Invertible and Non-Invertible Information Sets in Dynamic Stochastic General Equilibrium

Brad Baxter,† Liam Graham and Stephen Wright‡

February 4, 2008

Abstract

Most DSGE modelling derives rational expectations solutions on the assumption that all state variables relevant to optimising decisions are directly observable. This can be rationalised, as in Mehra & Prescott (1990) by the assumption that the states are "...an invertible function of observables." We show, by application of an endogenous version of the Kalman Filter, that the conditions for a more general concept of "asymptotic invertibility" of an information set can be less stringent in the context of DSGE models. We also show that non-invertibility of the information set can have significant implications for the time-series properties of economies. We provide a Matlab toolkit which allows the easy implementation of a wide class of models conditional upon a given information set.

JEL classifications: D80; E32.

Keywords: imperfect information; invertibility; fundamental versus nonfundamental time series representations; Kalman filter; dynamic stochastic general equilibrium

---

*School of Economics, Maths & Statistics Birkbeck College, University of London, Malet Street, London W1E 7HX, UK. b.baxter@bbk.ac.uk
†Department of Economics, University College London, Gower Street, London WC1E 6BT, UK. Liam.Graham@ucl.ac.uk
‡Corresponding author: School of Economics, Maths & Statistics Birkbeck College, University of London, Malet Street, London W1E 7HX, UK. s.wright@bbk.ac.uk
1 Introduction

Underlying most dynamic stochastic general equilibrium models is the strong assumption that all state variables relevant to optimising decisions are directly observable. This can be rationalised, as in Mehra & Prescott (1990) by the assumption that the states are "...an invertible function of observables." In this paper we consider rational expectations solutions of DSGE models that condition explicitly on the information set available to optimising agents. We state conditions under which an information set is invertible and we analyse the implications of non-invertible information sets in DSGE models.

We distinguish between instantaneous and asymptotic invertibility of an information set. The former only requires \( t \)-dated information, the latter exploits the full history of observables. We show that both can be nested within a general DSGE framework in which optimising agents solve a signal extraction problem, given available information, using a version of the Kalman Filter modified to take account of the endogeneity of both states and observables to agents' optimising decisions. While some past research (Pearlman et al, 1986; Pearlman, 1992; Svensson & Woodford, 2003, 2004) has addressed this problem, we 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. This insight allows us to derive minimal conditions for asymptotic invertibility, and also to explore the implications of information sets that are not invertible and hence do not replicate full information.

Our key results are:

1. We show that asymptotic invertibility of an information set requires (as a necessary but not sufficient condition) that the number of observables be equal to the "stochastic dimension" of the system (i.e., the number of structural innovations). Since DSGE models usually have pre-determined endogenous variables, their stochastic dimension is less than their state dimension, so the condition for asymptotic invertibility can be less stringent than that for instantaneous invertibility which requires equality with the state dimension.

2. We show that necessary and sufficient conditions for asymptotic invertibility also put important additional restrictions on the nature of the measurement process, and its interaction with the dynamics of the underlying states.

3. Asymptotic invertibility of an information set is closely related to, but not identical to invertibility (or "fundamentalness") in time series analysis. The solution to a Kalman Filter problem always results in a fundamental time series representation of the observables, for which
the innovations can be recovered from the history of the observables (see, e.g., Hamilton, 1994, p 391). But asymptotic invertibility makes the stronger requirement that structural innovations that drive the true states must themselves be "fundamental" in the time series sense.

4. When the information set is non-invertible we show that the impact of limited information is transitory, but can be highly persistent, and introduces new (but, in real time at least, intrinsically unobservable) sources of dynamics in response to structural shocks.

5. The only observable dynamics will be those of the estimated, as opposed to the actual states. We show that the estimated states follow the same vector autoregressive process as would the true states in a notional full information economy but, if the information set is non-invertible, with a different covariance pattern of shocks. This can introduce significant differences in time series properties. In particular, estimates of pre-determined variables like capital may be subject to "shocks", that are logically impossible under full information.

6. In standard Kalman Filter problems, when information on the economy is very noisy, in the limit it becomes optimal to ignore it. In the endogenous Kalman Filter, in contrast, we show that it may be optimal to update estimates of some states, however poor the quality of the information.

To complement the paper we provide a Matlab toolkit which allows the easy application of our techniques to a wide class of linear models. Limited information sets 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 (2004), Keen (2004), Collard & Dellas (2006) where representative consumers are assumed to face informational restrictions. Woodford (2003) and Nimark, (2007a) address the problem of heterogenous agents facing a symmetric filtering problem of inferring aggregates which requires them to form estimates of a "hierarchy of average expectations" (Townsend, 1983). Graham and Wright (2007) show that when such models include endogenous states a filtering problem of the form discussed in the present paper is central to the solution. We discuss this particular case further in Section 6.2.

The remainder of the paper is organised as follows. Section 2 derives the endogenous Kalman filter for a general DSGE model. In Section 3 we give conditions for asymptotic invertibility, and relate these to the nature of time series representations. In Section 4 we discuss some implications of non-invertible information sets. In Section 5 we show how our techniques can be applied to an analytical example based on the benchmark stochastic
growth model. Section 6 discusses applications and extensions 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 as in McCallum (1998) as:

\[
\begin{align*}
A_{yy}y_{t+1} &= B_{yy}y_t + B_{yk}k_t + B_{yz}z_t \\
k_{t+1} &= B_{ky}y_t + B_{kk}k_t + B_{kz}z_t \\
z_{t+1} &= B_{zz}z_t + \zeta_{t+1}
\end{align*}
\]

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, \( k_t \), while the third describes the evolution of an \( r_z \times 1 \) vector of exogenous stochastic processes, \( z_t \), that can be represented by a first order vector autoregression, with \( \zeta_t \) an \( r_z \times 1 \) vector of iid innovations with covariance matrix \( S_{zz} = E(\zeta_t'\zeta_t) \) which we assume is full rank.\(^1\) Using results such as those in Hamilton (1994, pp 678-9) a process as in (3) can also represent a wide range of stochastic processes such as Markov Chains by an appropriate specification of the process for the innovations.

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

\[
i_t = C_{ik}k_t + C_{iz}z_t + C_{iy}y_t + C_{iw}w_t
\]

where \( i_t \) is an \( n \times 1 \) vector of observed variables, \( \Xi \) 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 \( 0 \leq r_w \leq n \).\(^2\)

For generality we can in principle allow these to be serially correlated by representing them as a vector autoregression of the form

\[
w_{t+1} = B_{ww}w_t + \omega_{t+1}
\]

\(^1\)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 \( S_{zz} \) is full rank.

\(^2\)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.
where $\omega_t$ has the full rank covariance matrix $S_{\omega\omega} = E[\omega_t\omega_t']$. We assume that the eigenvalues of $B_{ww}$ have real parts less than or equal to unity. The two innovations $\omega_t$ and $\zeta_t$ may in principle be contemporaneously correlated, with $E(\zeta_t w_t') = S_{\zeta\omega}$, but are assumed uncorrelated at all other leads and lags.

We noted in the introduction that different information sets can be given different rationales, which we do not attempt to explore further here, since they are typically model-specific. However, we shall show that different information sets can have very different implications for the properties of the economy. Hence the nature of the information set should be stated explicitly from the outset, and preferably in a way that can be related to the underlying structure of the economy. It should also be borne in mind that, as we show in Appendix A, the structure of the model in (1) to (3) needs to be both informationally consistent and informationally feasible (most obviously for forward-looking variables).

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.

2.2 The 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:

\begin{align*}
\xi_{t+1} &= F_\xi \xi_t + F_c c_t + v_{t+1} \\
\zeta_t &= H_\xi' \xi_t + H_c c_t
\end{align*}

where

\begin{align*}
\xi_t &= \begin{bmatrix} k_t \\ z_t \\ w_t \end{bmatrix};
\zeta_t &= \begin{bmatrix} 0 \\ \zeta_t \\ \omega_t \end{bmatrix};
\end{align*}

\begin{align*}
Q &= E(v_t v_t') = \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{align*}

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.

We assume that optimising agents solve the signal extraction problem using the Kalman Filter. However, the system in (6) and (7) differs from
that underlying standard derivations of the Kalman filter (as in, e.g., Harvey, 1981; 1989; Hamilton, 1994) where the system is usually written as

\[ \xi_{t+1} = F\xi_t + v_{t+1} \quad (6a) \]
\[ i_t = H\xi_t + w_t \quad (7a) \]

where \(\xi_t\) 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, there are two forms of endogeneity in our system since 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.

It might appear that a state equation of the form in (6) can be transformed into the standard form, (6a), by letting the optimal values of the dynamic choice variables, \(c_t\), depend on the states, and thus derive a system of the same form as (6a) and (7a) by substitution. In section 2.3, we shall show that this is indeed the way in which the rational expectations solution is derived under the standard assumption of full information. But in this case, of course, the informational problem entirely disappears, and hence the Kalman Filter is redundant. For the general case, in which the states are not directly observed (or, equivalently, the information set is not instantaneously invertible), the Kalman Filter is required, because such a straightforward substitution is impossible.

Second, in standard applications of the Kalman Filter, where the \(\xi_t\) are 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 \(F\xi\) 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, \(F\xi\) 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 latent 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, \(\xi_t\). 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 other information sets.

**Definition 1** (Full information) Full information implies the state variables are known, \( i_t = \xi_t \).

Full information is a special case of the system (6) and (7) with \( n = r \), \( H_{\xi} = I_r \), \( H_c = 0 \). The Kalman Filter is again redundant.

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

\[
\begin{align*}
    c_t^* &= \eta' \xi_t^* \\
    \text{where all elements in the } i\text{th row of } \eta \text{ are zero for } i > r_k + r_z \text{ (measurement errors have no impact on under full information), and for any variable } x_t, \quad x_t^* \text{ denotes its value under full information. The matrix } \eta, \text{ which in general depends on all structural and preference parameters of the model, can be computed using standard techniques (following, e.g. Blanchard and Kahn, 1980; McCallum, 1998). For the rest of the paper we treat it as parametric.}
\end{align*}
\]

Given (8), the full information states follow a first order vector autoregressive process in reduced form:

\[
\xi_{t+1}^* = G \xi_t^* + \nu_{t+1}
\]

where

\[
G = F_{\xi} + F_c \eta'
\]

The behaviour of the dynamic choice variables, \( c_t \), is crucial for the stability of the states, under any information set. As noted in the previous section,
in a model with endogenous capital \( F_\xi \) 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 stabilisation follows directly from the standard rational expectations solution. Under standard conditions the matrices \( \eta \) 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 \( \Lambda \) is a diagonal matrix of eigenvalues and \( V \) the corresponding matrix of eigenvectors. For any strictly unit eigenvalue in \( \Lambda \) 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).\(^3\)

These features of the solution under full information turn out to be equally crucial for the stability of the solution even when the information set is non-invertible, and thus does not replicate full information.

### 2.4 Instantaneous Invertibility

The full information solution can also be replicated straightforwardly under the following conditions:

**Definition 2** *(Instantaneous invertibility).* An information set in instantaneously invertible if

1. \( n = r \), the number of observables is equal to the state dimension
2. \( H_\xi \) in the measurement equation (7) is invertible

and hence full information can be replicated using only \( t \)-dated information.

In this case the state variables can be replaced in the state equation by setting \( \xi_t = \left(H_t^{-1}\right)^{(i_t - H_c c_t)} \) and can therefore be treated as known, hence the Kalman Filter is redundant. The counterpart to this is that all lagged information on the observables \( \{i_{t-j}\} \) for \( j > 0 \) is also redundant.

\(^3\)Assumption 2 follows naturally from the underlying structural model, since the dependence of state variables on \( c_t \) is only via \( k_{t+1} \), hence all elements of the \( i \)th row of \( F_c \) are zero for \( i > r - s \).
2.5 The Endogenous Kalman Filter

For all other information sets, including, as we shall show, some that are invertible, we need to apply the Kalman Filter allowing for the endogeneity of the dynamic choice variables to the filtering process.

Following Pearlman (1992) and Svensson and Woodford (2004) we conjecture that for the general case optimal choices will be “certainty-equivalent”, which in this context is defined as:

\[ c_t = \eta \tilde{\xi}_t \]  

(11)

where \( \tilde{\xi}_t = E_{t}\xi_t|I_t \) is the optimal estimate of the current state vector\(^4\) given the available information set \( I_t \), which evolves as in (7) and \( \eta \) 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.

\[ M_t = E \left[ \left( \xi_t - \tilde{\xi}_t \right) \left( \xi_t - \tilde{\xi}_t \right)^\prime \right] \]  

(12)

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

\[ P_{t+1} = E \left[ \left( \xi_{t+1} - E_t\xi_{t+1} \right) \left( \xi_{t+1} - E_t\xi_{t+1} \right)^\prime \right] \]  

(13)

is the covariance matrix of the one-step ahead state forecast errors.\(^5\)

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 a parallel exogenous Kalman Filter problem

\[ \tilde{\xi}_{t+1} = F\tilde{\xi}_t + v_t \]
\[ \tilde{\eta}_t = H\tilde{\xi}_t \]

(i.e., setting \( F_c = H_c = 0 \) in (6) and (7)). They are thus given by the

\(^4\)For compactness of notation we write the period \( t \) estimate of the states as \( \tilde{\xi}_t \); whereas the standard Kalman filter literature commonly uses \( \tilde{\xi}_{i[t]} \). For the forecast at time \( t \) of the states at period \( t + 1 \) we write \( E_t\xi_{t+1} \) instead of the standard \( \tilde{\xi}_{t+1[i]} \).

\(^5\)\( P_{t+1} \) is commonly denoted \( P_{t+1[i]} \), and using the same notation \( M_t = P_{t[i]} \), but we separate the two for clarity.
standard exogenous Kalman filter recursion

\[ P_{t+1} = F_t M_t F'_t + Q \]  
\[ M_t = \left[ I_r - \beta_t H'_t \right] P_t \]  
\[ \beta_t = P_t H' \left[ H'_t P_t H'_t \right]^{-1} \]

However, in the solution to the actual endogenous Kalman Filter problem, conditional upon certainty-equivalence of the dynamic choice variables in (11), the estimated states follow the process

\[ \hat{\xi}_{t+1} = G \hat{\xi}_t + \beta_t \varepsilon_{t+1} \]  

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

\[ \varepsilon_{t+1} \equiv i_{t+1} - E_t i_{t+1} \]  

and

\[ \beta_t = \beta_t \left[ I_n + H_c \beta_t \right]^{-1} \]

**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 \( e_\xi \) and a notional set of measured variables \( e_i \) for which both effects are absent, so that the standard Kalman filter formulae can be applied.\(^6\)

Proposition 1 shows that, conditional upon the solution for \( \beta_t \) in (19), the estimated states \( \hat{\xi}_t \) 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 \( \hat{\xi}_t \) 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.

We can derive the Kalman filter from the properties of the parallel, rather than the true process, because, 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

\(^6\)Our formulae for \( P_{t+1} \) and \( \beta_t \) 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.

9
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 matrix $\beta_{t}$ for the true problem is not the same as its counterpart $\tilde{\beta}_{t}$ 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 $P_{t+1}$, thus these can be derived under the assumption that $H_{c} = 0$.

In Appendix B.6 we show that the solution to the filtering problem given by Proposition 1 is equivalent to that in Svensson & Woodford (2003) and in Pearlman et al (1986). However neither set of authors draws the link with the “parallel problem”, and do not derive the remainder of our results.

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 $F_{\xi}$. 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 $\beta$, $P$ 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

$$v_{t} = F_{u}u_{t}$$

where $F_{u} = \begin{bmatrix} 0_{(r-s)\times s} \\ I_{s} \end{bmatrix}$; $u_{t} = \begin{bmatrix} \zeta_{t} \\ \omega_{t} \end{bmatrix}$; $E\left(u_{t}u'_{t}\right) = S$

7We 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 $K = F_{\xi}\beta_{t}$, 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 $\beta_{t}$ but also in the autoregressive representation in (17).
then the two conditions can be written as

\begin{align*}
\text{stabilisability:} & \quad |\lambda_i (F \xi + F_u L_1^i) | < 1 \quad \forall i \\
\text{detectability:} & \quad |\lambda_i (F \xi + L_2 H \xi) | < 1 \quad \forall i
\end{align*}

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 $i$th 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 $F \xi$, so that the condition for stabilisability can only be met if the relevant row of $F \xi$ 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.

Note that, as is standard in all applications of the Kalman Filter, the recursion in equations (14) to (16) does not depend on the data. It is common practice to carry out the recursion in pseudo time and then solve the model using steady-state values of $\beta$ and $P$. It should be borne in mind, however, that this methodology makes the implicit assumption that there is a very long history of observations of $i_t$ in the information set.

### 3 Asymptotic invertibility: replicating full information by the Kalman Filter

#### 3.1 Instantaneous versus asymptotic invertibility of the information set

We have derived the endogenous Kalman Filter for a general signal extraction problem where we have assumed only that the information set is not instantaneously invertible (if it were, as we saw in Section 2.4, the Kalman Filter would be redundant). For the general case the information set will be non-invertible: ie state estimates will differ from the true states. However, under certain conditions we now show that the Kalman Filter may converge to a steady state which does replicate full information.

Definition 2 showed that that instantaneous invertibility requires that
the number of measured variables, equal \( r \), the number of states. In most standard exogenous Kalman Filter problems, in which the stochastic dimension, \( s \), equals the state dimension, \( r \), this is the only way that full information can be replicated. However, the endogenous nature of the states in our modified version of the Kalman Filter implies that this condition is overly restrictive: it is a sufficient, but not a necessary condition.

The dynamic endogeneity of some of the states (the \( k_t \)) in our framework means that, because they are pre-determined, the stochastic dimension \( s \), is less than \( r \), the state dimension. Under conditions which we shall state below, there may as a result be cases in which the Kalman Filter yields state estimates that become arbitrarily close to the true states as the history of the information set increases over time. This property can be defined formally in three ways that are all logically equivalent.

**Definition 3 (asymptotic invertibility)** An information set \( I_t \) is asymptotically invertible if

\[
M \equiv \lim_{t \to \infty} M_t(I_t) = 0 \iff P \equiv \lim_{t \to \infty} P_t(I_t) = Q \iff \lim_{t \to \infty} \hat{\xi}_t(I_t) = \xi^*_t
\]

where, from Proposition 2, \( M \) and \( P \) satisfy the steady state of the recursion in (14) and (15).

The definition explicitly notes the dependence, not only on current values of the observed variables, \( i_t \), but on their entire history, \( I_t \). 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; and \( P_t \), the covariance matrix of one-period-ahead state forecast errors, tends to its irreducible minimum of \( Q \). Given the nature of the recursion in for \( M_t \) and \( P_t \) in (14) and (15) this means that the property of asymptotic invertibility depends potentially on the entire structure of the model and its interaction with the measurement process.

The first two conditions taken in isolation require only that the estimated states must converge on the true states, whereas the third condition appears somewhat stronger, since it requires that the true states in turn must converge on their full information values (thus any initial errors in estimating the states must have strictly transitory impact). However we show below that the three features are in fact logically identical.\(^8\)

This implies significant restrictions on the nature of the information set, and its relation to the underlying structural model, which we summarise in the following proposition:

\(^8\)See Corollary 5 below. Note however that the states will usually converge to their full information values more slowly than \( P_t \) and \( M_t \) because past errors will have a persistent effect on the true states (see discussion below in Section 4).
Proposition 3 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 invertible are

1. $n = s$

2. $\left| H'_\xi F_u \right| \neq 0$

3. $\Lambda_i \left( \left( I - \tilde{\beta}(Q) H'_\xi \right) F_\xi \right) < 1 \ \forall i$

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

Proof. See Appendix E. ■

To illustrate, assume that all three conditions in Proposition 3 do indeed hold, and that there is a sufficiently long history of the observables that state estimates in period $t$ have converged on their true values. To simplify, assume we have "dynamic endogeneity" of the states ($F_c \neq 0$) but that $H_c = 0$ (there is no endogeneity in the measurement equation). Manipulation of equations (6), (7), (18) and (20) then implies that in period $t + 1$, the innovations $\varepsilon_{t+1}$ in the observables relate to the underlying structural innovations, $u_{t+1}$ by

$$\lim_{t \to \infty} \varepsilon_{t+1} \equiv \lim_{t \to \infty} (\hat{\xi}_{t+1} - E_t \hat{\xi}_{t+1}) = H'_\xi (\hat{\xi}_{t+1} - E_t \hat{\xi}_{t+1}) = H'_\xi u_{t+1} = H'_\xi F_u u_{t+1}$$

(21)

So, if the first two conditions hold, such that $H'_\xi F_u$ is both square (given $n = s$) and invertible, then if the Kalman filter reveals the true states in period $t$, the innovations in the observables reveal the structural innovations, and hence it continues to reveal the states in period $t + 1$, and so on indefinitely.

However, this simply tells us that if the first two conditions in Proposition 3 are satisfied, then $P = Q$ is a steady state of the Kalman Filter. The third condition tells us whether this steady state is stable, and shows that the nature of the measurement process, and its interaction with the dynamics of the underlying states, is crucial. We shall show below that if the third condition does not hold, then in the neighbourhood of the full information solution even arbitrarily small errors in state estimates in period $t$ must ultimately multiply up and dominate the innovations to the observables, thus rendering the argument underlying (21) invalid.

9 This assumption is made purely to simplify the algebra; the same qualitative properties hold with both forms of endogeneity.
The conditions set by Proposition 3 state the minimal assumptions necessary for the assumption of full information to be at least asymptotically valid given a sufficiently long history of observables in the information set. All three conditions can in principle relate to interesting economic features of DSGE models.

1. Dynamic endogeneity of the states will usually imply that $s$, the stochastic dimension of an economic system, is lower than $r$, the dimension of the states. A notable example is the benchmark stochastic growth model (which we analyse in Section 5) which is driven by a single stochastic process for technology. In such cases, by Definition 2, instantaneous invertibility requires that $n$, the number of observed variables, equals $r$, the number of states. However, the first condition in Proposition 3 shows that (subject to the other two conditions also holding) it may be possible to replicate full information asymptotically with even a single observable variable.

2. A long-recognised inference problem, dating back at least to Muth (1961), and re-visited more recently by Bomfim (2004), arises when technology is subject to shocks with different dynamic effects: e.g. if there are both transitory and persistent shocks. Since such shocks have an identical initial impact on technology, they will also typically affect all observable variables identically on impact. In such cases, even if the first condition in Proposition 3 is satisfied, so that there are as many observable variables as shocks, the second condition will not be satisfied. Thus inference problems of this type will never disappear, however long the history of the observables.

3. Even if both the first two conditions are satisfied, the third condition can also be crucial. In Section 5 we give examples of two information sets, both of which satisfy conditions 1 and 2, but only one of which satisfies the third condition.

3.2 Asymptotic invertibility and time series representations of the observables

A corollary of Proposition 3 relates the concept of asymptotic invertibility of the information set to the role of the true structural innovations, $u_t$, in the time series representation of the observables, $i_t$, and in so doing illustrates the crucial nature of the third condition in Proposition 3.

**Corollary 1** Assume that the first two conditions in Proposition 3 are satisfied. The Endogenous Kalman Filter converges to a solution in which:

a) if the third condition is also satisfied (and hence the information set is asymptotically invertible) then the true structural innovations, $u_t$ are in-
novations to a fundamental time series representation of the observables, $i_t$, and hence of $c_t$.

b) if the third condition is not satisfied (and hence the information set is not asymptotically invertible) they are the innovations to a nonfundamental, nonbasic representation, as in Lippi & Reichlin (1994).

**Proof.** See Appendix E

The first part of the corollary is straightforward. It is a standard property (see, e.g., Hamilton, 1994, p 391) that the observable innovations, $\varepsilon_t$ to the measured variables, $i_t$, can always be used to derive a fundamental time series representation of $i_t$. This representation is unique up to an orthonormal transformation of the innovations (Lippi & Reichlin, 1994; Hannan, 1970). It then follows automatically that if the conditions for asymptotic invertibility hold, and $u_t$ can be recovered from the $\varepsilon_t$, as in (21), then the $u_t$ are also innovations to a fundamental representation. As such they can be recovered from the history of the observables, $i_t$. Thus asymptotic invertibility and fundamentalness of the $u_t$ are two sides of the same coin.

The nonfundamentalness of $u_t$ in the case where the third condition for asymptotic invertibility is not satisfied is similarly just another way of saying that, despite the fact that both sets of innovations have the same dimension, the $u_t$ cannot be recovered from the history of the observables. As a consequence, there is no static, invertible transformation of the form $\varepsilon_t = Au_t$ for some constant matrix $A$, as there is in the limiting case of asymptotic invertibility. Instead we have $\varepsilon_t = A(L)u_t$, where $A(L)$ is a matrix polynomial in the lag operator that maps a vector white noise process to another vector white noise process of the same dimension.

This last feature means that the $u_t$ are not only nonfundamental but nonbasic (as defined by Lippi & Reichlin, 1994), in the sense that the associated VARMA representation of the observables, $i_t$, is of higher order than the fundamental representation. Nonfundamentalness means that innovations cannot be recovered from the history of the observables. If the representation is "basic" impulse responses can at least be inferred from the fundamental representation; however for nonbasic representations, as here, even this is ruled out. Lippi and Reichlin show that there is an infinite set of such representations; however Corollary 1 implies that in only one of these representations can the innovations be derived as a static linear transformation of the true structural innovations. This is of some interest in the light of Lippi & Reichlin’s conclusion that nonbasic nonfundamental representations are "not likely to occur in models based on economic theory" (Lippi & Reichlin, 1994, p 315); our qualification to this conclusion is that they may well occur when optimising agents face informational problems.
4 Implications of non-invertible information sets

While the conditions for asymptotic invertibility (and hence replication of full information) are less stringent than those for instantaneous invertibility they nonetheless place significant restrictions on the nature of the information set. We now address the implications of an information set being non-invertible. We first show that even in this case the observable dynamics of the economy can be represented by those of a notional full information economy, but with a different covariance pattern of innovations; we then examine the implied true (but, in real time, intrinsically unobservable) dynamics of the economy, which are distinctly more complex.

4.1 An isomorphic representation

Using (17) and (9) whether or not the information set is invertible, we always have

\[ E_t \hat{\xi}_{t+1} = E_t \tilde{\xi}_{t+1} = G \xi_t \]

since if the information set is invertible, and thus replicates full information, trivially \( \hat{\xi}_t = \xi^*_t \). Since the law of motion for the estimated states has the same autoregressive structure as under full information, the coefficient matrix \( \eta \) in the conjectured solution (11) will be the same as under full information, i.e. as in (8). Thus optimal choices are certainty-equivalent, verifying our conjecture.\(^{10}\)

Given this property, and the convergence of the Kalman Filter to its steady state value, we have the following property as a direct implication of Propositions 1 and 2:

**Corollary 2** Assuming convergence of the Endogenous Kalman Filter, the behaviour of the estimated states \( \hat{\xi}_t \) and the dynamic choice variables, \( c_t \) is isomorphic to the behaviour of the true states, \( \xi_t \) and dynamic choice variables under full information, if \( Q \), the covariance matrix of underlying structural innovations in (6), is replaced by the matrix \( Q + F \xi M F^\xi - M \).

In this representation there may be shocks to states that are in reality predetermined.

**Proof.** See Appendix C

This property follows directly from the representation of the state estimates in (17), which shows that the \( \hat{\xi}_t \) have an autoregressive representation which, if it differs from that of the full information states, \( \xi^*_t \) in (9) does so

\(^{10}\)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.
only via its innovations. Given the efficiency of the Kalman Filter, the innovations to the estimated states, which are a linear combination of the innovations to the observable variables, are vector white noise conditional upon the $t$-dated information set. Given convergence of the Kalman Filter, these observable innovations have a time-invariant distribution. And, given the certainty equivalent nature of the optimal choices of the dynamic choice variables in (11) the estimated states are sufficient for a time series representation of the dynamic choice variables, just as the true states are under full information. But despite its identical autoregressive representation, the notional full information economy that represents the behaviour of the true economy with a non-invertible information set can have innovations with very different properties: most notably, as stated in Corollary 2, there may be "shocks" to pre-predetermined variables that are logically impossible under full information.

Given certainty equivalence, any differences in the dynamics of the states in the isomorphic representation in Corollary 2 in turn determine the impact of non-invertible information sets on the time series properties of dynamic choice variables. The additional terms in the covariance matrix of the notional states in the isomorphic representation do not always sum to a positive definite matrix. Hence non-invertibility of an information set can in principle result in a process for dynamic choice variables with higher, or lower, variance than under full information.\footnote{A point also made by Pearlman et al (1986) and Pearlman (1992), and relevant to the results of Bomfim (2004).}

### 4.2 The dynamics of the true states, and some implications of non-invertible information sets

While there is, as stated in Corollary 2, an autoregressive representation of the state estimates in terms of observable innovations, this representation does not describe the dynamics of the true states, except of course in the special case of invertibility. For the general, and hence non-invertible case, the true dynamics of the economy are richer once expressed in terms of the true structural innovations. Analysis of these dynamics provides further important insights.

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

$$f_t \equiv \xi_t - \hat{\xi}_t \tag{23}$$

We show in Appendix D that the joint process for $f_t$ and the true states $\xi_t$...
can be expressed in the vector autoregressive form

$$
\begin{bmatrix}
\xi_{t+1} \\
\epsilon_{t+1}
\end{bmatrix} =
\begin{bmatrix}
G & -F_c \eta' \\
0 & (I - \bar{\beta} H'_{\xi}) F_{\xi}
\end{bmatrix}
\begin{bmatrix}
\xi_t \\
\epsilon_t
\end{bmatrix}
+ 
\begin{bmatrix}
I \\
I - \bar{\beta} H'_{\xi}
\end{bmatrix}
u_{t+1}
$$

(24)

The top block is entirely independent of filtering parameters, and transparently reduces to the full information process (9) when filtering error disappears. In the special case of asymptotic invertibility via the Kalman Filter, as in Proposition 3, it is straightforward to show (see Appendix E) that the filtering error, \( \epsilon_t \), has zero innovation variance, and thus there is no "contamination" of the full information solution. In general, however, filtering error “contaminates” state dynamics via the off-diagonal element of the autoregressive matrix for the joint process for \( \xi_t \) and \( \epsilon_t \).

In contrast the process for the state filtering error \( \epsilon_t \) 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 (24):

**Corollary 3** In the autoregressive representation (24), the autoregressive matrix \( \left( I - \bar{\beta} H'_{\xi} \right) F_{\xi} \) of the filtering errors, \( \epsilon_t \equiv \xi_t - \bar{\xi}_t \), has at most \( r - n \) non-zero eigenvalues, all of which have real parts strictly less than unity in absolute value.

**Proof.** See Appendix C.

Thus convergence of the Kalman Filter to a unique steady state automatically implies that filtering error is stationary (and vice versa).

Corollary 3 also sheds light on the third condition for asymptotic invertibility in Proposition 3. By inspection this can be interpreted as a condition on the eigenvalues of the autoregressive matrix of filtering errors in the neighbourhood of the solution that replicates full information. If these eigenvalues are not stable then, even if full information is a steady state of the Kalman Filter, it will not be a stable steady state: even arbitrarily small filtering errors must ultimately multiply up, and dominate the innovations to the observables. Assuming the conditions in Proposition 2 hold, there will still of course be a stable steady state Kalman Filter, in the neighbourhood of which, from Corollary 3, filtering errors will be stationary, but it will not be one that replicates full information.

The joint process for \( \xi_t \) and \( \epsilon_t \) in (24) provides a complete description of the true process for the dynamic choice variables \( c_t \), since, using (11) we can write \( c_t = \eta' \bar{\xi}_t \equiv \eta' (\xi_t - \epsilon_t) \) : thus with a non-invertible information set 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, \( \eta \) has zeros in its \( i \)th row for \( i > r_k + r_z \). The more persistent is the filtering error process (the closer are the non-zero eigenvalues of \( (I - \beta H_E^t) F_\xi \) to unity), the more prolonged will be the additional dynamics introduced by the filtering problem.

Using Corollary 3, the representation in (24) implies a number of key features of non-invertible information sets that are direct corollaries of Propositions 1 and 2, given Assumptions 1 and 2.

**Corollary 4** The solution with a non-invertible information set is non-explosive.

**Corollary 5** Non-invertibility of an information set has no permanent effects, even when there are permanent structural shocks (i.e., if \( B_{zz} \) has one or more unit eigenvalues).

**Proof.** See Appendix D

These first two features of the solution with non-invertible information are unsurprising, if reassuring. 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 solution that arises from the enogenous Kalman Filter, even when the information set is non-invertible. Corollary 5 states that impulse responses under full information and non-invertible information must converge. Thus while short-run dynamics may be significantly affected by informational problems, long-run responses derived from models that assume full information may approximate quite well those from models with non-invertible information sets.

**Corollary 6** The filtering errors \( f_t \) satisfy \( H_\xi f_t = 0 \).

**Proof.** See Appendix D

Linear dependence between the elements of the vector of filtering errors arises from 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

\[
\begin{align*}
i_t & = H_\xi \xi_t + H_c \eta \hat{\xi}_t = (H_\xi + H_c \eta) \hat{\xi}_t \\
& \Rightarrow H_\xi' \xi_t = H_\xi' \hat{\xi}_t = H_\xi f_t = 0
\end{align*}
\]  

(25)

thus agents know that filtering errors for any given state variable must be precisely offset by some combination of other estimation errors. By impli-
cation 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 $r - n$.

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

**Proof.** See Appendix D

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 informationally optimal to update estimates of predetermined variables.

Mathematically, this result follows directly from the stationarity of the filtering error process: when $F_\xi$ has explosive eigenvalues associated with the evolution of the pre-determined variables, $k_t$, the autoregressive matrix of the filtering error in (24), $(I - \tilde{\beta}H_0')F_\xi$, could not have stable eigenvalues with $\tilde{\beta}_k = 0$. Since, from Corollary 3, stability of the filtering errors is directly equivalent to convergence of the Kalman Filter, this in turn implies that in this case $\beta_k \neq 0$ can never be a solution of the filtering problem. This has interesting 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 $F_\xi$ 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, $\beta_k$ and $\beta_z$ will both tend to zero, the state estimates $\hat{k}_t$ and $\hat{z}_t$ will tend to a constant, and the filtering errors for these state variables become the processes themselves. In the endogenous Kalman Filter problem the same feature applies when $F_\xi$ has stable or unit eigenvalues. In contrast, if $F_\xi$ has explosive eigenvalues, the state process in the parallel problem set out in Proposition 1 also has explosive eigenvalues, so for $\beta_k$ 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, $\beta_k$ tends to a fixed, non-zero matrix. Thus however poor the information, it is always optimal to respond to it.

---

13 In this case $\beta_w$ will tend to $I_{rw}$: all innovations to measured variables will be interpreted as due to measurement error. In the borderline case where $F_\xi$ has strictly unit eigenvalues the filtering error process for the corresponding states will tend towards a unit root process.

14 This feature is noted in Svensson & Woodford (2003, p711) where it is assumed to hold as a general result.

15 Since this feature derives from the parallel problem it clearly also applies to any exogenous Kalman filtering problem where the autoregressive representation is explosive.
4.3 A caveat on impulse responses

The joint process for $\xi_t$ and $f_t$ in (24) discussed in the previous section shows that non-invertible information introduces more complicated dynamics than under full information. However, impulse responses derived from the full reduced form representation in (24) would not be observable in real time. The only observable impulse responses would be those to the isomorphic representation of Corollary 2, in which, as already discussed, the innovation covariance properties of this notional economy may be very different from those of the true structural shocks.

5 The benchmark stochastic growth model with alternative information sets

In this section we present an analytical example that shows how our techniques can be applied to a model which has Campbell’s (1994) log-linearised stochastic growth model as its underlying structure. We then explore the conditions for alternative forms of invertibility under different information sets and ask whether alternative information sets are invertible. There are two states, evolving as

$$
k_{t+1} = \lambda_1 k_t + \lambda_2 a_t + (1 - \lambda_1 - \lambda_2) c_t
$$

$$
a_{t+1} = \phi a_t + u_{t+1}
$$

where capital $k_t$ is pre-determined, $c_t$ is consumption, and $a_t$ is an AR(1) aggregate technology process. The linearisation parameters $\lambda_1 > 1; \lambda_2 > 0$ are as defined in Campbell.

We assume that the vector of possible measured variables is given by

$$
\begin{bmatrix}
w_t \\
r_t
\end{bmatrix} =
\begin{bmatrix}
1 - \alpha & \alpha \\
-\lambda_3 & \lambda_3
\end{bmatrix}
\begin{bmatrix}
k_t \\
a_t
\end{bmatrix}
$$

where $w_t$ is the wage, $r_t$ the return on capital, $\alpha$ is the exponent on labour in a Cobb-Douglas production function, and $\lambda_3 > 0$ is Campbell’s third linearisation parameter. The first line of (27) is the marginal product relation for labour, the second uses Campbell’s linearisation of the net return on capital.

The specification both of the law of motion of capital in (26) and of the measurement process is simplified by following Campbell’s assumption in the first half of his paper that the labour supply curve is vertical.

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.
To simplify we assume there is no extraneous noise, and consider the implications of three possible information sets.

### 5.1 An instantaneously invertible information set

In the first case we assume that $i_t = \begin{bmatrix} w_t & r_t \end{bmatrix}'$; a straightforward result that arises naturally from, eg, observability of factor prices in a homogeneous economy in which all households earn the same wage and the same return.\(^{16}\)

Given this common information set on the aggregate economy, we have $n = r = 2$ and by inspection of (27) the conditions for instantaneous invertibility in Definition 2 are satisfied as long as $\alpha \neq \frac{1}{2}$.

### 5.2 An asymptotically invertible information set

We next consider an artificially censored version of the information set, in which $n = 1$, with the single observable variable $i_t = w_t$, the aggregate wage. This example may appear artificial, and is indeed used primarily to demonstrate the potential for asymptotic invertibility. But it is also of some independent interest, since in any solution that does replicate full information the pre-determined nature of capital means that innovations to the two measured variables will be perfectly correlated, hence one observable variable will always be informationally redundant.

Since capital is pre-determined we have $n = s = 1$ hence the first condition in Proposition 3 is satisfied. The second condition is also trivially satisfied since $H_\xi' F' u = \begin{bmatrix} 1 & 0 & \alpha \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix}' = \alpha > 0$. To derive the third condition, which tests whether the fixed point of the Kalman Filter is stable, we first need to derive the Kalman Gain matrix, $\hat{\beta}(Q) = \begin{bmatrix} \beta_k(Q) & \beta_a(Q) \end{bmatrix}'$ that replicates full information. Since capital is pre-determined we must have $\beta_k(Q) = 0$; it then follows straightforwardly that we must also have $\beta_a(Q) = \beta_a(1) = \frac{1}{\alpha}$.\(^{17}\) We then have

$$
\left( I - \hat{\beta}(Q) H_\xi' \right) F_\xi = \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & \alpha \\ \frac{1}{\alpha} & 0 & \phi \end{bmatrix} \right) \begin{bmatrix} \lambda_1 & \lambda_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{1 - \alpha} & \lambda_1 \frac{1}{1 - \alpha} \\ \lambda_2 - \frac{1}{\alpha} \end{bmatrix}
$$

of which the single non-zero eigenvalue is given by $\lambda_1 - \frac{1 - \alpha}{\alpha} \lambda_2$. Using Campbell’s (1994, p. 469) equation (14), which derives the linearisation constants from underlying structural parameters, we have

$$
\lambda_1 = \frac{1 - \alpha}{\alpha} \lambda_2 = \frac{1 - \delta}{1 + g} < 1
$$

\(^{16}\)Graham & Wright (2007) show that the same information set will also be common knowledge in a heterogeneous economy with complete markets.

\(^{17}\)The same result can also be derived by explicit use of the more general formula given in Appendix E, equation (67). Note that with fixed labour supply $c_t$ does not enter the measurement equation hence $\beta = \hat{\beta}$. 

22
where \( \delta \) is the depreciation rate and \( g \) the steady-state growth rate. Hence the fixed point is stable.

Thus an information set consisting only of the history of the aggregate wage is asymptotically invertible, and, by implication, the history of the return on capital is informationally redundant. As a further direct implication, noise in measuring returns on capital (possibly due, for example, to market frictions or noise traders in financial markets) has no impact on the dynamics of the stochastic growth model as long as the aggregate wage is observable.

### 5.3 A non-invertible information set

We now consider the alternative censored information set, \( i_t = r_t \). In Graham & Wright (2007) we show that in an incomplete markets version of the stochastic growth model in which households earn a wage with an idiosyncratic component,\(^{18}\) this is the limiting case of the public information set on the aggregate economy as the variance of the idiosyncratic component goes to infinity, so that each household’s individual wage is an arbitrarily poor signal of the aggregate wage. An obvious question is whether, despite this, the history of the return on capital is sufficient to replicate perfect information. We shall see that, despite apparent similarities with the previous case, it is not.

The first two conditions in 3 are again trivially satisfied, with \( n = s = 1 \); and

\[
H^T_F u = \begin{bmatrix} -\lambda_3 & \lambda_3 \\ 0 & 1 \end{bmatrix}' = \lambda_3 > 0.
\]

Thus full information is again a fixed point of the Kalman Filter given this information set.\(^{19}\) However the third condition is not satisfied: it is not a stable fixed point. To see this, we again derive the Kalman Gain matrix for this information set, that replicates full information, which by similar arguments is given by

\[
\beta_k(Q) = 0; \quad \beta_a(Q) = \tilde{\beta}_a(Q) = \frac{1}{\lambda_3}.
\]

This in turn implies

\[
(I - \tilde{\beta}(Q) H^T_F) F_\xi = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{\lambda_3} \end{bmatrix} \begin{bmatrix} -\lambda_3 & \lambda_3 \\ 0 & \phi \end{bmatrix} = \begin{bmatrix} \lambda_1 & \lambda_2 \\ \lambda_1 & \lambda_2 \end{bmatrix}
\]

of which the single non-zero eigenvalue is given by \( \lambda_1 + \lambda_2 > 1 \). Thus the fixed point is not stable, and hence this restricted information set is non-invertible.

---

\(^{18}\)This model, which applies and extends the techniques set out in this paper, is discussed in more detail in Section 6.2 below.

\(^{19}\)Note that, given the underlying assumptions of the model in Graham & Wright (2007), aggregate consumption, \( c_t \), which feeds into the law of motion for aggregate capital in (26), is not in general observable to any individual household, since, given incomplete risk-sharing idiosyncratic consumption depends on idiosyncratic as well as aggregate state variables. However if full information is replicated, aggregate consumption thereby becomes common knowledge, hence it is valid to analyse this fixed point by looking only at the aggregate economy and the common information set.
To see the intuition for this result, recall that, from the representation in (24), the matrices in (28) and (30) are both the autoregressive matrices of the filtering error $f_t$ in the neighbourhood of the steady state that replicates full information. In the information set with only the history of wages filtering errors are stationary, hence any initial error in estimating the state decays back to zero. By contrast, when the information set consists only of the history of the return on capital, filtering errors are explosive in the neighbourhood of this steady state. By implication, and, again in contrast to the previous case, the history of the aggregate wage, however poorly measured, is never informationally redundant.

While filtering errors are explosive in the neighbourhood of the full information steady state of the Kalman Filter, it can be shown quite straightforwardly that this information set does satisfy the two conditions in Proposition 2 for the existence of a unique steady state Kalman Filter. Thus, for the associated steady state Kalman Gain Matrix $\beta$, while filtering errors clearly have non-zero variance, Corollary 3 implies that they must be stationary.

6 Extensions and Applications

6.1 Data Vintages

A common assumption (see for example Collard & Dellas, 2006) is that data quality may improve over time, such that lagged values of state variables may be measured with less error than current values. There is also some evidence for this as an empirical phenomenon with published data (see, e.g., Orphanides and Van Norden, 2002).

It is straightforward to incorporate data vintages into our framework by extending the state space to include lags of the underlying states, with any associated measurement error in the relevant line of the measurement equation having lower variance than in the line relating to current states. However, this begs the question of how any such improvement in data 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.

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.

6.2 Heterogeneity

Our framework has considered a class of models where the Kalman Filter is applied to endogenous states, in which the filtering problem is either that of a single optimising agent, or a set of such agents who share the same information set. The two types of problem are essentially identical, since, if, for example, agents A and B both know the structural model and have the same information set, agent A does not need to observe Agent B’s behaviour: A can simply infer it from the common information set and structural properties (and vice versa for B).\textsuperscript{20}

There is potentially a much wider class of models in which individual agents, or types of agent, have overlapping, but not common information sets. Even in such models the techniques outlined in this paper are still a crucial part of the solution when states are endogenous.

In Graham & Wright (2007), for example, we solve a dynamic general equilibrium economy in which a large number of heterogeneous households interact. In this economy we show that each household faces a symmetric filtering problem of the same general form as (6) and (7). Each household has an idiosyncratic capital stock which is endogenous to their own consumption and labour supply choices, and faces idiosyncratic labour supply shocks; all other elements of the state vector are aggregate, and hence exogenous, and common to each agent’s filtering problem. Each household’s filtering problem is solved as a parallel problem as in Proposition 1, taking the dynamics of the aggregate economy as given.

The attraction of this approach is that the filtering problem arises from the structure of the economy, rather than from any extraneous noise or exogenously assumed informational restrictions. But a significant complication is that, if we assume that all households know that all other households are solving the same problem, each household also knows that the dynamics of the aggregate economy must be affected (via the aggregate capital stock) by the dynamic choices of consumers in aggregate. This aggregate behaviour is unobservable to any individual household; but households can infer the nature of the choices (and hence the dynamics of the aggregate economy) from their own behaviour. We show that the resulting solution requires each household to form estimates of a "hierarchy of average expectations" (Townsend, 1983, Woodford, 2003, Nimark, 2007a) of the underlying non-expectational states, resulting in an infinite dimensional state vector. The model can however be solved, to an arbitrary degree of precision, with a finite state vector, by extending new techniques developed by Nimark (2007b).

\textsuperscript{20}This is in essence the basis for the solution method used by Svensson & Woodford (2004).
In such an economy optimal behaviour for each household remains, as in this paper, certainty-equivalent with respect to the full state vector (including the hierarchy of expectations), but is not, in general, certainty-equivalent with respect to the underlying non-expectational states. Thus in contrast to the solution here, the aggregate economy cannot be given an isomorphic representation in terms of a notional full information economy with a different set of innovations, as in Corollary 2. However, if there is sufficient heterogeneity, there is a limiting case that is certainty-equivalent. Furthermore, on available evidence on the relative variance of idiosyncratic versus aggregate shocks, the calibrated economy is very close to this limiting case. Thus certainty equivalence appears to provide at least a reasonable approximation even of an economy with both structural and informational heterogeneity.

The informational problem in this extended framework also has some interesting features that can be related to results in this paper. As already noted, as heterogeneity increases, the filtering problem for the aggregate economy tends to a limiting case very close to that of the benchmark stochastic growth model with an artificially censored non-invertible information set that we analyse in Section 5.3. The steady-state Kalman Filter results in a seemingly perverse response to aggregate productivity shocks, which are interpreted on impact largely as negative "capital shocks", and which therefore cause aggregate consumption to fall, rather than rise, as under full information.

7 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. We derive a number of new results relating to the nature of the "endogenous Kalman Filter", focussing in particular on whether the information set is "invertible", i.e., whether it can replicate full information, at least asymptotically. Since we have summarised these properties in the introduction we shall not repeat them here.

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. However most of our results are quite general, and in principle applicable in a wide variety of contexts where dynamic optimisation problems involve states that are both endogenous, and not directly observable. As such the techniques set out in this paper broaden out further the already wide scope for application of the Kalman Filter.
References

Anderson, D and John B. Moore, (1979), "Optimal Filtering", Dover


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


28
Appendix

Note to referees: In the current version of the appendix we include fairly lengthy derivations as an aid to the refereeing process. We would envisage that in any published version of the paper the appendix would be significantly shorter.

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 respectively (4) accordingly, giving

\[ \xi_{t+1} = D_{\xi \xi} \xi_t + D_{\xi y} y_t + v_{t+1} \]

\[ i_t = D_{i \xi} \xi_t + D_{i y} y_t \]

where

\[ D_{\xi \xi} = \begin{bmatrix} B_{kk} & B_{kz} & 0 \\ 0 & B_{zz} & 0 \\ 0 & 0 & B_{ww} \end{bmatrix} ; \quad D_{\xi y} = \begin{bmatrix} B_{ky} \\ 0 \\ 0 \end{bmatrix} ; \quad D_{i \xi} = \begin{bmatrix} C_{ik} & C_{iz} & C_{iw} \end{bmatrix} ; \quad D_{i y} = C_{iy} ; \]

\[ Q = E(v_t v_t') = E(v_t v_t') = \begin{bmatrix} 0_{r_k \times r_k} & 0_{r_k \times s} \\ 0_{s \times r_k} & S \end{bmatrix} ; \quad S = \begin{bmatrix} S_{\zeta \zeta} & S_{\zeta \omega} \\ S_{\zeta \omega}' & S_{\omega \omega} \end{bmatrix} \]

We next partition \( y_t \) as \( y_t = \begin{bmatrix} c_t' & x_t' \end{bmatrix}' \) and express (1) conformably as

\[ A_{xx} E_t \begin{bmatrix} c_{t+1} \\ x_{t+1} \end{bmatrix} = \begin{bmatrix} B_{xx} \\ B_{xx} \end{bmatrix} \begin{bmatrix} c_t \\ x_t \end{bmatrix} + \begin{bmatrix} B_{x \xi} \\ B_{x \xi} \end{bmatrix} \xi_t \]

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 we can substitute out straightforwardly, using

\[ x_t = -B_{xx}^{-1} [B_{x c} c_t + B_{x \xi} \xi_t] = D_{x c} c_t + D_{x \xi} \xi_t \]

and write the state and measurement equation in their final form in the

21 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 reweighting of the elements of \( y_t \). The sub-matrix \( A_{cc} \) may also in principle contain columns of zeros.

22 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.
main text as

\[ \xi_{t+1} = F_{\xi} \xi_t + F_{c} c_t + v_{t+1} \quad (6) \]

\[ i_t = H'_{\xi} \xi_t + H_c c_t \quad (7) \]

where \( D_{\xi y} = \begin{bmatrix} D_{\xi c} & D_{\xi x} \end{bmatrix} \), \( D_{iy} = \begin{bmatrix} D_{ic} & D_{ix} \end{bmatrix} \), etc, and

\[ F_{\xi} = D_{\xi \xi} + D_{\xi x} D_{\xi} ; \quad F_c = D_{\xi c} + D_{\xi x} D_{xc} ; \]

\[ H'_{\xi} = D_{\xi \xi} + D_{ix} D_{\xi} \quad H_c = D_{ic} + D_{ix} D_{xc} \]

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 (33), certain structures may require assumptions about the nature of the information set. Thus if we substitute out for the static relations using (34) 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

\[ \{ A c \eta' + A_{xc} (D_{xc} \eta' + D_{xc}) \} \tilde{\xi}_t = \{ B_{cc} + B_{cx} D_{xc} \} \eta' + \{ B_{c\xi} + B_{cx} D_{xc} \} \tilde{\xi}_t \]

which depends on \( \xi_t \) as well as \( \tilde{\xi}_t \). For such a formulation to be informationally feasible in this precise form, the linear combination of states given by \( \{ B_{c\xi} + B_{cx} D_{xc} \} \xi_t \) 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 \( \{ B_{c\xi} + B_{cx} D_{xc} \} \xi_t = \{ B_{c\xi} + B_{cx} D_{xc} \} \tilde{\xi}_t \) thus allowing the top block to be written entirely in terms of state estimates, as

\[ \{ A c \eta' + A_{xc} (D_{xc} \eta' + D_{xc}) \} \tilde{\xi}_t = \{ B_{cc} + B_{cx} D_{xc} \} \eta' + \{ B_{c\xi} + B_{cx} D_{xc} \} \tilde{\xi}_t \]

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) $B_c x$ and $B_c \xi$ are zero.

B Proof of Proposition 1

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

B.1 Forecasting $i_t$

We have (7) which we reproduce here,

$$i_t = H' \xi_t + H_c \eta' \xi_t$$

hence

$$E_{t-1} i_t = H' E_{t-1} \xi_t + H_c \eta' E_{t-1} \xi_t$$

$$= \left( H' + H_c \eta' \right) E_{t-1} \xi_t \quad (35)$$

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

$$i_t - E_{t-1} i_t = H' \xi_t + H_c \eta' \xi_t - \left( H' + H_c \eta' \right) E_{t-1} \xi_t$$

$$= H' \left[ \xi_t - E_{t-1} \xi_t \right] + H_c \eta' \left( \xi_t - E_{t-1} \xi_t \right)$$

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

$$i_t - E_{t-1} i_t = H' \xi_t - E_{t-1} \xi_t + H_c \eta' \beta_t (i_t - E_{t-1} i_t)$$

$$= J'_1 \left[ \xi_t - E_{t-1} \xi_t \right] \quad (36)$$

where

$$J'_1 = [I_n - H_c \eta' \beta_t]^{-1} H'$$

Thus we have, using (36) and (13)

$$E \left[ (i_t - E_{t-1} i_t) (i_t - E_{t-1} i_t) \right] = J'_1 E \left[ (\xi_t - E_{t-1} \xi_t) (\xi_t - E_{t-1} \xi_t) \right] J_t$$

$$= J'_1 P_t J_t \quad (38)$$
B.2 Deriving the updating equation (17).

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

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

and, using (36) and (13)

$$E \left[ (\xi_t - E_{t-1}\xi_t) (i_t - E_{t-1}i_t)' \right] = E \left[ (\xi_t - E_{t-1}\xi_t) (J_t' [\xi_t - E_{t-1}\xi_t])' \right] J_t$$

hence, using (38) and (40),

$$\beta_t = P_t J_t [J_t' P_t J_t]^{-1}$$

and the MSE of the state estimates can be written as

$$M_t = E \left[ (\xi_t - E_{t-1}\xi_t)' (\xi_t - E_{t-1}\xi_t)' \right]$$

$$\beta_t E \left[ (i_t - E_{t-1}i_t) (\xi_t - E_{t-1}\xi_t)' \right] = P_t - \beta_t J_t' P_t = [I_r + \beta_t J_t'] P_t$$

however these do not yet constitute closed form solutions since, via (37), $J_t$ depends on $\beta_t$.

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),

$$\xi_{t+1} - E_t \hat{\xi}_{t+1} = F_c \xi_t + F_c\eta' \xi_t + v_{t+1} - (F_c + F_c\eta') \hat{\xi}_t$$

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

$$P_{t+1} = F_c M_t F_c' + Q.$$
B.4 $P_{t+1}$ and $M_t$ are $H_c$-independent

If we define

$$J_t' = K_t^{-1} H_t',$$

where $K_t$ is the (as yet unknown) matrix that satisfies

$$K_t = (I_n - H_c \eta' \beta_t) \in \mathbb{R}^{n \times n}.$$  \hspace{1cm} (46)

then

$$(J' P_t J_t)^{-1} = \left( K_t^{-1} H_t' P_t H_t \left( K_t^{-1} \right)' \right)^{-1} = K_t' (H_t' P_t H_t)^{-1} K_t$$

and hence

$$\beta_t = P_t H_t' K K_t' (H_t' P_t H_t)^{-1} K_t$$

$$= P_t H_t' (H_t' P_t H_t)^{-1} K_t$$

and thus

$$\beta_t J_t' = J_t (J_t' P_t J_t)^{-1} J_t' = H_t \left( K_t^{-1} \right)' K_t' (H_t' P_t H_t)^{-1} K_t K_t^{-1} H_t'$$

$$= H_t' (H_t' P_t H_t)^{-1} H_t'.$$  \hspace{1cm} (49)

We thus have

$$M_t = \left( I_r - P_t H_t' \left( H_t' P_t H_t \right)^{-1} H_t' \right) P_t,$$  \hspace{1cm} (50)

which implies that

$$P_{t+1} = F_t \left( I_r - P_t H_t' \left( H_t' P_t H_t \right)^{-1} H_t' \right) P_t F_t' + Q,$$  \hspace{1cm} (51)

Both of these expressions imply that the recursions for $M_t$ and $P_t$ do not depend on $K_t$ 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 $\tilde{\beta}_t$ as in (16) then the above formulae are identical to (14) and (15) in Proposition 1. We also have, using (49)

$$\beta_t J_t' = \tilde{\beta}_t H_t'$$

B.5 Derivation of $\beta_t$

Finally we need to obtain an expression for $J_t$ itself, and hence for $\beta_t$.

Equations (37) and (41) imply the seemingly nonlinear equation

$$J_t' = (I_n - H_c \eta' P_t (J_t P_t J_t^{-1})^{-1} H_t').$$
However, using (45) and (48), we obtain

$$K_{t-1}^{-1}H_\xi' = J_t' = \left(I_n - H_c\eta'P_tH_\xi (H_\xi'P_tH_\xi)^{-1} K_t\right)^{-1} H_\xi'$$

which implies

$$\left(I_n - H_c\eta'P_tH_\xi (H_\xi'P_tH_\xi)^{-1} K_t\right) K_{t-1}^{-1} H_\xi' = H_\xi'$$

that is,

$$K_{t-1}^{-1}H_\xi' - H_c\eta'P_tH_\xi (H_\xi'P_tH_\xi)^{-1} H_\xi' = H_\xi'.$$

Recalling (45) once again, we find

$$J_t' = H_\xi' + H_c\eta'P_tH_\xi (H_\xi'P_tH_\xi)^{-1} H_\xi'$$

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

and thus

$$K_t = \left[I_n + H_c\eta'P_tH_\xi (H_\xi'P_tH_\xi)^{-1}\right]^{-1} \quad (54)$$

Using (16), these can be expressed as

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

$$K_t = \left(I_n + H_c\eta' \tilde{\beta}_t\right) H_\xi' \quad (56)$$

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


Svensson & Woodford (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 $\hat{\xi}_t$ 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 forecast error. Using their (16), the true forecast error
in their framework is
\[ Z_t - E_{t-1}Z_t = L (X_t - X_{t|t-1}) + M (X_{t|t} - X_{t|t-1}) + v_t \]
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
\[ \hat{\xi}_t - E_{t-1}\hat{\xi}_t = K \left[ H'_{\xi} \left( \xi_t - E_{t-1}\hat{\xi}_t \right) \right] \]
whereas we show that, in our notation, from (36), 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
\[ \hat{\xi}_t - E_{t-1}\hat{\xi}_t = \beta (i_t - E_{t-1}i_t) = \beta J' \left( \xi_t - E_{t-1}\hat{\xi}_t \right) \]

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,
\[ K = \tilde{\beta} \]
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 \( \beta J' = \tilde{\beta}H'_{\xi} \), hence
\[ \hat{\xi}_t - E_{t-1}\hat{\xi}_t = \beta J' \left( \xi_t - E_{t-1}\hat{\xi}_t \right) = \tilde{\beta}H'_{\xi} \left( \xi_t - E_{t-1}\hat{\xi}_t \right) \]
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). However, neither of these papers note the equivalence of \( M_t \) and \( P_{t+1} \) in the parallel problem or derive convergence conditions.

C Proof of Proposition 2 and Corollary 3.

C.1 Proof of Proposition 2

Since \( \beta_t \) 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, pp 77-81) provide a proof of a unique stable steady state given controllability and detectability as defined...
in the main text for any invertible $P_0$.

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

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

In other words, we are iterating the function $g : P_r \rightarrow P_r$, where $P_r$ denotes the set of all non-negative definite symmetric, real $r \times r$ matrices, and

$$g(P_t) = F \left(I_r - P_t H (H' P_t H)^{-1} H'\right) P_tF' + Q, \quad P_t \in P_r. \quad (58)$$

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

$$\tilde{F}(P_t) = F \left(I_r - P_t H (H' P_t H)^{-1} H'\right) = F \left(I_r - \beta(P_t) H'\right) \quad (59)$$

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

**Lemma 1** The function $g : P_r \rightarrow P_r$ defined by (58) can be expressed, using (59), in the symmetric form

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

**Proof.** Using (59), we have

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

and

$$\tilde{F}P\tilde{F}' = \tilde{F}PF' - F \left(I_r - \beta H'\right) PH\beta' F'$$

but

$$\left(I_r - \beta H'\right) PH\beta' = PH\beta' - \beta H' PH\beta' = PH\beta' - PH (H' PH)^{-1} H' PH\beta' = 0$$

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

**Lemma 2** If $E \in P_r$, then, letting $\tilde{F}_P = \tilde{F}(P)$

$$g(P + E) = g(P) + \tilde{F}_P E \tilde{F}_P' + O(E^2). \quad (61)$$
Thus if we let $Dg_P$ denote the Fréchet derivative of the matrix function $g$ at the point $P \in \mathbb{P}_r$, then

$$Dg_P(E) = \tilde{F}_P E \tilde{F}_P', \quad E \in \mathbb{P}_r.$$ \hfill (62)

**Proof.** We have, using (58)

$$g(P + E) = F(P + E)F' - F(P + E)H(H'PH + H'EH)^{-1}H'(P + E)F' + Q$$

$$= F(P + E)F' - F(P + E)H [(H'PH)(I + (H'PH)^{-1}(H'EH))]^{-1} H'(P + E)F' + Q$$

$$= F(P + E)F' - F(P + E)H [I - (H'PH)^{-1}(H'EH)] (H'PH)^{-1}H'(P + E)F' + Q + O(E^2)$$

$$= g(P) + \tilde{F}_P E \tilde{F}_P' + O(E^2). \hfill (63)$$

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

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

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

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

and thus as a corollary of Proposition 2, stability of the steady state implies that the matrix $F_\xi \left(I_r - \beta(P)H'_\xi\right)$ 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 $\left(I_r - \beta(P)H'_\xi\right) F_\xi$, thus proving the corollary.$\blacksquare$

**C.3 Proof of Corollary 2.**

From the autoregressive representation of the estimated states in (17) it is evident that they have the same autoregressive form as the true states under full information in (9). To derive the implied innovation covariance matrix we have, using (18), (36), (38) and (52),

$$E \left[\begin{array}{c}
\hat{\xi}_{t+1} - E_{t+1} \hat{\xi}_{t+1} \\
\hat{\xi}_{t+1} - E_{t+1} \hat{\xi}_{t+1}
\end{array}\right] = \beta E \left(\varepsilon_{t+1} \varepsilon'_{t+1}\right) \beta' = \beta J'PJ\beta' = \beta H'_\xi PH_\xi \beta'$$

37
but hence, using (16), (15), and exploiting symmetry of $P$ and $M$

\[
E \left[ (\xi_{t+1} - E_t \tilde{\xi}_{t+1}) (\xi_{t+1} - E_t \tilde{\xi}_{t+1})' \right] = PH_{\xi} [H_{\xi}' PH_{\xi}]^{-1} PH_{\xi} \tilde{\beta}' = PH_{\xi} \tilde{\beta}' = (\beta' P)' = (P - M)' = P - M
\]

\[
= Q + F_t M F_t' - M
\]

where the last line follows, after substituting from the steady state of (14).

**D Derivation of joint process for $\xi_t$ and $f_t$ in (24) and proofs of corollaries 4 to 7**

**D.1 Derivation of (24).**

Using (6), (11) and (23) we have

\[
\xi_{t+1} = F_\xi \xi_t + F_\eta \xi_t + v_{t+1}
\]

\[
= (F_\xi + F_\eta') \xi_t - F_\eta' f_t + v_{t+1}
\]

\[
= G \xi_t - F_\eta' f_t + v_{t+1}
\]

For the estimated states we have, using (17), the definition of $G$ in (9), (36) and (6)

\[
\tilde{\xi}_{t+1} = G \tilde{\xi}_t + \beta \tilde{\xi}_{t+1}
\]

\[
= G \tilde{\xi}_t + \beta J' \left[ \xi_{t+1} - G \tilde{\xi}_t \right]
\]

\[
= G \tilde{\xi}_t + \beta J' \left[ \xi_{t+1} - (F_\xi + F_\eta') \xi_t \right]
\]

\[
= G \tilde{\xi}_t + \beta J' \left[ \xi_{t+1} - F_\xi \xi_t - F_\eta' \xi_t + F_\xi \left( \xi_t - \tilde{\xi}_t \right) \right]
\]

\[
= G \tilde{\xi}_t + \beta J' [v_{t+1} + F_\xi f_t]
\]

hence, using (10),

\[
f_{t+1} = (G - F_\eta' - \beta J' F_\xi) f_t + (I - \beta J') v_{t+1}
\]

\[
= [I - \beta J'] F_\xi f_t + (I - \beta J') v_{t+1}
\]

which, using (52) we can also express as

\[
f_{t+1} = \left[ I - \tilde{\beta} H_\xi \right] F_\xi f_t + \left( I - \tilde{\beta} H_\xi \right) v_{t+1}
\]

(65)
Thus we have the joint process for $\xi_{t+1}$ and $f_{t+1}$ as in (24) which we reproduce here:

$$\begin{bmatrix} \xi_{t+1} \\ f_{t+1} \end{bmatrix} = \begin{bmatrix} G & -Fc\eta' \\ 0 & I - \beta H'_{\xi} \end{bmatrix} \begin{bmatrix} \xi_t \\ f_t \end{bmatrix} + \begin{bmatrix} I \\ I - \beta H'_{\xi} \end{bmatrix} v_{t+1}$$

**D.2 Proofs of Corollaries 4 to 7**

**D.2.1 Proof of Corollary 4**

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

$$\xi_{t+1} = [I - GL]^{-1} v_{t+1} - [I - GL]^{-1} Fc\eta' f_t$$

$$= \xi^*_{t+1} - [I - GL]^{-1} Fc\eta' f_t$$

$$= \xi_{t+1} - [I - GL]^{-1} Fc\eta' [I - (I - \beta J') F\xi L]^{-1} (I - \beta J') v_t$$

where $\xi^*_{t+1}$, the full information state process, is as given by (9), and the last line uses the lower block of (24). 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.

**D.2.2 Proof of Corollary 5**

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 $F_\xi$ 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.

**D.2.3 Proof of Corollary 6**

Using (16) we have

$$H^T_\xi \beta = H^T_\xi P H_\xi (H^T_\xi P H_\xi)^{-1} = I_n$$

(66)

If we pre-multiply (65) by $H^T_\xi$ and use (66) we have

$$H^T_\xi f_{t+1} = H^T_\xi [I - \beta H^T_\xi] F_\xi f_t + H^T_\xi (I - \beta H^T_\xi) v_{t+1} = 0$$

thus proving Corollary 6.
D.2.4 Proof of Corollary 7

By inspection of (24), if $F$ has explosive eigenvalues in its sub-matrix $F_{kk}$ and $\beta_k = 0$, then the matrix $\left( I - \tilde{\beta}H^t \right) F_\xi$ will also have explosive eigenvalues, which, from Corollary 3, contradicts stability of the recursion for $P_{t+1}$, thus proving Corollary 7.

E Proof of Proposition 3 and Corollary 1

E.1 Proof of Proposition 3

Using (39) our original definition of $\beta_t$, and (18), we have

$$\beta_t = E \left( (\xi_{t+1} - E_t \xi_{t+1}) \varepsilon_t^t \right) \left( E (\varepsilon_{t+1} \varepsilon_{t+1}^t) \right)^{-1}$$

Conjecture that the proposition is correct, and thus a limiting steady state exists with $P = Q \Rightarrow M = 0$. Recall that since both conditions are identical in the parallel problem of Proposition 1, so we set $F_c = H_c = 0$. Since this steady state replicates full information, the equivalent expressions to (38) and (40) are

$$E (\varepsilon_{t+1} \varepsilon_{t+1}^t) = E \left( H'F_u u_{t+1} (H'F_u u_{t+1})' \right) = H'F_u SF_u^t H$$

and hence, using (20), and (16) we have

$$E (\xi_{t+1} - E_t \xi_{t+1}) \varepsilon_t = E \left( F_u u_{t+1} (H'F_u u_{t+1})' \right) = F_u SF_u^t H$$

and hence, using (20), and (16) we have

$$\beta (Q) = F_u SF_u^t H \left( H'F_u SF_u^t H \right)^{-1}$$
$$= F_u SF_u^t H \left( F_u^t H \right)^{-1} \left( H'F_u \right)^{-1}$$
$$= F_u \left( H'F_u \right)^{-1}$$

and hence

$$\beta (Q) H'Q = F_u \left( H'F_u \right)^{-1} H'F_u SF_u^t = F_u SF_u^t = Q$$
$$\Rightarrow M(Q) = \left( I_r - \beta (Q) H' \right) Q = 0$$
$$\Rightarrow P = Q$$

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 3, that the matrix $F_\xi \left( I_r - \tilde{\beta}(Q)H_\xi^t \right)$ have eigenvalues with real parts strictly less than zero in absolute value. This
third condition is in general non-trivial. Since

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

(since each of its elements are rank \( n \)),

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

and hence

\[ \text{rank} \left( F \xi \left( I_r - \beta'(Q) H' \right) \right) \leq r - n \]

Thus the matrix \( F \xi \left( I_r - \beta'(Q) H' \right) \) may have up to \( r - n \) non-zero eigenvalues. These must be less than unity in absolute value to ensure stability of the recursion (and hence, via Corollary 3, of the filtering errors, \( f_t \)).

\[ \blacksquare \]

**E.2 Proof of Corollary 1**

Using (36) and the same expressions used in deriving (24) we have

\[ \varepsilon_t = A(L) u_t \]  \hspace{1cm} (68)

where \( A(L) = J' \left[ I_r + F \xi \left( I_r - (I_r - \beta J') F \xi L \right)^{-1} (I_r - \beta J') L \right] F_u \)

Given efficiency of the Kalman Filter \( \varepsilon_t \) is a vector white noise process, which, given the first condition in Proposition 3 is of the same dimension, \( s \), as the structural innovations. We can also follow Lippi and Reichlin (1994) and write in the orthonormal form

\[ \nu_t = B(L) \nu_t \]

where \( E(\nu_t \nu'_t) = I_s; \varepsilon_t = \Sigma_e \nu_t; u_t = \Sigma_u \nu_t \)

\[ \Rightarrow \Sigma_e \Sigma_e' = J'P; \Sigma_u \Sigma_u' = S; B(L) = \Sigma_e^{-1} A(L) \Sigma_u \]

such that \( B(L) \) is a Blaschke Matrix (defined as in Lippi & Reichlin, p 310-311).

In case a), where all three conditions for asymptotic invertibility are satisfied, we have the simple result that \( A(L) \) tends to the scalar matrix \( A = J'F_u \). Since both \( \varepsilon_t \) and hence \( u_t \) can be recovered from the history of the observables the representation of the \( i_t \) is fundamental.

In case b) the white noise property of both \( u_t \) and \( \varepsilon_t \) directly implies that \( B(L) \) must be a Blaschke matrix (Lippi & Reichlin, p 311). The order of the AR and MA polynomials in the representation of \( i_t \) is increased by the number of "nonbasic" roots in \( B(L) \) (Lippi & Reichlin, p315) hence in this case the representation is nonbasic as well as nonfundamental.\[ \blacksquare \]
E.3 Asymptotic invertibility implies filtering errors have zero variance

Using (24) and (20) the true filtering error process can be written as

\[ f_{t+1} = \left[ I - \left( I - \tilde{\beta}H_\xi' \right) F_\xi L \right]^{-1} \left[ I - \tilde{\beta}H_\xi' \right] F_u u_{t+1} \]

but, using (67) we have

\[ \left[ I - \tilde{\beta}H_\xi' \right] F_u = \left[ I - F_u (H' F_u)^{-1} H_\xi' \right] F_u = 0 \]

hence the filtering error has zero innovations hence is always zero.