ISSN 1745-8587

Birkbeck
UNIVERSITY OF LONDON

School of Economics, Mathematics and Statistics

Birkbeck Working Papers in Economics & Finance

BWPEF 0719

The Endogenous Kalman Filter

Brad Baxter
Birkbeck,University of London

Liam Graham
University College London

Stephen Wright
Birkbeck, University of London

April 2007

▪ Birkbeck, University of London ▪ Malet Street ▪ London ▪ WC1E 7HX ▪

The endogenous Kalman filter.

Brad Baxter*, Liam Graham† and Stephen Wright‡

12 April 2007

Abstract

We relax the assumption of full information that underlies most dynamic general equilibrium models, and instead assume agents optimally form estimates of the states from an incomplete information set. We derive a version of the Kalman filter that is endogenous to agents' optimising decisions, and state conditions for its convergence. We show the (restrictive) conditions under which the endogenous Kalman filter will at least asymptotically reveal the true states. In general we show that incomplete information can have significant implications for the time-series properties of economies. We provide a Matlab toolkit which allows the easy implementation of models with incomplete information.

JEL Classification: E27; E37

Keywords: Dynamic general equilibrium; Kalman filter; imperfect information; signal extraction

1 Introduction

Underlying most dynamic general equilibrium models is the strong assumption that agents can perfectly observe the state variables. In the recent literature a growing number of papers[1] assume instead that agents use the Kalman Filter to obtain estimates of the states from the incomplete information available to them. Most of this work assumes that incomplete information applies only to exogenous variables. Only a few of these papers (Pearlman et al, 1986; Pearlman, 1992; Svensson & Woodford, 2003, 2004), have examined the signal extraction problem when it is endogenous to agents' decisions.

In this paper we present a new derivation of this endogenous version of the Kalman Filter. We distinguish between two potential types of endogeneity, firstly when some state variables are at least dynamically endogenous (i.e., are affected

*School of Economics, Maths & Statistics Birkbeck College, University of London, Malet Street, London W1E 7HX, UK. b.baxter@bbk.ac.uk

†Corresponding author: Department of Economics, University College London, Gower Street, London WC1E 6BT, UK. Liam.Graham@ucl.ac.uk

‡School of Economics, Maths & Statistics Birkbeck College, University of London, Malet Street, London W1E 7HX, UK. s.wright@bbk.ac.uk

[1] Aoki (2003, 2006), Bomfim (2001), Collard and Dellas (2006), Keen (2004), Svennson and Woodford (2003, 2004)

by the filtering problem with a time lag), and secondly when observable variables may themselves be affected by the optimising response to the filtering problem. We then show how the endogenous problem can be related to a parallel problem in which the states are exogenous, and hence standard formulae can be applied. By exploiting the nature of the parallel problem we obtain a number of results:

1. We state conditions under which the endogenous Kalman filter will converge to a unique steady state, even when (as will usually be the case) some of the state processes in the parallel problem are explosive.

2. The explosive nature of the parallel problem implies that however poor the information on the economy, it is always optimal to respond to it.

3. Incomplete information can have only transitory impacts but these can be highly persistent, and introduce new sources of dynamics in response to structural shocks. However, by their nature, these additional dynamics will not be observable in real time.

4. The only observable impulse responses will be those that can be related to the dynamics of estimated, as opposed to actual states. The estimated states follow the same vector autoregressive process as the states in a full information economy, but with a different covariance pattern of shocks. This can introduce significant differences in time series properties. In particular, estimated states will usually be subject to "shocks" to estimates of pre-determined variables like capital, that are logically impossible under full information.

5. Since full information is such an important benchmark assumption, we state the minimum conditions under which the endogenous Kalman Filter will at least asymptotically converge on full information. In an economy with one or more pre-determined state variables (implying that the "stochastic dimension" of the economy is less than its "state dimension") the number of observable variables required may sometimes be quite low; however we show that the *nature* of the measurement process, and the way it interacts with state dynamics, is also crucial.

Further, we provide a Matlab toolkit which allows the easy application of our techniques to a wide class of linear models.

Models involving the endogenous Kalman filter are particularly interesting in the light of our related work (Graham and Wright, 2007) in which we show that, while complete markets imply full information, incomplete markets in general imply incomplete information. So any work which involves incomplete markets needs to address the issue of how agents acquire and use information.

The remainder of the paper is organised as follows. Section 2 states the general form of the endogenous Kalman filter, and derives its properties. In Section 3 we present a simple analytical example that illustrates some of the features of the endogenous Kalman filter. Section 4 discusses some of the implications, and possible future applications, of our results. Appendices provide proofs and algebraic derivations.

2 The signal extraction problem in stochastic dynamic general equilibrium

2.1 A general system representation

A general linearised dynamic stochastic general equilibrium model can be written, following McCallum (1998) as:

$$A_{yy}E_t y_{t+1} = B_{yy}y_t + B_{yk}k_t + B_{yz}z_t \quad (1)$$
$$k_{t+1} = B_{ky}y_t + B_{kk}k_t + B_{kz}z_t \quad (2)$$
$$z_{t+1} = B_{zz}z_t + \zeta_{t+1} \quad (3)$$

In the first block of equations y_t is a $q \times 1$ vector of non-predetermined variables. The matrix A_{yy} may not be invertible. The second block describes the evolution of an $r_k \times 1$ vector of predetermined variables, while the third describes the evolution of an $r_z \times 1$ vector of exogenous variables that are assumed to follow a first order vector autoregression, with ζ_t an $r_z \times 1$ vector of innovations with covariance matrix $S_{zz} = E\left(\zeta_t \zeta_t'\right)$ which we assume is full rank[2]

We assume that agents form expectations based on an information set $I_t = \{\{i_{t-j},\ j \geq 0\};\Xi\}$ that evolves by

$$i_t = C_{ik}k_t + C_{iz}z_t + C_{iy}y_t + C_{iw}w_t \quad (4)$$

where i_t is an $n \times 1$ vector of observed variables, Ξ contains the (time-invariant) structure and parameters of equations (1) to (4), and w_t is an $r_w \times 1$ vector of measurement errors, with $0 \leq r_w \leq n$.[3] For generality we can in principle allow these to be serially correlated by representing them as a vector autoregression of the form

$$w_{t+1} = B_{ww}w_t + \omega_{t+1} \quad (5)$$

where ω_t has the full rank covariance matrix $S_{\omega\omega} = E\left[\omega_t \omega_t'\right]$. We assume that the eigenvalues of B_{ww} have real parts less than or equal to unity. The two innovations ω_t and ζ_t may in principle be contemporaneously correlated, with $E\left(\zeta_t w_t'\right) = S_{\zeta\omega}$, but are assumed uncorrelated at all other leads and lags.

Informational restrictions may arise, as in, e.g. Svensson & Woodford (2003) and Pearlman (1992), where a policymaker sets policy variables with incomplete information on the underlying state variables in the economy, or, as in Bomfim (2001), Keen (2004*),* Collard & Dellas (2006) where representative consumers are assumed to face informational restrictions.[4] Graham & Wright (2007) show that a filtering problem of the same general form can also arise in an incomplete

[2] Higher order VARMA representations of exogenous variables may in principle be captured by including lags of z_t and current or lagged values of u_t in k_t, and allowing z_{t+1} to depend on k_t. With this small amendment there is no loss of generality in assuming that S is full rank.

[3] Measurement errors may be of lower dimension than the measured variables themselves, if, for example, some linear combination of k_t, z_t and y_t is measured without error, or if measurement errors in different variables are systematically related.

[4] As in Svensson & Woodford (2003) we consider here only the case where the policymaker and the private sector have a common information set.

markets version of the stochastic growth model, where heterogeneous agents face a symmetric filtering problem of inferring aggregate magnitudes from the disaggregated prices they themselves observe directly.

A Matlab toolkit, provided as a companion to this paper, takes as input a system in the form specified in equations (1) to (5), and implements all the transformation and solution methods that follow.[5]

2.2 The transformed filtering problem

For compactness of notation we incorporate predetermined and exogenous variables, k_t and z_t, together with measurement error, w_t, into a vector of state variables of dimension $r = r_k + r_z + r_w$. In Appendix A we show that we can then use (2) to (4) to derive the following compact representation of the state evolution and measurement equations:

$$\xi_{t+1} = F_\xi \xi_t + F_c c_t + v_{t+1} \tag{6}$$

$$i_t = H'_\xi \xi_t + H_c c_t \tag{7}$$

where

$$\xi_t = \begin{bmatrix} k_t \\ z_t \\ w_t \end{bmatrix}; \; v_t = \begin{bmatrix} 0 \\ \zeta_t \\ \omega_t \end{bmatrix}; \; Q = E\left(v_t v'_t\right) = \begin{bmatrix} 0_{r_k \times r_k} & 0_{r_k \times s} \\ 0_{s \times r_k} & S \end{bmatrix}; \; S = \begin{bmatrix} S_{\zeta\zeta} & S_{\zeta\omega} \\ S'_{\zeta\omega} & S_{\omega\omega} \end{bmatrix}$$

and where c_t (a sub-vector of y_t) is a vector of dynamic choice variables such as consumption or policy variables that satisfy expectational difference equations. Precise definitions of the matrices in (6) and (7) are given in Appendix A.

By contrast, in the standard derivation of the Kalman filter (e.g. Harvey, 1981; 1989; Hamilton, 1994) the system is written as

$$\begin{aligned} \xi_{t+1} &= F_\xi \xi_t + v_{t+1} \\ i_t &= H'_\xi \xi_t + w_t \end{aligned}$$

where ξ_t are the state variables. Comparing this with our system in (6) and (7) reveals a number of differences that are of significance for our results.

First, in our system both equations depend on the dynamic choice variables c_t. An important distinction between the two forms of endogeneity is that the states have only lagged dependence on c_t (the most obvious example being the impact of current consumption decisions on future capital) while the measured variables may have contemporaneous dependence (for example via intratemporal optimality conditions). Both forms of endogeneity have important implications for the nature of the endogenous Kalman Filter, and for the consequences of the filtering problem for the economy in which it takes place.

Second, in standard applications of the Kalman Filter, where the ξ_t are

[5] In implementing the toolkit, it should be borne in mind that, as we show in Appendix A, the nature of the process assumed for the non-predetermined variables in (1) may have implications for the nature of the information set (or vice versa).

exogenous state processes, it is typically assumed that these are either stationary or at worst may have unit roots. Thus the eigenvalues of the matrix F_ξ are assumed to be not greater than unity in absolute value. In contrast, in the Endogenous Kalman Filter problem generated by a typical dynamic stochastic general equilibrium model, F_ξ may have at least one explosive eigenvalue, due to the dynamics of capital under dynamic efficiency (see for example, Campbell, 1994). We shall show that this feature interacts in interesting ways with the endogeneity of the Kalman Filter (and indeed *requires* some endogeneity, if any potentially explosive roots are to be stabilised).

Third, Q, the covariance matrix of the innovations of the redefined states (defined after (7)) is of rank $s = r_z + r_w \leq r$, with the inequality holding in strict form when there are pre-determined variables (if $r_k > 0$).

Fourth, the measurement errors, w_t, have been absorbed into the redefined states, ξ_t. This allows us to accommodate, in principle, both serial correlation of measurement errors and contemporaneous correlation with the structural innovations, u_t.

While we have derived the filtering problem from a standard dynamic stochastic general equilibrium model, in which the different elements of the state variables have a clear interpretation (and imply a number of restrictions on the structure of the problem) in what follows the only features of the system in (6) and (7) that are crucial to our results are the overall “state dimension", given by r, the "stochastic dimension" given by $s \leq r$, and the number of measured variables, $n \leq r$ (where in most of what follows we shall assume that both inequalities hold in strict form), along with the endogeneity of both states and measured variables to the dynamic choice variables, c_t. Thus in principle the results that follow may apply to a wider class of models that fit within the general framework of (6) and (7).

2.3 Full information solution

As a first stage in our derivation we derive the solution for the special case of full information, which provides a crucial analytical building block for the more general solution under incomplete information.

The full information case is a special case of the system (6) and (7) with $n = r$, $H_\xi = I_r, H_c = 0$ implying $i_t = \xi_t$, so that the Kalman Filter is, trivially, redundant.

The solution for the dynamic choice variables under full information can be expressed in the form

$$c_t^* = \eta' \xi_t^* \tag{8}$$

where all elements in the ith row of η are zero for $i > r_k + r_z$ (measurement errors have no impact on under full information), and for any variable x_t, x_t^* denotes its value under full information. The matrix η can be computed using standard techniques (for example Blanchard and Kahn, 1980 ; McCallum, 1998). For the rest of the paper we treat it as a parameter. Given (8), the full information states follow a first order vector autoregressive process in reduced form:

$$\xi^*_{t+1} = G\xi^*_t + v_{t+1} \tag{9}$$

where

$$G = F_\xi + F_c\eta' \tag{10}$$

The behaviour of the dynamic choice variables, c_t, is crucial for the stability of the states, under both full and incomplete information. As noted in the previous section, in a model with endogenous capital F_ξ will usually have at least one explosive eigenvalue. This latent explosive property can only be controlled by the behaviour of the dynamic choice variables. Under full information this stabilization follows directly from the standard rational expectations solution. Under standard conditions the matrices η and F_c (8) always satisfy the following conditions:

Assumption 1. *All the eigenvalues of the matrix $G = F_\xi + F_c\eta'$ have real parts less than or equal to unity*

Assumption 2. *Let $G = V\Lambda V^{-1}$ where Λ is a diagonal matrix of eigenvalues and V the corresponding matrix of eigenvectors. For any strictly unit eigenvalue in Λ the corresponding row of F_c is zero.*

Assumption 1 rules out explosive rational expectations solutions; Assumption 2 states that, to the extent that any innovations have permanent effects, these are innovations to strictly exogenous processes (e.g., there may be a unit root component in technology).[6]

These features of the solution under full information turn out to be equally crucial for the stability of the solution under incomplete information.

2.4 Indirect observability

The full information solution can also be replicated straightforwardly in the special case that $n = r$, and the matrix H_ξ in the measurement equation (7), (which is therefore square) is also invertible. In this case the state variables can be replaced in the state equation by setting $\xi_t = \left(H'_\xi\right)^{-1}(i_t - H_c c_t)$ and can therefore be treated as known. Mehra and Prescott (1980) refer to this case when they write "...the state variables are observed, or are an invertible function of observables...". In this, quite restrictive, special case the Kalman Filter is again redundant.

2.5 Incomplete information solution: the Endogenous Kalman Filter

For the general case we need to apply the Kalman Filter, but allowing for the endogeneity of the dynamic choice variables to the filtering process.

[6] Assumption 2 follows naturally from the underlying structural model, since in equations (1) and (2) the dependence of state variables on c_t is only via k_{t+1}, hence all elements of the ith row of F_c in equation (6) are zero for $i > r - s$.

Following Pearlman (1992) and Svennson and Woodford (2004) we conjecture that under incomplete information optimal choices will be certainty-equivalent:

$$c_t = \eta' \widehat{\xi}_t \tag{11}$$

where $\widehat{\xi}_t = E_t \xi_t | I_t$ is the optimal estimate of the current state vector[7] given the available information set I_t, which evolves as in (7) and η is identical to the matrix for the full information case in (8). We show below that this conjecture is verified.

We first define two key matrices that characterise the properties of the state estimates and state forecasts.

$$M_t = E\left[\left(\xi_t - \widehat{\xi}_t\right)\left(\xi_t - \widehat{\xi}_t\right)'\right] \tag{12}$$

is the covariance matrix of the filtering error in current state estimates, and

$$P_{t+1} = E\left[\left(\xi_{t+1} - E_t \xi_{t+1}\right)\left(\xi_{t+1} - E_t \xi_{t+1}\right)'\right] \tag{13}$$

is the covariance matrix of the one-step ahead state forecast errors.[8]

The nature of the solution to the endogenous Kalman Filter problem is summarised in the following proposition:

Proposition 1 *In the solution to the endogenous Kalman Filter problem given by (6) and (7), the mean squared error matrices P_{t+1} and M_t are identical to those derived from the parallel exogenous Kalman Filter problem*

$$\begin{aligned} \widetilde{\xi}_{t+1} &= F_\xi \widetilde{\xi}_t + v_t \\ \widetilde{i}_t &= H_\xi' \widetilde{\xi}_t \end{aligned}$$

(i.e., setting $F_c = H_c = 0$ in (6) and (7)). They are thus given by the standard exogenous Kalman filter recursion

$$P_{t+1} = F_\xi M_t F_\xi' + Q \tag{14}$$

$$M_t = \left[I_r - \widetilde{\beta}_t H_\xi'\right] P_t \tag{15}$$

$$\widetilde{\beta}_t = P_t H_\xi \left[H_\xi' P_t H_\xi\right]^{-1} \tag{16}$$

However, in the solution to the actual endogenous Kalman Filter problem, conditional upon the certainty-equivalent consumption function (11), the estimated states follow the process

$$\widehat{\xi}_{t+1} = G \widehat{\xi}_t + \beta_t \varepsilon_{t+1} \tag{17}$$

[7] For compactness of notation we write period $t's$ estimate of the states at t as $\widehat{\xi}_t$, the standard Kalman filter literature commonly uses $\widehat{\xi}_{t|t}$. For the forecast at time t of the states at period $t+1$ we write $E_t \xi_{t+1} \left(= E_t \widehat{\xi}_{t+1}\right)$ instead of the standard $\widehat{\xi}_{t+1|t}$

[8] P_{t+1} is commonly denoted $P_{t+1|t}$, and using the same notation $M_t = P_{t|t}$, but we separate the two for clarity.

where G is as defined in (10), ε_t is the innovation to the measured variables, given by

$$\varepsilon_{t+1} \equiv i_{t+1} - E_t i_{t+1} \tag{18}$$

and

$$\beta_t = \widetilde{\beta}_t \left[I_n + H_c \eta' \widetilde{\beta}_t \right]^{-1} \tag{19}$$

Proof. See Appendix B. ■

We noted above that the filtering problem set out in equations (6) and (7) displays two forms of endogeneity: the dependence of the states on the lagged dynamic control variables, via F_c in (6) and the dependence of the measured variables on the contemporaneous dynamic control variables, via H_c in (7). Proposition 1 states that the solution to this problem can be derived from the solution to a parallel filtering problem for a notional state process $\widetilde{\xi}_t$ and a notional set of measured variables $\widetilde{i}_t$ for which both effects are absent, so that the standard Kalman filter formulae can be applied.[9]

Proposition 1 shows that, conditional upon the solution for β_t in (19), the estimated states $\widehat{\xi}_t$ follow a first order vector autoregression given by (17) with the same non-explosive autoregressive matrix G as in the process of the true states under full information, in (9). In the parallel problem, in contrast, the notional state process $\widetilde{\xi}_t$ has autoregressive matrix $F_{\xi,}$ which, as noted above may have explosive eigenvalues. We shall see below that this rather unusual feature of the state process in the parallel problem has significant implications for the nature of information processing.

The intuition for this feature is that while the dynamic choice variables determine future states in the true problem via the matrix F_c this does *not* impact on one-step ahead uncertainty (since the marginal impact of today's choices on tomorrow's states is known today even if current states are unknown). As a result the expression for P_{t+1} only allows for the direct impact of uncertainty about today's states transmitting to uncertainty about tomorrow's states, via the matrix $F_{\xi.}$ Since the matrix F_c does not affect the solution to the filtering problem, it can be solved under the assumption that $F_c = 0$.

Further, the Kalman gain[10] matrix β_t for the true problem is not the same as its counterpart $\widetilde{\beta}_t$ in the parallel problem because the signal conveyed by innovations to the measured variables also affects the dynamic choice variables c_t. But this has no impact on the mean squared error matrices M_t and P_{t+1}, hence these can be derived under the assumption that $H_c = 0$.[11]

[9] Our formulae for P_{t+1} and $\widetilde{\beta}_t$ are more compact than the more common formulation, given our absorption of measurent error into the states, but can be easily shown to be identical.

[10] We use the definition of the Kalman gain as in Harvey (1981), in which it can be interpreted as a matrix of regression coefficients updating current state estimates in response to forecast errors in predicting measured variables. The term is also frequently applied (as in for example, Hamilton, 1994) to a matrix, often denoted K, that updates *forecasts* of the states in response **to** the same forecast errors. In the parallel exogenous problem $\widetilde{K} = F_\xi \widetilde{\beta}_t$ in our notation, however in the actual endogenous problem $K = G\beta_t$, since it would incorporate the endogenous response of dynamic choice variables both in β_t but also in the autoregressive representation in (17).

[11] In Appendix B we show that the solution to the filtering problem given by Proposition 1 is identical to that in Svensson & Woodford (2003) (albeit derived by a distinctly different

2.6 The steady state endogenous Kalman filter

Equations (14) to (16) are a set of recursive matrix equations, for which it is natural to look for a stable steady state. Standard proofs of convergence (see for example, those in Hamilton, 1994) cannot be applied given the presence of explosive eigenvalues in F_ξ. However, even in the presence of explosive eigenvalues, under conditions that will usually be satisfied in dynamic general equilibrium models, a unique stable state does exist:

Proposition 2 *If the parallel problem in Proposition 1 is stabilisable and detectable in the sense of Anderson & Moore (1979), then for any initial positive definite matrix P_0, a unique stable steady state endogenous Kalman Filter exists, with matrices β, P and M that satisfy the steady state of equations (14) to (16).*

Proof. See Appendix C ■

The twin conditions of stabilisability and detectability can both be related to the nature of the underlying shock processes driving the state process. If the innovations to the state process in (6) are expressed in the form

$$\begin{aligned} v_t &= F_u u_t \qquad (20) \\ \text{where } F_u &= \begin{bmatrix} 0_{(r-s)\times s} \\ I_s \end{bmatrix}; \;\; u_t = \begin{bmatrix} \zeta_t \\ \omega_t \end{bmatrix}; E\left(u_t u_t'\right) = S \end{aligned}$$

and ζ_t and ω_t are the innovations from equations (3) and (5) respectively, then the two conditions can be written as

$$\begin{aligned} \text{stabilisability:} \quad & |\lambda_i \left(F_\xi + F_u L_1'\right)| < \; 1 \;\; \forall i \\ \text{detectability:} \quad & |\lambda_i \left(F_\xi + L_2 H_\xi'\right)| < \; 1 \;\; \forall i \end{aligned}$$

where the two conditions are satisfied for some matrices L_1 and L_2 of dimensions $r \times s$ and $r \times n$ respectively, and where $\lambda_i(.)$ denotes the ith eigenvalue of a square matrix. Note that these conditions apply to the parallel problem, and hence are entirely unaffected by endogeneity in the true filtering problem.

The first condition is trivial if there are no pre-determined variables, and hence u_t the underlying innovations in (20) are of dimension $s = r$, since in that case F_u is a full rank $r \times r$ matrix. Where there are pre-determined variables ($s < r$) it is not so straightforward. In this case F_u is likely to contain a row of zeros in exactly the row corresponding to an explosive eigenvalue in F_ξ, so that the condition for stabilisability can only be met if the relevant row of F_ξ contains off-diagonal elements. A simple example might be that capital must depend not only on lagged consumption, but also on lags of stochastic exogenous state variables (for example technology).

The second condition is more straightforward: essentially it requires that there must be some observable indicator, however poor, of any state variables with associated explosive (or unit) eigenvalues.

approach) and very close to that in Pearlman et al (1986). However neither set of authors draws out the implications of the "parallel problem", that are crucial to our remaining results.

2.7 Time series implications of incomplete information

Conditional upon convergence of the endogenous Kalman filter, if we define the state filtering error as

$$f_t = \xi_t - \widehat{\xi}_t \tag{21}$$

then we show in Appendix D that the joint process for ξ_t and f_t under incomplete information can be expressed in the vector autoregressive form

$$\begin{bmatrix} \xi_{t+1} \\ f_{t+1} \end{bmatrix} = \begin{bmatrix} G & -F_c\eta' \\ 0 & \left[I - \widetilde{\beta}H'_\xi\right] F_\xi \end{bmatrix} \begin{bmatrix} \xi_t \\ f_t \end{bmatrix} + \begin{bmatrix} I \\ I - \widetilde{\beta}H'_\xi \end{bmatrix} v_{t+1} \tag{22}$$

The top block is entirely independent of filtering parameters, and transparently reduces to the full information process (9) when filtering error disappears. In general, however, filtering error "contaminates" state dynamics via the off-diagonal element of the autoregressive matrix for the joint process for ξ_t and f_t.

In contrast the process for the state filtering error f_t is block recursive. Furthermore, consistent with Proposition 1, it follows an identical process to the state filtering error in the parallel exogenous problem (i.e., it does not depend on F_c or H_c) and hence is also invariant to the properties of the c_t, the dynamic choice variables.

Proposition 2 has an important corollary that is crucial to the time series properties summarised in (22):

Corollary 3 *The matrices* $F_\xi\left(I - \widetilde{\beta}H'_\xi\right)$ *and* $\left(I - \widetilde{\beta}H'_\xi\right)F_\xi$ *have at most* $r - n$ *non-zero eigenvalues, that have real parts strictly less than unity in absolute value.*

Proof. See Appendix C. ■

Thus stability of the filtering problem automatically implies that the filtering error is stationary. The more persistent is the filtering error process (the closer are the non-zero eigenvalues of $\left(I - \widetilde{\beta}H'_\xi\right)F_\xi$ to unity), the more prolonged will be the additional dynamics introduced by the filtering problem.

The joint process for ξ_t and f_t also provides a complete description of the process for the dynamic choice variables c_t, since, using (11) we can write $c_t = \eta'\widehat{\xi}_t = \eta'(\xi_t - f_t)$. Thus under incomplete information the process for c_t differs from the process under full information both because of the direct effect of filtering error on the estimated states and because the true states differ from their full information values. Note that only filtering errors in the underlying states k_t and z_t have any direct impact on c_t since, as noted previously, η has zeros in its ith row for $i > r_k + r_z$.

The representation in (22) implies a number of key features of the endogenous Kalman Filter that are direct corollaries of Propositions 1 and 2, given Assumptions 1 and 2.

Corollary 4 *The incomplete information solution is non-explosive.*

Corollary 5 *Incomplete information has no permanent effects.*

Corollary 6 *The filtering errors f_t satisfy $H'_\xi f_t = 0$.*

Corollary 7 *Let $\beta = \begin{bmatrix} \beta_k & \beta_z & \beta_w \end{bmatrix}'$. If F_ξ has explosive eigenvalues, $\beta_k = 0$ can never be a convergent solution of the endogenous Kalman Filter problem.*

For proofs see Appendix D.

The first two of these results are unsurprising, if reassuring features of the incomplete information solution. Corollary 4 states that the stabilising features of the dynamic choice variables in the full information solution discussed in Section 2.3 follow through into the incomplete information solution. Corollary 5 states that impulse responses under full information and incomplete information must converge.

The third feature summarised in Corollary 6 implies that there is linear dependence between the elements of the vector of filtering errors, an issue we discuss further in Section 4.2 below. Since filtering errors are identical to those in the parallel process this linear dependence is unaffected by the endogeneity of the filtering problem.

Corollary 7 is a more distinctive feature of the endogenous Kalman Filter solution, with important implications for the nature of information processing: in essence, however poor the information set, it is always optimal to update estimates of predetermined variables.

Mathematically, this result follows directly from the stationarity of the filtering error process: when F_ξ has explosive eigenvalues associated with the evolution of the pre-determined variables, k_t, the autoregressive matrix of the filtering error, $\left(I - \widetilde{\beta} H'_\xi\right) F_\xi$ could not have stable eigenvalues with $\widetilde{\beta}_k = 0$. From (19) this in turn implies that in this case $\beta_k \neq 0$ can never be a solution of the filtering problem. This has quite significant implications for the nature of the optimal response to information, as the quality of that information deteriorates.

In standard exogenous Kalman filter problems, in which $F_c = 0$ and F_ξ usually has at worst borderline unit eigenvalues, the lower the quality of the information, the smaller is the optimal response to that information. As $S_{\omega\omega}$, the covariance matrix of structural measurement error innovations tends to infinity, β_k and β_z will both tend to zero, the state estimates $\widehat{k}_t$ and $\widehat{z}_t$ will tend to a constant, and the filtering errors for these state variables become the processes themselves.[12] In the endogenous Kalman Filter problem the same feature applies when F_ξ has stable or unit eigenvalues.[13]

In contrast, if F_ξ has explosive eigenvalues, the state process in the parallel problem set out in Proposition 1 also has explosive eigenvalues, so for β_k sufficiently close to zero, the filtering error process would itself be explosive, contradicting Corollary 3. In this case, as $S_{\omega\omega}$ tends to infinity, β_k tends to a

[12] In this case β_w will tend to I_{r_w} : all innovations to measured variables will be interpreted as due to measurement error. In the borderline case where F_ξ has strictly unit eigenvalues the filtering error process for the corresponding states will tend towards a unit root process.

[13] This feature is noted in Svensson & Woodford (2003, p711). It is essentially the basis for the principle of policy gradualism under uncertainty, originally noted by Brainard (1967).

fixed, non-zero matrix. Thus however poor the information, it is always optimal to respond to it.[14]

2.8 Certainty equivalence

Forecasts of the estimated states, given by (17) have an identical autoregressive form to forecasts of the *true* states under full information: i.e., under both incomplete and complete information

$$E_t \xi_{t+1} = E_t \widehat{\xi}_{t+1} = G \widehat{\xi}_t \tag{23}$$

where under full information, $\widehat{\xi}_t = \xi_t = \xi_t^*$. Since forecasts of neither estimated nor actual states depend on β, the Kalman Gain matrix, incomplete observability has no impact on optimal forecasts, which depend only on structural parameters. As a result the coefficient matrix η in the conjectured form for optimal choices under incomplete observability, (11) can be derived under the assumption of full information, i.e. as in (8). Thus optimal choices are certainty-equivalent, verifying our conjecture.[15]

2.9 Can incomplete information ever replicate full information?

We have already seen two special cases of our framework in which the states are either directly (Section 2.3) or indirectly (Section 2.4) observable, and thus the Kalman Filter is redundant. Both of these cases require that n, the number of measured variables, equal r, the number of states. However, when there are predetermined variables, and hence s, the "stochastic dimension" of the system, is less than r, the "state dimension", there may also be cases in which the Kalman Filter may yield state estimates that become arbitrarily close to the true states as the history of the information set increases over time. This feature can be defined formally in three ways that are all logically equivalent.

Definition 8 *An information set I_t is asymptotically revealing if*

$$\lim_{t\to\infty} M_t(I_t) = 0 \Longleftrightarrow \lim_{t\to\infty} P_t(I_t) = Q \Longleftrightarrow \lim_{t\to\infty} \widehat{\xi}_t(I_t) = \xi_t^*$$

where M_t and P_t satisfy the recursion in (14) and (15).

[14] See Section 3 for a simple example. Since this feature derives from the parallel problem it clearly also applies to any *ex*ogenous Kalman filtering problem where the autoregressive representation is explosive. We are not however aware of any discussion of this feature in the existing Kalman filter literature, presumably because explosive representations are so unusual. Harvey (1989) notes that Anderson & Moore's (1979) analysis dismisses even the borderline unit root case as of limited interest.

[15] Certainty equivalence arises naturally from the fact that we first linearise the model (including Euler equations) and then solve the filtering problem. To the extent that state uncertainty introduces new sources of variance in dynamic choice variables (an issue we discuss in Section 4.3) incorporation of state uncertainty into the optimisation problem before linearisation would presumably result in effects analogous to those in the precautionary consumption literature.

The definition explicitly notes the dependence, not only on the history of the observed variables, i_t, but also the structure of the model. If the information set satisfies this definition, M_t, the mean squared error matrix of state filtering errors converges to a steady state value of zero; P_t, the covariance matrix of one-period-ahead state forecast errors, tends to its irreducible minimum of Q, and estimated states converge on the true states. Since the impact of filtering error is, by Corollary 5 only transitory, ultimately the true states in turn must converge on their full information values.

This implies significant restrictions on the nature of the information set, which we summarise in our final proposition.

Proposition 9 *Assume that the Endogenous Kalman Filter of Proposition 1 satisfies the conditions for convergence given by Proposition 2. Necessary and sufficient conditions for an information set I_t to be asymptotically revealing are*

$$\begin{aligned} n &= s \\ \left|H'_\xi F_u\right| &\neq 0 \\ \lambda_i\left(\left(I - \widetilde{\beta}\left(Q\right)H'_\xi\right)F_\xi\right) &< 1 \ \forall i \end{aligned}$$

where: n is the number of observed variables; s =rank(S) =rank(Q) is the "stochastic dimension" of the state variables; H_ξ and F_u are as given in equations (7) and (20); $\lambda_i\left(A\right)$ are the eigenvalues of a matrix A; and $\widetilde{\beta}\left(Q\right)$ satisfies (16) setting $P = Q$.

Proof. See Appendix E. ■

The first and second conditions given in Proposition 9 are quite intuitive. However large r, the number of underlying state variables may be, uncertainty as to their true values must ultimately depend on s, the number of innovations actually driving the system.[16]. To see this, assume that all of the conditions in Proposition 9 are indeed satisfied. Since these restrictions relate only to the limiting properties of M_t and P_t, which from Proposition 1 are identical in both the true Endogenous Kalman Filter problem and its parallel problem, only the properties of the latter problem matter, so we can ignore the endogeneity of the states, and set $H_c = F_c = 0$. If we apply this restriction, then, using equations (18), (7), (6) and (20), with a sufficiently long history of observations of the measured variables, as state estimates converge on their true values,

$$\varepsilon_{t+1} \equiv i_{t+1} - E_t i_{t+1} = H'_\xi\left(\xi_{t+1} - E_t\xi_{t+1}\right) \rightarrow H'v_{t+1} = H'F_u u_{t+1}$$

So, if $H'F_u$ is square (hence if $n = s$) and invertible, then in the limit we can recover the true innovations from the observable innovations to the measured variables, and thus recover the states themselves.

However, these are necessary but not sufficient conditions. We have already seen, in Corollary 3 that convergence of the Kalman filter implies that the filter-

[16] In this context there is no logical distinction between measurement errors ω_t and structural errors ζ_t so we do not distinguish between the two.

ing error process, f_t in (22) must be stationary. The converse must also apply. The first two conditions imply that $P = Q$ is a steady state of the Kalman Filter, but they do not necessarily imply that it is a stable steady state. If the filtering errors in the neighbourhood of $P_t = Q$ are not stable, then this is equivalent to the statement that the Kalman Filter will not converge to this steady state: thus the third condition is also crucial.

The conditions set by Proposition 9 are particularly interesting because full information is the maintained assumption in virtually all analysis of DGE models, and it is useful to be reminded of the minimal informational assumptions necessary for this assumption to be valid. It is also quite frequently the case in DGE models that s, the "stochastic dimension" of an economic system, is indeed distinctly lower than r, the dimension of the states. Most notably, the benchmark stochastic growth model is typically driven by a single technology shock, but will have at least two state variables, technology itself and capital, and possibly more if there are other forms of inertia such as capital installation costs or sticky prices.[17] In these cases, we saw in Section 2.4 that for an information set to reveal the full set of states instantaneously requires that n, the number of observed variables, equals r, the number of states. However, it may reveal them asymptotically with even a single observable variable.

However, the third condition set by Proposition 9 means that for this to be the case requires more than simply counting the number of observable variables: the *nature* of the measurement process associated with that variable is crucial. In the filtering problem of a stochastic growth model with a single aggregate technology shock under incomplete information analysed by Graham & Wright (2007), for example, an information set with the history of just output or just wages satisfies all the conditions set by Proposition 9, and thus asymptotically reveals all states; whereas the history of just returns on capital satisfies the first two conditions, but not the third.

There is an interesting contrast between the case where an information set only asymptotically reveals the states, and the case of indirect observability discussed in Section 2.4. In that case, because the information set instantaneously reveals the states, the optimal values of the dynamic choice variables c_t can be represented as a static linear weighting of the r observable variables, given by

$$c_t^* = \mu' i_t^*; \quad i_t^* \in \mathbb{R}^{r\times 1}$$

where $\mu' = \eta' \left[I + \left(H'_\xi \right)^{-1} H_c \eta' \right] \left(H'_\xi \right)^{-1}$,whereas if an information set is only asymptotically revealing, the relationship tends in the limit to the form

$$c_t^* = \nu'(L)\, i_t^* \quad i_t^* \in \mathbb{R}^{s\times 1}$$

[17]Or, in a case analysed by Campbell (1994), if technology follows a higher order AR or ARMA process than the usual AR(1) assumption.

where $\nu(L)$ is a matrix polynomial in the lag operator,[18] and the observed variables are now only of dimension s. Thus while the number of observable variables is reduced, this is offset by a significant dynamic complexity in the nature of the relationship. This has potentially interesting empirical implications. Since both forms of the relationship represent optimising behaviour neither contains an error term. The case of indirect observability should therefore imply that any simple static regression equation that contains i_t^* amongst its regressors should fit the data perfectly. In the case of asymptotic revealing there should also be in principle be some dynamic equation that fits perfectly, but only if specified with sufficient (possibly infinitely many) lags, and with the correct parameter restrictions: it would therefore be a much harder hypothesis to test.

3 A simple analytical example

We illustrate some of the features set out above by looking at a very simple case where there is a single state variable, capital k_t and a single noisy signal of it, i_t, is observed. We allow the single dynamic choice variable, consumption, c_t to impact on both the state equation (with a lag) and on the measurement equation. We assume that k_t is hit by a white noise shock, u_t in each period so is not perfectly predictable from its own past and the past value of consumption.[19] The state and measurement equations are:

$$k_{t+1} = \lambda k_t - \mu c_t + u_{t+1} \tag{24}$$
$$i_t = k_t + hc_t + w_t \tag{25}$$

where $u_t \sim N(0, S)$ and $w_t \sim N(0, R)$ are two mutually and serially uncorrelated white noise processes, with S and R both scalars.

In solving the filtering problem we exploit Proposition 1 and solve the parallel filtering problem in which capital is exogenous (i.e., we can set $\mu = h = 0$), so that the problem reduces to an (almost) standard single exogenous state variable signal extraction problem. The only non-standard feature we wish to allow for is that λ, the coefficient on lagged capital in (24) may be greater than unity (a standard feature of linearised models under dynamic efficiency - see for example Campbell, 1994). Given the single state variable the certainty-equivalent consumption function has the very simple form

$$c_t = \eta_k \widehat{k}_t \tag{26}$$

[18] For the general case the endogenous Kalman filter implies

$$\nu(L) = \eta' [I - GL]^{-1} \beta \left[I + \left(H'_\xi + H_c \eta'\right) G (I - GL)^{-1} \beta L\right]^{-1}$$

where additionally, in the special case of asymptotic revealing, $\beta = F_u \left(H'_\xi F_u\right)^{-1}$

[19] In Appendix F we show that the process for capital can be derived from a production process in which new technology is white noise, and is only embodied in new capital. The parameters of the problem can be related directly to those in the linearised stochastic growth model of Campbell (1994)

where $\eta_k \in (0,1)$. In Appendix F we show that by respecifying (24) and (25) in the same form as (6) and (7) (thus incorporating the measurement error w_t into the states), applying the formulae from Proposition 1,and assuming that the Kalman Filter has converged[20] the updating equation for capital is

$$\widehat{k}_{t+1} = E_t\widehat{k}_{t+1} + \beta_k \left(i_{t+1} - E_t i_{t+1}\right)$$

where β_k, the first element of β, is given by

$$\beta_k = \frac{\widetilde{\beta}_k}{1 + h\eta_k\widetilde{\beta}_k}$$

and the Kalman gain for capital in the parallel problem, $\widetilde{\beta}_k$, has the usual form for a single exogenous state variable measured with noise:

$$\widetilde{\beta}_k = \frac{P_k}{P_k + R}$$

where P_k,the top left element of the matrix P (which in this case is 2×2) satisfies the steady state of the recursion

$$P_{kt+1} = \lambda^2 P_{kt}\left(\frac{R}{P_{kt} + R}\right) + S$$

implying that P_k solves the quadratic equation

$$\left(\frac{P_k}{R}\right)^2 + \left[1 - \lambda^2 - \frac{S}{R}\right]\frac{P_k}{R} - \frac{S}{R} = 0.$$

The joint process for capital and the and its associated filtering error $f_{kt} = k_t - \widehat{k}_t$ under incomplete information is given by:[21]

$$\begin{bmatrix} k_{t+1} \\ f_{kt+1} \end{bmatrix} = \begin{bmatrix} \lambda - \mu\eta & \mu\eta \\ 0 & \lambda\left(1 - \widetilde{\beta}_k\right) \end{bmatrix} \begin{bmatrix} k_t \\ f_{kt} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 - \widetilde{\beta}_k & -\widetilde{\beta}_k \end{bmatrix} \begin{bmatrix} u_{t+1} \\ w_{t+1} \end{bmatrix} \quad (27)$$

The stability of this process is governed by two eigenvalues: the first, $\lambda - \mu\eta$ is the single stable eigenvalue for capital under full information, the second, $\lambda\left(1 - \beta_k\right)$, determines the stability of the filtering error. It is entirely unaffected by the behaviour of consumption, and is thus identical to the filtering error of the exogenous process in the parallel problem. Transparently, for $\lambda > 1$ stability of the filtering error always requires a non-zero value of β_k. Thus, however poor the information set, it is always optimal to respond to it. We show in Appendix

[20] The parallel problem is (trivially) both stabilisable and detectable in this simple case: both conditions collapse to the trivial condition that $\lambda - \alpha < 1$ for some constant α

[21] The vector of filtering errors f_t has a second element f_{wt}, but this can be derived straightforwardly using the adding up constraint in Corollary 6, as $f_{wt} = -f_{kt}$.

F that

$$\lim_{R\to\infty}\beta_k = 0 \text{ if } \lambda \le 1$$
$$= \frac{\lambda^2-1}{\lambda^2+h\eta\left(\lambda^2-1\right)} > 0 \text{ if } \lambda > 1 \tag{28}$$

As R, the variance of the measurement error becomes large, the signal to noise ratio $\frac{S}{R}$ tends to zero. In this limiting case, as $R \to \infty$, if capital is non-explosive in the parallel problem ($\lambda \le 1$) it is optimal for the forecast of capital not to respond to the observation: $\beta_k \to 0$. However if the parallel process for capital is explosive, not responding to the signal leads to a much worse estimate for the capital stock since if the forecast did not respond, the filtering error f_{kt} would inherit the explosive nature of the parallel capital process and hence its variance would be infinite.

Of course the *true* capital stock in this example will not be explosive, but only because, in line with Corollary 7, $\beta_k = 0$ can never be a steady state of the Kalman Filter. If we set $\beta_k = 0 \Rightarrow \widetilde{\beta}_k = 0$ in (27) this would imply $k_{t+1} = (1-\lambda L)^{-1} u_{t+1}$ and so capital *would* transparently be explosive (indeed it would follow the identical process to that assumed in the parallel problem).[22] Thus the filtering process is crucial to the stability of capital.

While the joint process in (27) shows that filtering error "contaminates" the capital stock process, via the off-diagonal term in the autoregressive matrix, it is evident that, in line with Corollary 5, this has no permanent effects. After substituting for lagged filtering error, if we let $k^*_{t+1} = (1-(\lambda-\mu\eta)L)\,u_{t+1}$ be the value of capital under full information, then

$$k_{t+1} - k^*_{t+1} = \kappa(L)\left[\left(1-\widetilde{\beta}_k\right)u_t - \widetilde{\beta}_k w_t\right] \tag{29}$$

where

$$\kappa(L) = \frac{\mu\eta}{(1-(\lambda-\mu\eta)L)\left(1-\lambda\left(1-\widetilde{\beta}_k\right)L\right)}$$

thus incomplete information will cause the capital stock to differ from its full information value by a stationary AR(2) process. Note that while there are indeed no permanent effects, the impact of imperfect information is highly persistent. The first AR root is, as noted above, simply the single stable eigenvalue under perfect information, which is typically very close to unity. The second AR root is that of the filtering error. This will tend to zero as the variance of measurement error tends to zero, but we show in Appendix F that as measurement error variance tends to infinity it will tend to a limiting value of λ^{-1}, again, very close to unity. Thus in the limiting case of extremely poor information the capital stock will differ from its full information value by a process which will be close to having two unit roots (i.e., will be close to being $I(2)$).

[22] To see this note that, with $\beta_k = 0$, the estimated capital stock would be equal to zero in all periods, hence $k_{t+1} = f_{kt+1} = (1-\lambda L)^{-1} u_{t+1}$, after setting $\beta_k = 0$ in the bottom row of (27).

4 Discussion

4.1 Real vs pseudo time

As is standard in all applications of the Kalman Filter, the recursion in equations (14) to (16) does not depend on the data. In principle the whole recursion can therefore be carried out in pseudo time and then the model can be solved using converged values of β and P. This is standard practice in most, if not all, of the recent papers cited in the introduction. It should however be borne in mind that this methodology makes the implicit assumption that there is a very long history of observations of i_t in the information set.

An alternative approach that does not make such strong assumptions is to solve the model and do the recursion together, period by period. In this approach forecasts would be more uncertain initially (i.e., $P_0 > P$) because of the need to draw from the unconditional distribution initially, but would become less uncertain as time proceeded. This reduction in uncertainty would not depend on the data but *would* depend on the passage of time. Even in the special case that the information set asymptotically reveals full information, the endogeneity of states would imply that the lagged impact of initial uncertainty may persist for distinctly longer.

4.2 Adding-up constraints

We have already noted, in Corollary 6, that the filtering errors are linearly dependent. This reflects the efficient use of the structural knowledge of the economy that underpins the Kalman Filter. To see the intuition for this result, note that, if we take t-dated expectations of the measurement equation (7), using (8) this implies

$$\begin{aligned} i_t &= H_\xi \xi_t + H_c \eta' \widehat{\xi}_t = \left(H_\xi + H_c \eta'\right) \widehat{\xi}_t \\ &\Rightarrow H'_\xi \xi_t = H'_\xi \widehat{\xi}_t \Rightarrow H_\xi f_t = 0 \end{aligned}$$

thus agents know that filtering errors for any given state variable must be precisely offset by some combination of other estimation errors. By implication neither the innovation matrix of the filtering error, f_t, nor its autoregressive matrix can be of full rank, and thus the n-dimensional vector f_t can always be expressed in terms of a sub-vector of dimension $r - n$.[23]

4.3 Observable innovations, impulse responses and time series properties

The joint process for ξ_t and f_t in (22) shows that incomplete information introduces more complicated dynamics than under full information. However care needs to be applied in interpreting impulse responses from models with incom-

[23] In Appendix F we illustrate the nature of these restrictions for the analytical example discussed in Section 3.

plete information. Impulse responses derived from the full reduced form representation in (22) would not be observable in real time.

The only observable impulse responses in real time would be those implied by the simpler representation of the estimated states in (17), which are hit by the observable innovations $\beta_t \varepsilon_t$. Using (17), the estimated states have the same vector autoregressive form as the true states under full information. However, innovations to the estimated states may differ substantially, so that there is no guarantee that real-time observable impulse responses will resemble those under full information. From (6), the covariance matrix of innovations to full information states is $E(v_t v_t') = Q$ whereas we show in Appendix D that the steady state covariance matrix of innovations to the estimated states is given by

$$E\left[\left(\widehat{\xi}_{t+1} - E_t\widehat{\xi}_{t+1}\right)\left(\widehat{\xi}_{t+1} - E_t\widehat{\xi}_{t+1}\right)'\right] = Q + F_\xi M F_\xi' - M \tag{30}$$

where, from (12) and (21) M is the covariance matrix of f_t, the vector of filtering errors.

In the example of Section 3, in which there is only a single persistent true state variable, the implied difference in innovation variance is relatively modest, and in any case impulse responses of the estimated state to observable innovations will simply be a scaling of the true impulse response under full information.[24] This in itself is somewhat paradoxical, first because the observable innovations in that example will of course not be simply technological shocks, but will include a component due to measurement error, that will become dominant as the variance of measurement error increases, and second, because the true (but unobservable) impulse responses to a true technological shock will (from (27)) differ significantly from those under full information.

But once we move to a system with multiple states, the impact of differences in the nature of the innovation matrix to estimated states also becomes potentially more significant, since the structure of the matrix can in principle change radically. Most notably, while by definition pre-determined variables have no innovations under full information, this restriction does not apply to *estimates* of pre-determined variables. Thus impulse responses of estimated states may trace out responses to shocks that would simply not occur under full information.

As an extreme example of this Graham and Wright (2007) show that in a two state version of the stochastic growth model, there exists a set of parameter values, very close to those commonly used in the RBC literature, such that, whereas the true innovations may have a lower triangular covariance matrix of the form $\begin{bmatrix} 0 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}$ (all true innovations are technology shocks) the observable innova-

[24] The innovation variance of the estimated capital stock in that example will be given by

$$E\left(\widehat{k}_{t+1} - E_t\widehat{k}_{t+1}\right)^2 = S + \left(\lambda^2 - 1\right)\left(1 - \widetilde{\beta}_k\right) P_k$$

which will be strictly greater thann S, the innovation variance of the true capital stock under full information. But with increased variance of measurement error there are offsetting effects: P_k rises but β_k falls, thus dampening the impact of measurement error.

tions can have the form $\begin{bmatrix} \sigma_2^2 & 0 \\ 0 & 0 \end{bmatrix}$ (all observable innovations are attributed to mis-measurement of capital).

Since dynamic choice variables under incomplete information are, from (11) the same linear weighting of estimated states as of true states under full information, and the vector autoregressive representation of the former only differs from that of the latter by its innovation covariance matrix, this in turn determines the impact of incomplete information on the time series properties of dynamic choice variables. In both our analytical example and the example from Graham & Wright cited above, incomplete information can result in a significant increase in the variance of the single dynamic choice variable, consumption. This increase in variance need not, however, necessarily arise. For the general case the last two terms in (30) do not necessarily sum to a positive definite matrix, hence incomplete information can in principle result in a process for dynamic choice variables with higher, or lower, variance than under full information.[25]

4.4 Data Vintages

Our specification allows for the possibility of persistence in the structural measurement error process w_t, as given in (5). One possible example of how such persistence might arise is if statistical offices measure capital by perpetual inventory techniques that may result in cumulation of measurement errors in investment flows. An alternative, more structural source of persistence in Graham & Wright (2007) arises when agents use their own observable wage (which may be hit by persistent idiosyncratic shocks) as a proxy for the aggregate wage.

Any such process is however quite distinct from the common assumption that data quality may improve over time, such that lagged values of state variables may be measured with lower (or possibly no) error, compared to current values (as assumed by, for example, Collard & Dellas, 2006). A small amendment to the analytical example of section 3 provides an illustration of the impact of different data vintages on the filtering problem.

Assume that we introduce a second measured variable such that $i_t = \begin{bmatrix} i_{1t} & i_{2t} \end{bmatrix}'$ where $i_{1t} = k_t + w_t$ as in (25), and

$$i_{2t} = \lambda k_{t-1} \tag{31}$$

thus to simplify matters we assume that the capital stock in the previous period can be measured without error. Substituting from the capital evolution equation (24) this implies

$$i_{2t} = k_t - u_t$$

so that the measurement error in the new measured variable is perfectly negatively correlated with the underlying innovation to capital. In the appendix we show that in this case the signal extraction problem becomes static with $\widetilde{\beta}_{k1} = S/(S+R)$ (which is strictly less than the equivalent expression when only data on i_{1t} are available) and the filtering error for capital, f_{kt}, is white noise.

[25] A point also made by Pearlman et al (1986) and Pearlman (1992).

As a result, the impact of filtering error in "contaminating" state dynamics is distinctly weakened.[26] Thus if data quality is assumed to improve rapidly with new vintages of data, we would expect relatively more modest deviations from full information.

However, this begs the obvious question of how any such improvement in quality may arise. One answer may be that statistical offices effectively engage in an informal version of "backward-smoothing", whereby the Kalman filter can be used to derive improved estimates of underlying states by working backwards in time (see for example, Harvey, 1989, Section 3.6) and thus exploiting the benefits of hindsight. But, to the extent that this is the explanation, it clearly should have no impact on forward-looking behaviour at all since (at best) it implies that later vintages of data simply exploit the same information set that is used in deriving the current best estimates of the state variables and thus cannot improve the accuracy of these estimates.[27]

5 Conclusions

In this paper we have derived a general method of solving the signal extraction problem in linearised dynamic stochastic general equilibrium economies, which allows for potential endogeneity between dynamic choice variables and both measured and state variables. and derive a number of key results (summarised in our introduction) relating to the nature of the "endogenous Kalman Filter".

Most existing papers in the literature simply assume that the representative agent faces informational restrictions. However in related work (Graham and Wright, 2007) we cast some doubt on this approach by showing that the complete markets which underlie the assumption of a representative agent imply full information. Instead we go on to microfound incomplete information in a model of idiosyncratic shocks and incomplete markets. One implication of this is that models of incomplete markets usually also involve incomplete information, and thus need to be solved use the techniques outlined in this paper.

While we have derived the general analytical framework of the endogenous Kalman Filter from a standard linearised dynamic stochastic general equilibrium model of the type analysed by, e.g., McCallum (1998). and have emphasised the application of our techniques to this type of model, most of our results are quite general, and in principle applicable in a wide variety of contexts where dynamic optimisation problems involve both imperfectly observable states and endogeneity. As such the techniques set out in this paper broaden out further the already wide scope for application of the Kalman Filter.[28]

[26] We showed in Section 3 that with only the single measured variable, k_t differed from its full information value k_t^* by an AR(2). In this case it differs by an AR(1) with strictly lower innovation variance.

[27] Some form of "backward-smoothing" by statistical offices may also mean that econometric impulse responses estimated using historic data may be closer to true impulse responses (ie, including the impact of the filtering process), but, as discussed in Section 4.3 this does not mean that such impulse responses would be observable in real time.

[28] Consider, as a simple example, a dynamic traffic congestion problem, in which it may be possible to give some information on exogenous or stochastic factors affecting traffic flow. The

Appendix

A Derivation of Equations (6) and (7)

As a first stage in the derivation we stack equations (1) to 3) to derive the law of motion for the state variables, defined, as in the main text, by $\xi_t = \begin{bmatrix} k_t & z_t & w_t \end{bmatrix}'$, and respecify (4) accordingly, giving

$$\xi_{t+1} = D_{\xi\xi}\xi_t + D_{\xi y}y_t + v_{t+1} \tag{32}$$

$$i_t = D_{i\xi}\xi_t + D_{iy}y_t \tag{33}$$

where

$$\begin{aligned} D_{\xi\xi} &= \begin{bmatrix} B_{kk} & B_{kz} & 0 \\ 0 & B_{zz} & 0 \\ 0 & 0 & B_{ww} \end{bmatrix}; \; D_{\xi y} = \begin{bmatrix} B_{ky} \\ 0 \\ 0 \end{bmatrix}; \\ D_{i\xi} &= \begin{bmatrix} C_{ik} & C_{iz} & C_{iw} \end{bmatrix}; \quad D_{iy} = C_{iy}; \\ Q &= E\left(v_t v_t'\right) = E\left(v_t v_t'\right) = \begin{bmatrix} 0_{r_k \times r_k} & 0_{r_k \times s} \\ 0_{s \times r_k} & S \end{bmatrix}; \; S = \begin{bmatrix} S_{\zeta\zeta} & S_{\zeta\omega} \\ S_{\zeta\omega}' & S_{\omega\omega} \end{bmatrix} \end{aligned}$$

We next partition (1) into the following form[29]

$$\begin{bmatrix} A_{cc} & A_{cx} \\ 0 & 0 \end{bmatrix} E_t \begin{bmatrix} c_{t+1} \\ x_{t+1} \end{bmatrix} = \begin{bmatrix} B_{cc} & B_{cx} \\ B_{xc} & B_{xx} \end{bmatrix} \begin{bmatrix} c_t \\ x_t \end{bmatrix} + \begin{bmatrix} B_{c\xi} \\ B_{x\xi} \end{bmatrix} \xi_t \tag{34}$$

where the first block of equations are expectational difference equations and thus represent dynamic choice variables such as consumption or policy variables The second block of equations represent purely static relationships (for example, intratemporal optimality conditions, production functions, identities, etc.) Using these, assuming B_{xx}^{-1} exists[30] we can substitute out straightforwardly, using

$$x_t = -B_{xx}^{-1}\left[B_{xc}c_t + B_{x\xi}\xi_t\right] = D_{xc}c_t + D_{x\xi}\xi_t \tag{35}$$

and write the state and measurement equation in their final form in the main text as

$$\xi_{t+1} = F_\xi \xi_t + F_c c_t + v_{t+1} \tag{6}$$

$$i_t = H_\xi' \xi_t + H_c c_t \tag{7}$$

optimising response to this information will affect the subsequent flow of traffic, and may also possibly affect the measurement process itself.

[29] This form for A_{yy} will usually follow naturally from the structure of the model, but as long as A_{yy} is singular this structure can always be achieved by an appropriate linear re-weighting of the elements of y_t. The sub-matrix A_{cx} may also in principle contain columns of zeros.

[30] The case where B_{xx}^{-1} does not exist implies that some elements of x_t can be expressed as linear combinations of other elements, and can thus be trivially dealt with by substitution.

where $D_{\xi y} = \left[\begin{array}{cc} D_{\xi c} & D_{\xi x} \end{array} \right]$, $D_{iy} = \left[\begin{array}{cc} D_{ic} & D_{ix} \end{array} \right]$, etc., and

$$\begin{aligned} F_{\xi} &= D_{\xi\xi} + D_{\xi x} D_{x\xi}; \quad F_c = D_{\xi c} + D_{\xi x} D_{xc}; \\ H'_{\xi} &= D_{i\xi} + D_{ix} D_{x\xi} \quad H_c = D_{ic} + D_{ix} D_{xc} \end{aligned}$$

Note that the substitutions involved in deriving (6) and (7) are by no means innocuous in informational terms.

First, even static relationships may require informational assumptions. Since they may involve linear combinations of state variables it may be of considerable importance whether these combinations, or the elements of x_t themselves, are in the information set I_t. The form of the measurement equation allows for the possibility that elements of x_t may be observable, whether directly or indirectly, but there may be interesting cases where they are not.

The nature of the expectational difference equations satisfied by c_t, the dynamic choice variables, may also have important informational implications. While this framework can in principle accommodate any structure to the top block of equations in (34), certain structures may require assumptions about the nature of the information set. Thus if we substitute out for the static relations using (35) and from state process (6) and use the conjectured form for optimal choices of c_t as in (11) we can write the top block, applying the law of iterated expectations, as

$$\left\{ A_{cc}\eta' + A_{cx}\left(D_{xc}\eta' + D_{x\xi} \right) \right\} G\widehat{\xi}_t = \left\{ B_{cc} + B_{cx} D_{xc} \right\} \eta' \widehat{\xi}_t + \left\{ B_{c\xi} + B_{cx} D_{x\xi} \right\} \xi_t$$

which depends on ξ_t as well as $\widehat{\xi}_t$. For such a formulation to be informationally feasible in this precise form, the linear combination of states given by $\left\{ B_{c\xi} + B_{cx} D_{x\xi} \right\} \xi_t$ must be observable, and therefore should also be an element of i_t. In principle this may significantly alter the information set and hence the nature of the filtering problem (although clearly the rationale for this combination being observable should of course be justifiable). However, it will not alter the certainty-equivalent nature of the consumption function. If this linear combination is indeed observable, then (from Corollary 6) efficiency of the state estimates requires that they satisfy the adding up constraint $\left\{ B_{c\xi} + B_{cx} D_{x\xi} \right\} \xi_t = \left\{ B_{c\xi} + B_{cx} D_{x\xi} \right\} \widehat{\xi}_t$ thus allowing the top block to be written entirely in terms of state estimates, as

$$\left\{ A_{cc}\eta' + A_{cx}\left(D_{xc}\eta' + D_{x\xi} \right) \right\} G\widehat{\xi}_t = \left[\left\{ B_{cc} + B_{cx} D_{xc} \right\} \eta' + B_{c\xi} + B_{cx} D_{x\xi} \right] \widehat{\xi}_t$$

which results in an undetermined coefficients problem identical to that under full information. Note also that the nature of the undetermined coefficients problem is unchanged if this linear combination of states is not observable, but is replaced by the same combination of state estimates.

This issue does not, of course, arise if, as in many contexts (for example consumption Euler equations) B_{cx} and $B_{c\xi}$ are zero.

B Proof of Proposition 1

We assume that in some period $t-1$ initial estimates of the states ξ_t and P_t are available, that must satisfy $E_{t-1}\widehat{\xi}_t = E_{t-1}\xi_t$ by the law of iterated expectations, given the definition of $\widehat{\xi}_t$. This condition will always be satisfied if, at $t = 0$, $E_0\widehat{\xi}_1 = E_0\xi_1$

B.1 Forecasting i_t

We have (7) which we reproduce here,

$$i_t = H_\xi'\xi_t + H_c\eta'\widehat{\xi}_t$$

hence

$$\begin{aligned} E_{t-1}i_t &= H_\xi' E_{t-1}\xi_t + H_c\eta' E_{t-1}\widehat{\xi}_t \\ &= \left(H_\xi' + H_c\eta'\right) E_{t-1}\xi_t \end{aligned} \tag{36}$$

where the second line follows by the law of iterated expectations. The error of this forecast is

$$\begin{aligned} i_t - E_{t-1}i_t &= H_\xi'\xi_t + H_c\eta'\widehat{\xi}_t - \left(H_\xi' + H_c\eta'\right) E_{t-1}\xi_t \\ &= H_\xi'\left[\xi_t - E_{t-1}\widehat{\xi}_t\right] + H_c\eta'\left(\widehat{\xi}_t - E_{t-1}\xi_t\right) \end{aligned}$$

We then treat (17), the process for the estimated states, as a conjectured solution to the filtering process, which we show to be verified by the actual solution. Conditional upon this conjectured solution

$$\begin{aligned} i_t - E_{t-1}i_t &= H_\xi'\left[\xi_t - E_{t-1}\widehat{\xi}_t\right] + H_c\eta'\beta_t\left(i_t - E_{t-1}i_t\right) \\ &= J_t'\left[\xi_t - E_{t-1}\xi_t\right] \end{aligned} \tag{37}$$

where

$$J_t' = [I_n - H_c\eta'\beta_t]^{-1}H_\xi' \tag{38}$$

Thus we have, using (37) and (13)

$$\begin{aligned} E\left[\left(i_t - E_{t-1}i_t\right)\left(i_t - E_{t-1}i_t\right)'\right] &= J_t'E\left[\left(\xi_t - E_{t-1}\xi_t\right)\left(\xi_t - E_{t-1}\xi_t\right)'\right]J_t \\ &= J_t'P_tJ_t \end{aligned} \tag{39}$$

B.2 Deriving the updating equation (17).

Since (conditional upon β_t and hence J_t) innovations to i_t depend only on unobservable errors in forecasting the states, the Kalman Gain matrix, β_t in the updating equation (17) is

$$\beta_t = \left\{E\left[\left(\xi_t - E_{t-1}\xi_t\right)\left(i_t - E_{t-1}i_t\right)'\right]\right\}\left\{E\left[\left(i_t - E_{t-1}i_t\right)\left(i_t - E_{t-1}i_t\right)'\right]\right\}^{-1} \tag{40}$$

and, using (37) and (13)

$$\begin{aligned} E\left[\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)(i_t - E_{t-1}i_t)'\right] &= E\left[\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)\left(J_t'\left[\xi_t - E_{t-1}\widehat{\xi}_t\right]\right)'\right] \\ &= E\left[\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)'\right] J_t \\ &= P_t J_t \end{aligned}$$

hence

$$\beta_t = P_t J_t \left[J_t' P_t J_t\right]^{-1} \tag{41}$$

and the covariance matrix of the filtering errors can be written as

$$\begin{aligned} M_t &= E\left[\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)'\right] \\ &\quad -\beta_t E\left[(i_t - E_{t-1}i_t)\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)'\right] \\ &= P_t - \beta_t J_t' P_t = \left[I_{r+n} - \beta_t J_t'\right] P_t \end{aligned} \tag{42}$$

however these do not yet constitute closed form solutions since, via (38), J_t depends on β_t.

B.3 The recursion for P_{t+1} is F_c-independent

Conditional upon updated estimates of the states in period t, the forecast error in predicting the states in period $t+1$ is, using (10), and (6),

$$\begin{aligned} \xi_{t+1} - E_t\widehat{\xi}_{t+1} &= F_\xi \xi_t + F_c \eta' \widehat{\xi}_t + v_{t+1} - \left(F_\xi + F_c\eta'\right)\widehat{\xi}_t \\ &= F_\xi\left(\xi_t - \widehat{\xi}_t\right) + v_{t+1} \end{aligned} \tag{43}$$

and is thus independent of F_c. Hence, using the orthogonality assumptions and (12),

$$P_{t+1} = F_\xi M_t F_\xi' + Q. \tag{44}$$

B.4 P_{t+1} and M_t are H_c-independent

[NB K_t replaced with L_t throughout to avoid confusion with Svensson & Woodford's K in Section B.6] If we define

$$J_t^{'} = L_t^{-1} H_\xi^{'}, \tag{45}$$

where L_t is the (as yet unknown) matrix that satisfies

$$L_t = \left(I_n - H_c\eta'\beta_t\right) \in \mathbb{R}^{n\times n}. \tag{46}$$

then

$$\left(J' P_t J_t\right)^{-1} = \left(L_t^{-1} H_\xi' P_t H_\xi \left(L_t^{-1}\right)'\right)^{-1} = L_t' \left(H_\xi' P_t H_\xi\right)^{-1} L_t \tag{47}$$

and hence

$$\begin{aligned}\beta_t &= P_t H_\xi L L_t' \left(H_\xi' P_t H_\xi\right)^{-1} L_t \\ &= P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} L_t \end{aligned} \tag{48}$$

and thus

$$\begin{aligned}\beta_t J_t' = J_t \left(J_t' P_t J_t\right)^{-1} J_t' &= H_\xi \left(L_t^{-1}\right)' L_t' \left(H_\xi' P_t H_\xi\right)^{-1} L_t L_t^{-1} H_\xi' \\ &= H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} H_\xi'. \end{aligned} \tag{49}$$

We thus have

$$M_t = \left(I_r - P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} H_\xi'\right) P_t, \tag{50}$$

which implies that

$$P_{t+1} = F_\xi \left(I_r - P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} H_\xi'\right) P_t F_\xi' + Q, \tag{51}$$

Both of these expressions imply that the recursions for M_t and P_t do not depend on L_t and hence are are H_c-independent, and can thus be derived by setting $H_c = 0$ as in the parallel problem. If we define $\widetilde{\beta}_t$ as in (16) then the above formulae are identical to (14) and (15) in Proposition 1. We also have, using (49)

$$\beta_t J_t' = \widetilde{\beta}_t H_\xi' \tag{52}$$

B.5 Derivation of β_t

Finally we need to obtain an expression for J_t itself, and hence for β_t.

Equations (38) and (41) imply the seemingly nonlinear equation

$$J_t' = \left(I_n - H_c \eta' P_t J_t \left(J_t P_t J_t^{-1}\right)\right)^{-1} H_\xi'.$$

However, using (45) and (48), we obtain

$$L_t^{-1} H_\xi^{'} = J_t' = \left(I_n - H_c \eta' P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} L_t\right)^{-1} H_\xi',$$

which implies

$$\left(I_n - H_c \eta' P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} L_t\right) L_t^{-1} H_\xi' = H_\xi'$$

that is,

$$L_t^{-1} H_\xi' - H_c \eta' P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} H_\xi' = H_\xi'.$$

Recalling (45) once again, we find

$$J_t' = H_\xi' + H_c\eta' P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1} H_\xi'$$
$$= \left(I_n + H_c\eta' P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1}\right) H_\xi', \tag{53}$$

and thus

$$L_t = \left[I_n + H_c\eta' P_t H_\xi \left(H_\xi' P_t H_\xi\right)^{-1}\right]^{-1} \tag{54}$$

Using (16), these can be expressed as

$$J_t' = \left(I_n + H_c\eta'\widetilde{\beta}_t\right) H_\xi' \tag{55}$$

$$L_t = \left(I_n + H_c\eta'\widetilde{\beta}_t\right) H_\xi' \tag{56}$$

which, after substituting from (56) into (48) gives (19) in Proposition 1.■

B.6 Comparison with Svensson & Woodford (2003) and Pearlman et al (1986)

Svensson & Woodford's (2003) have a structural model which in reduced form is extremely close to ours. Their equations (15) and (16) correspond directly to our state and measurement equations (6) and (7), after substituting from (11). Using their notation their equation (22) is

$$X_{t|t} = X_{t|t-1} + K\left[L\left(X_t - X_{t|t-1}\right) + v_t\right]$$

where, $X_{t|t}$ in their notation corresponds to $\widehat{\xi}_t$ in ours, and Z_t to our i_t. They then assert that this allows them to identify K as "(one form of) the Kalman Gain Matrix" (which they assume, without proof, will converge to a fixed matrix). However, by the usual convention in the literature the Kalman gain updates in response to a forecast error. The square bracketed expression is not a true forecast error. Using their (16), the true forecast error in their framework is

$$Z_t - E_{t-1}Z_t = L\left(X_t - X_{t|t-1}\right) + M\left(X_{t|t} - X_{t|t-1}\right) + v_t$$

where the endogeneity of the measured variables to the response of the estimated states is evident.

However, it turns out that, despite the somewhat unusual basis for their derivation, their final result is in fact identical to our own. If we re-express their (22) in our own notation (apart from the matrix K), it becomes

$$\widehat{\xi}_t - E_{t-1}\widehat{\xi}_t = K\left[H_\xi'\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)\right]$$

whereas we show that, in our notation, from (37), after substituting from the endogenous response of c_t, and assuming convergence, the updating rule in response to the forecast error in the measured variables is given by

$$\widehat{\xi}_t - E_{t-1}\widehat{\xi}_t = \beta\left(i_t - E_{t-1}i_t\right) = \beta J'\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)$$

But using their equations (24) and (25) (noting that we absorb the covariance matrix of measurement errors into Q, and hence P), their derivation implies, in our notation,

$$K = \widetilde{\beta}$$

thus in our notation K is identical to the Kalman gain matrix in the parallel, rather than the actual problem. But, from (52), we have $\beta J' = \widetilde{\beta}H'_{\xi}$, hence

$$\widehat{\xi}_t - E_{t-1}\widehat{\xi}_t = \beta J'\left(\xi_t - E_{t-1}\widehat{\xi}_t\right) = \widetilde{\beta}H'_{\xi}\left(\xi_t - E_{t-1}\widehat{\xi}_t\right)$$

thus Svensson & Woodford's updating rule is in fact identical to our own. An equivalent updating rule is also given in Pearlman et al (1986) equation (39) but without noting the equivalence of M_t and P_{t+1} in the parallel problem or deriving convergence conditions.

C Proof of Proposition 2 and Corollary 3.

C.1 Proof of Proposition 2

Since β_t and M_t can both be expressed in terms of P_t and structural parameters a necessary and sufficient condition for convergence of all three matrices to a unique steady state is convergence of P_t to a unique steady state. Since P_t can be derived from the the parallel problem of Proposition 1 in which the states are exogenous we only need be concerned with the stability properties of that problem. Anderson and Moore (1979, Section 4.4) provide a proof of a unique stable steady state given controllability and detectability as defined in the main text for any invertible P_0, although their derivation is unfortunately diffused throughout Anderson and Moore (1979, Section 4.4). The key point is that equations (14), (15) and (16) reduce to

$$P_{t+1} = F_{\xi}\left(I_r - P_tH_{\xi}(H'_{\xi}P_tH_{\xi})^{-1}H'_{\xi}\right)P_tF'_{\xi} + Q,$$

which is a special case of Anderson and Moore (1979, (4.1)). The proof of convergence for arbitrary P_0 then follows from the subsection *Convergence for arbitrary* Σ_{k_0/k_0-1}, Anderson and Moore (1979, p. 81).

C.2 Proof of Corollary 3

We first restate (14), writing $F \equiv F_{\xi}$, $H \equiv H_{\xi}$ in this section, for brevity, as

$$P_{t+1} = F\left(I_r - P_tH\left(H'P_tH\right)^{-1}H'\right)P_tF' + Q. \tag{57}$$

In other words, we are iterating the function $g : \mathbb{P}_r \to \mathbb{P}_r$, where $\mathbb{P}_r$ denotes the set of all non-negative definite symmetric, real $r \times r$ matrices, and

$$g(P_t) = F\left(I_r - P_t H\left(H' P_t H\right)^{-1} H'\right) P_t F' + Q, \qquad P_t \in \mathbb{P}_r. \tag{58}$$

If the conditions set by Proposition 2 are satisfied, then this iteration is stable around a unique fixed point P

We first note a convenient simplification. Let

$$\widehat{F}(P_t) = F\left(I_r - P_t H\left(H' P_t H\right)^{-1} H'\right) = F\left(I_r - \widetilde{\beta}(P_t) H'\right) \tag{59}$$

(where the second expression uses (16)) then:

Lemma 10 *The function $g : \mathbb{P}_r \to \mathbb{P}_r$ defined by (58) can be expressed, using (59), in the symmetric form*

$$g(P_t) = \widehat{F}(P_t) P \widehat{F}(P_t)' + Q, \tag{60}$$

Proof. Using (59), we have

$$g(P_t) = \widehat{F}(P_t) P_t F' + Q$$

and

$$\widehat{F} P \widehat{F}' = \widehat{F} P F' - F\left(I_r - \widetilde{\beta} H'\right) P H \widetilde{\beta}' F'$$

but

$$\begin{aligned} \left(I_r - \widetilde{\beta} H'\right) P H \widetilde{\beta}' &= P H \widetilde{\beta}' - \widetilde{\beta} H' P H \widetilde{\beta}' \\ &= P H \widetilde{\beta}' - P H\left(H' P H\right)^{-1} H' P H \widetilde{\beta}' = 0 \end{aligned}$$

■

As is usual in the analysis of fixed point iteration, we must calculate the (Fréchet) derivative of g at the fixed point P.

Lemma 11 *If $E \in \mathbb{P}_r$, then, letting $\widehat{F}_P = \widehat{F}(P)$*

$$g(P + E) = g(P) + \widehat{F}_P E \widehat{F}_P' + O(E^2). \tag{61}$$

Thus if we let Dg_P denote the Fréchet derivative of the matrix function g at the point $P \in \mathbb{P}_r$, then

$$Dg_P(E) = \widehat{F}_P E \widehat{F}_P' \qquad E \in \mathbb{P}_r. \tag{62}$$

Proof. We have, using (58)

$$
\begin{aligned}
& g(P+E) \\
& = F(P+E)F' - F(P+E)H(H'PH + H'EH)^{-1}H'(P+E)F' + Q \\
& = F(P+E)F' - F(P+E)H\left[(H'PH)(I + (H'PH)^{-1}(H'EH))\right]^{-1}H'(P+E)F' + Q \\
& = F(P+E)F' - F(P+E)H\left[I - (H'PH)^{-1}(H'EH)\right](H'PH)^{-1}H'(P+E)F' + Q + O(E^2) \\
& = g(P) + \widehat{F}_P E \widehat{F}_P' + O(E^2). \qquad (63)
\end{aligned}
$$

■

It is useful to restate (61) and (62) in Kronecker product notation, as

$$\text{vec}(g(P+E)) = \text{vec}(g(P)) + \widehat{F}_P \otimes \widehat{F}_P \text{vec}(E) + O(\text{vec}(E^2))$$

hence, in this form the Fréchet derivative is (using (59))

$$Dg_P = \widehat{F}_P \otimes \widehat{F}_P = F\left(I_r - \widetilde{\beta}(P)H'\right) \otimes F\left(I_r - \widetilde{\beta}(P)H'\right) \tag{64}$$

and thus as a corollary of Proposition 2, stability of the steady state implies that the matrix denoted $F_\xi\left(I_r - \widetilde{\beta}(P)H_\xi'\right)$ in the main text must have eigenvalues with real parts strictly less than one in absolute value. Since products of matrices have common non-zero eigenvalues irrespective of order of multiplication this condition must also apply to the matrix $\left(I_r - \widetilde{\beta}(P)H_\xi'\right)F_\xi$.■

D Derivation of joint process for ξ_t and f_t under incomplete information and proofs of corollaries

D.1 Derivation of (22).

Using (6), (11) and (21) we have

$$
\begin{aligned}
\xi_{t+1} &= F_\xi \xi_t + F_c \eta' \widehat{\xi}_t + v_{t+1} \\
&= \left(F_\xi + F_c\eta'\right)\xi_t - F_c\eta' f_t + v_{t+1} \\
&= G\xi_t - F_c\eta' f_t + v_{t+1}
\end{aligned}
$$

For the estimated states we have, using (17), the definition of G in (9), (37) and (6)

$$
\begin{aligned}
\widehat{\xi}_{t+1} &= G\widehat{\xi}_t + \beta\varepsilon_{t+1} \\
&= G\widehat{\xi}_t + \beta J'\left[\xi_{t+1} - G\widehat{\xi}_t\right] \\
&= G\widehat{\xi}_t + \beta J'\left[\xi_{t+1} - \left(F_\xi + F_c\eta'\right)\widehat{\xi}_t\right] \\
&= G\widehat{\xi}_t + \beta J'\left[\xi_{t+1} - F_\xi\xi_t - F_c\eta'\widehat{\xi}_t + F_\xi\left(\xi_t - \widehat{\xi}_t\right)\right] \\
&= G\widehat{\xi}_t + \beta J'\left[v_{t+1} + F_\xi f_t\right]
\end{aligned}
$$

hence

$$\begin{aligned} f_{t+1} &= \left(G - F_c\eta' - \beta J' F_\xi\right) f_t + \left(I - \beta J'\right) v_{t+1} \\ &= \left[I - \beta J'\right] F_\xi f_t + \left(I - \beta J'\right) v_{t+1} \end{aligned}$$

which, using (52) we can also express as

$$f_{t+1} = \left[I - \widetilde{\beta} H_\xi'\right] F_\xi f_t + \left(I - \widetilde{\beta} H_\xi'\right) v_{t+1} \tag{65}$$

Thus we have the joint process for ξ_{t+1} and f_{t+1} as in (22) which we reproduce here:

$$\begin{bmatrix} \xi_{t+1} \\ f_{t+1} \end{bmatrix} = \begin{bmatrix} G & -F_c\eta' \\ 0 & \left[I - \widetilde{\beta} H_\xi'\right] F_\xi \end{bmatrix} \begin{bmatrix} \xi_t \\ f_t \end{bmatrix} + \begin{bmatrix} I \\ I - \widetilde{\beta} H_\xi' \end{bmatrix} v_{t+1}$$

D.2 Proofs of Corollaries

Since the filtering error process f_t is block recursive we can write the top block of equations as

$$\begin{aligned} \xi_{t+1} &= [I - GL]^{-1} v_{t+1} - [I - GL]^{-1} F_c \eta' f_t \\ &= \xi_{t+1}^* - [I - GL]^{-1} F_c \eta' f_t \\ &= \xi_{t+1}^* - [I - GL]^{-1} F_c \eta' \left[I - \left(I - \beta J'\right) F_\xi L\right]^{-1} \left(I - \beta J'\right) v_t \end{aligned}$$

where ξ_{t+1}^*, the full information state process, is as given by (9). The incomplete information states are thus equal to the full information states plus a lag polynomial in the filtering error (itself a lag polynomial in the underlying shocks). Since the filtering error is stationary (from Corollary 3) and the full information process is non-explosive (from Assumptions 1 and 2) the incomplete information process is also non-explosive, thus proving Corollary 4.

Since there may be permanent productivity or other shocks, G may have unit eigenvalues, implying permanent effects of these shocks. But permanent effects will only arise with respect to rows of v_t for which the relevant rows of F_c are, by Assumption 1, zero (the shock processes are exogenous). Hence filtering error will only cause transitory deviations from the full information outcome, proving Corollary 5.

Using (16) we have

$$H_\xi' \widetilde{\beta} = H_\xi' P H_\xi \left(H_\xi' P H_\xi\right)^{-1} = I_n \tag{66}$$

If we pre-multiply (65) by H_ξ' and use (66) we have

$$H_\xi' f_{t+1} = H_\xi' \left[I - \widetilde{\beta} H_\xi'\right] F_\xi f_t + H_\xi' \left(I - \widetilde{\beta} H_\xi'\right) v_{t+1} = 0$$

thus proving Corollary 6.

By inspection of (22), if F_ξ has explosive eigenvalues in its sub-matrix F_{kk} and $\beta_k = 0$, then the matrix $\left(I - \widetilde{\beta} H'\right) F_\xi$ will also have explosive eigenvalues, which, from Corollary 3, contradicts stability of the recursion for P_{t+1}, thus proving Corollary 7.■

D.3 Derivation of (30)

From the autoregressive representation of the estimated states in (17) we have, using (18), (37), (39) and (52)

$$
\begin{aligned}
E\left[\left(\widehat{\xi}_{t+1}-E_t\widehat{\xi}_{t+1}\right)\left(\widehat{\xi}_{t+1}-E_t\widehat{\xi}_{t+1}\right)'\right] &= \beta E\left(\varepsilon_{t+1}\varepsilon_{t+1}'\right)\beta' \\
&= \beta J'PJ\beta' = \widetilde{\beta}H_\xi' PH_\xi\widetilde{\beta}'
\end{aligned}
$$

but hence, using (16), (15), and exploiting symmetry of P and M

$$
\begin{aligned}
E\left[\left(\widehat{\xi}_{t+1}-E_t\widehat{\xi}_{t+1}\right)\left(\widehat{\xi}_{t+1}-E_t\widehat{\xi}_{t+1}\right)'\right] &= PH_\xi\left[H_\xi' PH_\xi\right]^{-1}PH_\xi\widetilde{\beta}' \\
&= PH_\xi\widetilde{\beta}' = \left(\beta H'P\right)' \\
&= (P-M)' = P-M
\end{aligned}
$$

which, after substituting from the steady state of (14) gives (30).

E Proof of Proposition 9

Using (40) our original definition of β_t, and (18), we have

$$
\beta_t = E\left(\left(\xi_{t+1}-E_t\xi_{t+1}\right)\varepsilon_t'\right)\left(E\left(\varepsilon_{t+1}\varepsilon_{t+1}'\right)\right)^{-1}
$$

Conjecture that the proposition is correct, and thus a limiting steady state exists with $P = Q \Rightarrow M = 0$. Since both conditions are identical in the parallel problem of Proposition 1 we can set $F_c = H_c = 0$. Since this steady state replicates full information, we have

$$
\begin{aligned}
E\left(\left(\xi_{t+1}-E_t\xi_{t+1}\right)\varepsilon_t\right) &= E\left(F_u u_{t+1}\left(H'F_u u_{+1}\right)'\right) = F_u S F_u' H \\
E\left(\varepsilon_{t+1}\varepsilon_{t+1}'\right) &= E\left(H'F_u u_{t+1}\left(H'F_u u_{t+1}\right)'\right) = H'F_u S F_u' H
\end{aligned}
$$

and hence, using (20), we have

$$
\begin{aligned}
\beta(Q) &= F_u S F_u' H\left(H'F_u S F_u' H\right)^{-1} \\
&= F_u S F_u' H\left(F_u' H\right)^{-1}S^{-1}\left(H'F_u\right)^{-1} \\
&= F_u\left(H'F_u\right)^{-1}
\end{aligned}
$$

and hence

$$\begin{aligned}\beta(Q)H'Q &= F_u\left(H'F_u\right)^{-1}H'F_uSF_u' = F_uSF_u' = Q\\ &\Rightarrow M(Q) = \left(I_r - \beta(Q)H'\right)Q = 0\\ &\Rightarrow P = Q\end{aligned}$$

which verifies the conjecture that $|H'F_u| \neq 0$ (which transparently requires $n = s$) is a necessary condition for $P = Q$ to be a steady state.

However using our proof of Corollary 3 stability of a steady state requires the third condition in Proposition 9, that the matrix $F_\xi\left(I_r - \widetilde{\beta}(Q)H_\xi'\right)$ have eigenvalues with real parts strictly less than zero in absolute value. Since

$$\text{rank}\left(\widetilde{\beta}(Q)H'\right) = \text{rank}\left(F_u\left(H'F_u\right)^{-1}H'\right) = n$$

(since each of its elements are rank n),

$$\text{rank}\left(I_r - \beta(Q)H'\right) = \text{rank}\left(I_r - F_u\left(H'F_u\right)^{-1}H'\right) = r - n$$

and hence

$$\text{rank}\left(F_\xi\left(I_r - \beta(Q)H'\right)\right) \leq r - n$$

hence this third condition is in general non-trivial.■

F Derivation of analytical example

F.1 Basic model derivation

To motivate the example, assume the following underlying processes for output, Z_t and capital, K_t

$$\begin{aligned}Z_t &= K_t^\alpha\\ K_{t+1} &= \left[(1-\delta)K_t + Z_t - C_t\right]\exp(v_{t+1})\end{aligned}$$

thus new technology is embodied in new capital. If we assume for simplicity that v_t is white noise this means there is only a single state variable. Assume further that i_t is a noisy measure of log output (which we rescale for convenience into units of capital)

$$i_t = \alpha^{-1}z_t + w_t$$

After log linearisation this means we can write the model as in the main text, reproduced here, as

$$\begin{aligned}k_{t+1} &= \lambda k_t + \mu c_t + u_{t+1}\\ i_t &= k_t + w_t\end{aligned}$$

where $k_t = \log K_t$, and c_t is log consumption. Under standard assumptions (e.g. as in Campbell, 1994) we have $\lambda > 1$, and, assuming that we can represent aggregate consumption by that of a representative consumer, the consumption function (26) will have $1 > \eta_k > 0$.[31]

As in (6) we incorporate measurement errors into the states, and write

$$\xi_{t+1} = \begin{bmatrix} k_{t+1} \\ w_{t+1} \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} k_t \\ w_t \end{bmatrix} + \begin{bmatrix} -\mu \\ 0 \end{bmatrix} c_t + \begin{bmatrix} u_{t+1} \\ w_{t+1} \end{bmatrix}$$

(where clearly there is only 1 true state process, k_t, since F_ξ is rank 1). The measurement equation (25) can be written as in (7) as

$$i_t = \begin{bmatrix} 1 & 1 \end{bmatrix} \xi_t + hc_t$$

thus setting $H'_\xi = \begin{bmatrix} 1 & 1 \end{bmatrix}$.

Following the logic of Proposition 1 the parallel problem sets $h = \mu = 0$. The recursion for P_t is then of the form

$$\begin{aligned} P_{t+1} &= F_\xi M_t F'_\xi + Q \\ &= \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} M_{11t} & M_{12t} \\ M_{21t} & M_{22t} \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} S & 0 \\ 0 & R \end{bmatrix} \\ &= \begin{bmatrix} \lambda^2 M_{11t} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} S & 0 \\ 0 & R \end{bmatrix} \end{aligned}$$

from which it can be seen directly that P_t takes the form

$$P_t = \begin{bmatrix} P_{kt} & 0 \\ 0 & R \end{bmatrix}$$

Now derive $\widetilde{\beta}_t$ in the parallel problem using this form for P_t

$$\begin{aligned} \widetilde{\beta}_t &= P_t H_\xi \left[H'_\xi P_t H'_\xi \right]^{-1} \\ H'_\xi P_t H_\xi &= \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} P_{kt} & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = P_{kt} + R \end{aligned}$$

$$P_t H_\xi = \begin{bmatrix} P_{kt} & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} P_{kt} \\ R \end{bmatrix}$$

$$\Rightarrow \widetilde{\beta}_t \equiv \begin{bmatrix} \widetilde{\beta}_{kt} \\ \widetilde{\beta}_{wt} \end{bmatrix} = \frac{1}{P_{kt} + R} \begin{bmatrix} P_{kt} \\ R \end{bmatrix}$$

hence $\widetilde{\beta}_{wt} = 1 - \widetilde{\beta}_{kt}$, and hence

$$M_t = \left(I - \widetilde{\beta} H'_\xi \right) P_t = \begin{bmatrix} 1 - \beta_{kt} & -\beta_{kt} \\ -(1 - \beta_{kt}) & \beta_{kt} \end{bmatrix} \begin{bmatrix} P_{kt} & 0 \\ 0 & R \end{bmatrix}$$

[31] In Campbell's (1994) framework, the linearisation parameter λ corresponds to his λ_1 and $\mu = 1 - \lambda_1 - \lambda_2$ (we set $\lambda_2 = 0$). The consumption function parameter η_k corresponds to Campbell's η_{kk} despite the fact that capital is not predetermined, since u_{t+1} drops out of the Euler equation in expectation.

so the recursion for P_{kt} is simply

$$\begin{aligned} P_{kt+1} &= \lambda^2 P_{kt}\left(1-\widetilde{\beta}_{kt}\right)+S \\ &= \lambda^2 P_{kt}\left(\frac{R}{P_{kt}+R}\right)+S \end{aligned}$$

If we solve for the steady state, P_k, this implies the quadratic

$$P_k^2+\left[R(1-\lambda^2)-S\right]P_k-SR=0$$

which we can write as

$$\begin{aligned} \left(\frac{P_k}{R}\right)^2+\left[1-\lambda^2-\frac{S}{R}\right]\frac{P_k}{R}-\frac{S}{R} &= 0 \\ p_k^2+\left(1-\lambda^2-s\right)p_k-s &= 0 \end{aligned}$$

giving

$$p_k=\frac{s+\lambda^2-1+\sqrt{\left(1-\lambda^2-s\right)^2+4s}}{2}>s$$

where $p_k=\frac{P_k}{R}$ and $s=\frac{S}{R}$ (the signal to noise ratio). If we let $R\to\infty$ this implies $s\to 0$, and

$$\begin{aligned} \lim_{s\to 0} p_k &= \frac{-\left(1-\lambda^2\right)+\sqrt{\left(1-\lambda^2\right)^2}}{2} \\ &= 0, \lambda\leq 1 \\ &= \lambda^2-1, \dot{\lambda}>1 \end{aligned}$$

This implies the limiting steady state value of $\widetilde{\beta}_k$

$$\begin{aligned} \lim_{s\to 0}\widetilde{\beta}_k &= 0;\lambda\leq 1 \\ &= \frac{\lambda^2-1}{\lambda^2};\lambda>1 \end{aligned} \tag{67}$$

and hence a limiting steady state value of the autoregresssive coefficient of the filtering error for capital of

$$\lim_{s\to 0}\lambda\left(1-\widetilde{\beta}_k\right)=\frac{1}{\lambda}$$

For the general case, using (19) β_{kt} is given by

$$\beta_{kt}=\frac{\widetilde{\beta}_{kt}}{1+h\eta\widetilde{\beta}_{kt}}$$

Using (67), this implies the limiting steady state value as R goes to infinity, and hence s goes to zero,

$$\lim_{s\to 0}\beta_k = \frac{\lambda^2 - 1}{\lambda^2 + h\eta\left(\lambda^2 - 1\right)}$$

F.2 Adding-up constraints

The adding up constraint (66) discussed in Section 4.2 requires, for the parallel problem,

$$H'_\xi \widetilde{\beta} = 1 \Rightarrow \beta_w = 1 - \beta_k$$

which we have already shown is satisfied.

We can also write

$$I - \widetilde{\beta}_t H'_\xi = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 - \beta_{kt} & -\beta_{kt} \end{bmatrix}$$

implying that both $M_t = \left(I - \widetilde{\beta} H'_\xi\right) P_t$ and the steady-state autoregressive matrix of the filtering error $\left(I - \widetilde{\beta} H'_\xi\right) F_\xi$ are of rank 1.

F.3 Extension of example to allow for different data vintages

Given the additional measured variable the measurement equation becomes

$$i_t = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \xi_t$$

and in the state equation

$$F = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad Q = \begin{bmatrix} S & S & 0 \\ S & S & 0 \\ 0 & 0 & R \end{bmatrix}$$

from which it is straightforward to see that P_t is of the form

$$P_t = \begin{bmatrix} P_{1t} & S & 0 \\ S & S & 0 \\ 0 & 0 & R \end{bmatrix}$$

Thus we can derive $\widetilde{\beta}_t$ using this form for P_t

$$\left[H'_\xi P_t H_\xi\right]^{-1} = \frac{1}{(P_{1t}-S)(R+S)}\begin{bmatrix} -S+P_t & S-P_t \\ S-P_t & R+P_t \end{bmatrix}; \quad P_t H_\xi = \begin{bmatrix} P_t & -S+P_t \\ S & 0 \\ R & 0 \end{bmatrix}$$

$$\Rightarrow \widetilde{\beta}_t = P_t H_\xi \left[H'_\xi P_t H_\xi\right]^{-1} = \frac{1}{(P_{1t}-S)(R+S)}\begin{bmatrix} S(P_t-S) & R(P_t-S) \\ S(P_t-S) & -S(P_t-S) \\ R(P_t-S) & -R(P_t-S) \end{bmatrix}$$

$$= \frac{1}{(R+S)}\begin{bmatrix} S & R \\ S & -S \\ R & -R \end{bmatrix}$$

so the filtering problem becomes a purely static one: that of identifying two white noise disturbances.

Thus we have

$$\widetilde{\beta}_1 = \frac{S}{S+R}$$

which is strictly less than the equivalent figure assuming no data for i_{2t} since in that case $P_1 > S$ Using this form we can calculate the autoregressive matrix of the filtering error

$$\left(I - \widetilde{\beta} H'_\xi\right) F_\xi = \begin{bmatrix} 0 & \frac{R}{R+S} & \frac{-S}{R+S} \\ 0 & 1-\frac{S}{R+S} & \frac{-S}{R+S} \\ 0 & \frac{-R}{R+S} & 1-\frac{R}{R+S} \end{bmatrix}\begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

thus the filtering error is also vector white noise.

References

Anderson, Brian and John Moore, (1979) "Optimal Filtering", Prentice Hall

Aoki, Kosuke, (2006), "Optimal commitment policy under noisy information", Journal of Economic Dynamics and Control 30, pp.81-109.

Aoki, Kosuke, (2003), "On the optimal monetary policy response to noisy indicators", Journal of Monetary Economics 50 (3), pp.501-523.

Blanchard, Olivier and James A Kahn, (1980), "The Solution of Linear Difference Models under Rational Expectations", Econometrica 48 (5), pp.1305-1312.

Bomfim, Antulio N, (2001), "Measurement error in general equilibrium: the aggregate effects of noisy economic indicators", Journal of Monetary Economics 48, pp.505-603.

Brainard, William. (1967). "Uncertainty and the Effectiveness of Policy." American Economic Review 57, pp. 411-425.

Campbell, John Y, (1994), "Inspecting the mechanism: an analytical approach to the stochastic growth model", Journal of Monetary Economics 33, pp.463-506.

Collard, Fabrice and Harris Dellas, (2006), "Imperfect Information and Inflation Dynamics", Working Paper.

Graham, Liam and Stephen Wright, (2007), "Information, heterogeneity and market incompleteness in the stochastic growth model", Working Paper.

Harvey, Andrew (1981) "Time Series Models" Philip Allan, London

Harvey, Andrew (1989) "Forecasting, Structural Time Series Models and the Kalman Filter" Cambridge University Press

Hamilton, James, (1994), "Time Series Analysis", Princeton University Press

Keen, Benjamin D, (2004), "The Signal Extraction Problem Revisited: Its Impact on a Model of Monetary Policy", Working Paper.

Kydland, Finn E and Edward C Prescott, (1982), "Time to build and aggregate fluctuations", Econometrica 50 (6), pp.1345-1370.

McCallum, Bennett T, (1998), "Solutions to linear rational expectations models: a compact exposition", Economics Letters 61, pp.143-147.

Mehra, Rajnish and Edward C Prescott, (1980), "Recursive Competitive Equilibrium: The Case of Homogeneous Households", Econometrica 48 (6), pp.1365-1379.

Pearlman, J., D Currie and P Levine, (1986) "Rational Expectations Models with Private Information". Economic Modelling 3 (2), 90–105.

Pearlman, Joseph, (1992), "Reputational and nonreputational policies under partial information", Journal of Economic Dynamics and Control 16 (2), pp.339-357.

Svennson, Lars EO and Michael Woodford, (2004), "Indicator variables for optimal policy under asymmetric information.", Journal of Economic Dynamics and Control 28 (4), pp.661-690.

Svennson, Lars EO and Michael Woodford, (2003), "Indicator variables and optimal policy", Journal of Monetary Economics 50 (3), pp.691-720.