equiv (A_{22} - \tilde{E}G_{t+1}A_{12})^{-1}(\tilde{E}G_{t+1}B_1 - B_2)
\end{aligned}$$

(we assume that $A_{22} - \tilde{E}G_{t+1}A_{12}$ is invertible). Using (84) in the expectation of the upper block of (7) then gives

$$X_{t+1|t} = A_t^* X_{t|t} + B_t^* i_t, \tag{85}$$

where

$$\begin{aligned}
A_t^* &\equiv A_{11} + A_{12}\tilde{A}_t, \\
B_t^* &\equiv B_1 + A_{12}\tilde{B}_t.
\end{aligned}$$

Second, by (8) and (9), we can write

$$L_{t|t} = \begin{bmatrix} X_{t|t} \\ x_{t|t} \end{bmatrix}' Q \begin{bmatrix} X_{t|t} \\ x_{t|t} \end{bmatrix} + 2\begin{bmatrix} X_{t|t} \\ x_{t|t} \end{bmatrix}' U i_t + i_t' R i_t + l_t, \tag{86}$$

where

$$Q \equiv C'WC, \quad U \equiv C'WC_i, \quad R \equiv C_i'WC_i$$

$$l_t \equiv \mathrm{E}\left\{ \begin{bmatrix} X_t - X_{t|t} \\ x_t - x_{t|t} \end{bmatrix}' C^{1\prime}WC^1 \begin{bmatrix} X_t - X_{t|t} \\ x_t - x_{t|t} \end{bmatrix} |I_t \right\}. \tag{87}$$

Using (84) in (86) leads to

$$L_{t|t} = X_{t|t}'Q_t^* X_{t|t} + 2X_{t|t}'U_t^* i_t + i_t' R_t^* i_t + l_t, \tag{88}$$

where

$$
\begin{aligned}
Q_t^* &\equiv Q_{11} + Q_{12}\tilde{A}_t + \tilde{A}_t' Q_{21} + \tilde{A}_t' Q_{22}\tilde{A}_t, \\
U_t^* &\equiv Q_{12}\tilde{B}_t + \tilde{A}_t' Q_{22}\tilde{B}_t + U_1 + \tilde{A}_t' U_2, \\
R_t^* &\equiv R + \tilde{B}_t' Q_{22}\tilde{B}_t + \tilde{B}_t' U_2 + U_2' \tilde{B}_t,
\end{aligned}
$$

and Q and U are decomposed according to $X_{t|t}$ and $x_{t|t}$.

Third, since the loss function is quadratic and the constraints are linear, it follows that the optimal value of the problem will be quadratic. In period $t+1$ the optimal value will depend on the estimate $X_{t+1|t+1}$ and can hence be written $X_{t+1|t+1}' V_{t+1} X_{t+1|t+1} + w_{t+1}$, where V_{t+1} is a positive semidefinite matrix and w_{t+1} is a scalar. Then the optimal value of the problem in period t is associated with the positive semidefinite matrix V_t and the scalar w_t, and fulfills the Bellman equation

$$
X_{t|t}' V_t X_{t|t} + w_t \equiv \min_{i_t} \left\{ L_{t|t} + \delta \mathrm{E}[X_{t+1|t+1}' V_{t+1} X_{t+1|t+1} + w_{t+1} | I_t] \right\}, \tag{89}
$$

subject to (85) and (88). Indeed, the problem has been transformed to a standard linear regulator problem without forward-looking variables, albeit in terms of $X_{t|t}$ and with time-varying parameters. The first-order condition is, by (88) and (89),

$$
\begin{aligned}
0 &= X_{t|t}' U_t^* + i_t' R_t^* + \delta \mathrm{E}[X_{t+1|t+1}' V_{t+1} B_t^* | I_t] \\
&= X_{t|t}' U_t^* + i_t' R_t^* + \delta (X_{t|t}' A_t^{*\prime} + i_t' B_t^{*\prime}) V_{t+1} B_t^*.
\end{aligned}
$$

Here we have assumed that l_t is independent of i_t, which assumption is verified below. The first-order condition can be solved for the reaction function

$$
i_t = F_t X_{t|t}, \tag{90}
$$

where

$$
F_t \equiv -(R_t^* + \delta B_t^{*\prime} V_{t+1} B_t^*)^{-1} (U_t^{*\prime} + \delta B_t^{*\prime} V_{t+1} A_t^*)
$$

(we assume that $R_t^* + \delta B_t^{*\prime} V_{t+1} B_t^*$ is invertible). Using (90) in (84) gives

$$
x_{t|t} = G_t X_{t|t},
$$

where

$$
G_t \equiv \tilde{A}_t + \tilde{B}_t F_t.
$$

Furthermore, using (90) in (89) and identifying gives

$$V_t \equiv Q_t^* + U_t^* F_t + F_t' U_t^{*\prime} + F_t' R_t^* F_t + \delta \left(A_t^* + B_t^* F_t\right)' V_{t+1} \left(A_t^* + B_t^* F_t\right).$$

Finally, the above equations define a mapping from $(F_{t+1}, G_{t+1}, V_{t+1})$ to (F_t, G_t, V_t). The solution to the problem is a fixpoint (F, G, V) of the mapping. It is obtained as the limit of (F_t, G_t, V_t) when $t \to -\infty$. The solution thus fulfills the corresponding steady-state matrix equations. Thus, the instrument i_t and the estimate of the forward-looking variables $x_{t|t}$ will be linear functions, (15) and (16), of the estimate of the predetermined variables $X_{t|t}$, where the corresponding F and G fulfill the corresponding steady-state equations. In particular, G will fulfill (17).

It also follows that F, G and V only depend on A, B, C, C_i, $\tilde{E}$, W and δ and are independent of D, Σ_{uu} and Σ_{vv}. This demonstrates the certainty equivalence of the discretionary equilibrium.

It remains to verify the assumption that l_t in (87) is independent of i_t. Since by (18), $x_t - x_{t|t} = G^1(X_t - X_{t|t})$, it is sufficient to demonstrate that $\mathrm{E}[(X_t - X_{t|t})(X_t - X_{t|t})'|I_t]$ is independent of i_t. By (27),

$$X_t - X_{t|t} = X_t - X_{t|t-1} + K[L(X_t - X_{t|t-1}) + v_t] = (I + KL)(X_t - X_{t|t-1}) + K v_t.$$

Since X_t and $X_{t|t-1}$ are predetermined and v_t is exogenous, the assumption is true.

B Appendix: The case of a loss function independent of the instruments

Here we consider the extension of our results to the special case in which $C_i = 0$, so that the policymaker's objective function is independent of the path of the instruments i_t. In this case, $L_{ii} = 0$, so that the matrix is necessarily not invertible, as assumed in the text. In this case, we cannot solve (38) for i_t. However, in this case we also must have $L_{iy} = 0$; hence there is no need to solve (38) in order to eliminate i_t from (37), as the equation is already in the desired form. We can simply take the conditional expectation of (37) with respect to I_t and obtain

$$A' \varphi_{t+1|t} = -L_{yy} y_{t|t} + \delta^{-1} \tilde{I}' \varphi_{t|t}, \tag{91}$$

which is of the same form as the upper half of the system of equations (41). The sticking point is that we are unable to eliminate i_t from (7), as is also necessary in order to derive the system (41) in the text.

It is nonetheless possible to derive a system of difference equations of the same general form. We note that in the present case, (38) reduces to

$$B' \varphi_{t+1|t} = 0. \tag{92}$$

This is a set of restrictions of the same form as those in the system (41), and independent of (37). Let us suppose that B is of full rank;[24] then (92) is a system of n_i linearly independent restrictions, where n_i is the number of instruments. Let g be an $(n_X + n_x) \times (n_X + n_x - n_i)$ matrix, the columns of which all linearly independent of one another, and orthogonal to each of the columns of B (so that $g'B = 0$). Then premultiplying (7) by g' and taking the conditional expectation with respect to I_t, we obtain

$$g' \tilde{I} y_{t+1|t} = g' A y_{t|t}. \tag{93}$$

This provides an additional $n_X + n_x - n_i$ linearly independent restrictions, and equations (91), (92) and (93) jointly comprise a system of the form (41). In the generic case, the matrix on the left-hand side will be invertible as assumed in the text, and one will obtain a system of the form (42).

[24] This is purely a notational convenience, as our method here can be applied even when it is not. One simply must eliminate redundant equations from (92), and augment the number of columns in the matrix g accordingly.

References

[1] Aoki, K., 2000a. On the optimal monetary policy response to noisy indicators, working paper, Kobe University.

[2] Aoki, K., 2000b. An optimal commitment plan under incomplete information, working paper, Kobe University.

[3] Backus, D., Driffill, J., 1986. The consistency of optimal policy in stochastic rational expectations models, CEPR Discussion Paper no. 124.

[4] Chow, G.C., 1975. Analysis and Control of Dynamic Economic Systems (John Wiley & Sons, New York).

[5] Currie, D., Levine, P., 1993. Rules, Reputation and Macroeconomic Policy Coordination (Cambridge University Press, Cambridge).

[6] Dotsey, M., Hornstein, A., 2000. Is the behavior of money useful to a discretionary policymaker? working paper, Federal Reserve Bank of Richmond.

[7] Kalchbrenner, J.H., Tinsley, P.A., 1975. On the use of optimal control in the design of monetary policy, Special Studies Paper no. 76, Federal Reserve Board.

[8] LeRoy, S.F., Waud, R.N., 1977. Applications of the Kalman filter in short-run monetary control, International Economic Review 18, 195–207.

[9] Lucas, R.E., Jr., 1973. Some international evidence on output-inflation tradeoffs,American Economic Review 63, 326-334.

[10] Orphanides, A., 1998. Monetary policy rules based on real-time data, Finance and Economics Discussion Series no. 1998-03, Federal Reserve Board.

[11] Oudiz, G., Sachs, J., 1985. International policy coordination in dynamic macroeconomic models, In: Buiter, W.H., Marston, R.C., eds., International Economic Policy Coordination (Cambridge University Press, Cambridge).

[12] Pearlman, J.G., 1986. Diverse information and rational expectations models, Journal of Economic Dynamics and Control 10, 333–338.

[13] Pearlman, J.G., 1992. Reputational and nonreputational policies under partial information, Journal of Economic Dynamics and Control 16, 339–357.

[14] Pearlman, J.G., Currie, D., Levine, P., 1986. Rational expectations models with partial information, Economic Modelling 3, 90–105.

[15] Sargent, T.J., 1991. Equilibrium with signal extraction from endogenous variables, Journal of Economic Dynamics and Control 15, 245–273.

[16] Söderlind, P., 1999. Solution and estimation of RE macromodels with optimal policy, European Economic Review 41, 1111–1146.

[17] Svensson, L.E.O., 1999. Inflation targeting as a monetary policy rule, Journal of Monetary Economics 43, 607-654.

[18] Svensson, L.E.O., Woodford, M., 2002a. Indicator variables for optimal policy, working paper, Princeton University.

[19] Svensson, L.E.O., Woodford, M., 2002b. Optimal policy with partial information in a forward-looking model: certainty equivalence redux, working paper, Princeton University.

[20] Townsend, R.M., 1983. Forecasting the forecasts of others, Journal of Political Economy 91, 546–588.

[21] Woodford, M., 1999. Optimal monetary policy inertia, NBER Working Paper no. 7261.

[22] Woodford, M., 2000. Pitfalls of forward-looking monetary policy, American Economic Review 90(2), 100-104.

[23] Woodford, M., 2001. A neo-Wicksellian framework for the analysis of monetary policy, working paper, Princeton University.

[24] Woodford, M., 2002. Inflation stabilization and welfare, Contributions to Macroeconomics 2, paper 1.