RESERVE BANK OF NEW ZEALAND

DP2003/02

On applications of state-space modelling in macroeconomics

Olivier Basdevant

April 2003

JEL classification: C30, C50, E17

Discussion Paper Series

DP2003/02

On applications of state-space modelling in macroeconomics

Abstract[1]

This paper reviews the literature on applications of state-space modelling to macroeconomic questions, with four examples related to modelling unobserved trends, transition across different steady states, expectations formation and forecasting/data revision issues. Due to the flexibility of the state-space approach, it is both a useful tool for research purposes and highly useful in addressing practical issues. In many cases, state-space modelling offers the possibility of building encompassing models, and formulating rather complicated problems in a simple manner.

[1] I thank N Bjorksten, F Citu, S Hall, C Plantier and K West for their comments. Email address: olivier.basdevant@laposte.net. All errors and omissions are the responsibility of the author. The views expressed in the paper are those of the author and do not necessarily reflect those of the Reserve Bank of New Zealand.

1 Introduction

State-space modelling is increasingly used in economics, and there are already several exhaustive academic presentations of the Kalman filter and state-space models, such as the textbook of Harvey (1989). Nevertheless there are very few contributions that discuss how those models can be used in practice and why they are such a powerful tool for practitioners. Thus, the idea of this survey is to present state-space modelling in the field of macroeconomics, with policy-related implications.

There are two main types of problems in macroeconomics that can usefully be addressed using state-space models. Firstly, those in which some variables are unobservable, examples being the potential output or the NAIRU. Secondly, those containing coefficients that are inherently time-varying, making economic relationships potentially unstable (one example being the Phillips curve).

In all of the cases presented in this paper, there are two main advantages of state-space modelling. Firstly, it is straightforward to implement, and often provides a simple representation of relatively complex problems. Secondly, it provides an encompassing approach, which makes it possible to assess the relative merits of different approaches to the same problem.

This literature review is divided into five sections, four of which are related to specific economic problems. Section 2 provides a general introduction to state-space modelling, briefly presenting the general structure of state-space models and the Kalman filter. This section introduces the general notations adopted here, as well as the general assumptions – and their implications – that are usually made and how binding they are.

Section 3 discusses ways of modelling unobserved trends and fluctuations in the economy. All of the existing approaches, from the HP filter to common trends, can be represented in one unified way using state-space modelling. This has some implications for the way that we think about the distinction between cycles and trends, which – as an example – lies at the heart of the RBC literature and dynamic general equilibrium models.

Section 4 discusses how to model transition generically, which can equally be related to the transition of centrally planned systems into market economies, or to such issues as convergence. This is in fact one of the most natural applications of state-space modelling: in essence the technique captures how an equilibrium may be changing over time, or how an economy may converge from one state to another. As we will see within a simple framework, it is possible to make sense of various types of structural change.

Section 5 discusses how to model expectations formation and more generally how state-space modelling can be used to address the Lucas critique. While most macro-economic models are built either on adaptive or rational expectations (and sometimes both), a recent literature has begun to explore how to model a learning process in expectations formation. The advantage of learning processes is that they permit expectations to gradually converge towards rational, and make clear under which conditions they become rational.

Finally, in section 6 we address two issues that are particularly important for practitioners: forecasting and the problem of data revision. Forecasting properties of models can also be improved by using state-space modelling, when errors exhibit some patterns or when data are subject to revisions, which can be of particular importance (see Orphanides and Van Norden 1999).

2 State-space models

This section briefly presents the general characteristics of state-space models and the Kalman filter, and shows how the Kalman filter can be applied to a wide range of models.

2.1 State-space model

A standard state space formulation can be represented as follows.[2] Let

$$Y_t = Z_t A_t + \varepsilon_t \tag{1}$$

be the measurement equation, where Y_t is a vector of measured variables of dimension $n \times 1$, A_t is the state vector of unobserved variables of dimension[3] $p \times 1$, Z_t is a matrix of parameters of dimension $n \times p$ and $\varepsilon_t \sim N(0, H_t)$. The state equation is then given as:

$$A_t = T_t A_{t-1} + \eta_t \tag{2}$$

where T_t is a matrix of parameters and $\eta_t \sim N(0, Q_t)$.

Q_t and H_t are sometimes referred to as the hyper-parameters of the model, to distinguish them from the other parameters.

The specification of the state space system is completed by two further assumptions: first, that the initial vector A_0 has a mean a_0 and covariance matrix P_0 and second that the disturbances ε_t and η_t are uncorrelated with each other in all time periods, and uncorrelated with the initial state. This implies that:

$$\forall (s,t) \quad E(\varepsilon_t . \eta_s') = 0 \tag{3}$$

and

$$\forall t \quad E(\varepsilon_t . A_0') = 0 \tag{4}$$

---

2 See Kalman (1960, 1963) for the original contributions, Harvey (1987 and 1989) for exhaustive presentations and Cuthbertson et al (1992) or Hall (1993) for applied contributions.

3 Note that a priori there is no restriction on the sizes of Y and A.

2.2 The Kalman filter

Let a_t be the optimal estimator of A_t based on the observations up to and including Y_t, P_t its covariance matrix, $a_{t|t-1}$ the estimator based on the information available in $t-1$ and $P_{t|t-1}$ its covariance matrix.

The *predicted* estimate of A_t is $a_{t|t-1}$, and is defined by:

$$a_{t|t-1} = T_t a_{t-1} \tag{5}$$

With a covariance matrix defined as: $P_{t|t-1} = T_t P_{t-1} T_t' + Q_t$.

The *filtered* estimate[4] of A_t is a_t and is updated from $a_{t|t-1}$ when Y_t is known (see Harvey (1989) for details):

$$a_t = a_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1}\left(Y_t - Z_t a_{t|t-1}\right) \tag{6}$$

where $F_t = Z_t P_{t|t-1} Z_t' + H_t$ and $P_t = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1}$. Those three relations are the Kalman filter equations.

The *smoothed* estimate of A_t is $a_{t|T}$, and is updated from a_t using the whole set of information available:

$$a_{t|T} = a_t + P_t^*\left(a_{t+1/T} + T_{t+1} a_t\right) \tag{7}$$

The smoothed estimates are calculated working backwards from the last value of the filtered estimate with $P_t^* = P_t T_{t+1}' P_{t+1/t}^{-1}$,

$$P_{t|T} = P_t + P_t^*\left(P_{t+1/T} + P_{t+1/t}\right) P_t^{*\prime}, \; a_{T/T} = a_T \text{ and } P_{T/T} = P_T.$$

For a Gaussian model, the likelihood function can be written as (see Crowder 1976 and Schweppe 1965):

[4] The filtered estimate is also a minimum mean square estimate of A_t. It is unconditionally unbiased and the unconditional covariance matrix of the estimator is the P_t matrix given by the Kalman filter. Proof of these results can be found in Anderson and Moore (1979), Ducan and Horn (1972) or Harvey (1981).

$$\begin{aligned} Ln\ell = {} & -\frac{N(T-k)}{2} Ln(2\pi) - \frac{1}{2}\sum_{t=k}^{T} Ln\left|Z_t P_{t|t-1} Z_t' + H_t\right| \\ & -\frac{1}{2}\sum_{t=k}^{T} \nu_t'\left(Z_t P_{t|t-1} Z_t' + H_t\right)^{-1}\nu_t \end{aligned} \tag{8}$$

where N is the number of elements contained in the vector Y_t, k the number of periods needed to derive estimates of the state vector and with $\nu_t = Y_t - Z_t a_{t|t-1}$. The vector ν_t can be interpreted as the vector of prediction errors, as the conditional mean is also the minimum mean square estimation of Y_t. Hence, the likelihood function can be expressed as a function of the one-step-ahead prediction errors, suitably weighted. In general, then, the Kalman filter will provide estimates of the unobserved variable A_t, while estimates of any other desired parameters (including hyper-parameters) can obtained by MLE algorithm as adapted by Shumway and Stoffer (1982).

2.3 Properties

In the presentation above, some assumptions were made about correlation of residuals between measurement and state equations, initial conditions and linearity. In this section we briefly discuss those assumptions.

2.3.1 Correlation between disturbances in the state and measurement equations

If the measurement and transition equation disturbances are correlated, the Kalman filter needs to be modified. Derivations of the following results can be found in Jazwinski (1970). Consider: $E(\eta_t . \varepsilon_s') = G_t$ if $t=s$ and 0 otherwise, where G_t is a known matrix. The prediction equations remain unaltered as far as Y_t remains unknown, but the updating equation becomes:

$$a_t = a_{t|t-1} + \left(P_{t|t-1} Z_t' + G_t\right) F_t^{-1}\left(Y_t - Z_t a_{t|t-1}\right) \tag{9}$$

with $F_t = Z_t P_{t|t-1} Z_t' + Z_t G_t + G_t' Z_t' + H_t$ and

$$P_t = P_{t|t-1} - \left(P_{t|t-1} Z_t' + G_t\right) F_t^{-1}\left(Z_t P_{t|t-1} + G_t'\right).$$

2.3.2 Importance of initial conditions

If prior information is available on all the elements of A_0, then A_0 has a proper prior distribution with known mean a_0 and bounded covariance matrix P_0. The Kalman filter then yields the exact likelihood function of the observations, Y_t, via the prediction error decomposition. Genuine prior information is rarely available. For univariate series the Kalman filter can always be initialised with the mean and covariance matrix of the unconditional distribution of A_t when A_t is stationary (see Harvey 1981).

The Kalman filter can be viewed as a Bayesian approach to economic modelling: for given prior information it is possible to revise those priors, ie to get posterior estimates, and the Kalman filter gives the optimal ones. More precisely, the Kalman filter consists in choosing a_t so that it maximises the probability that the observed Y_t would take place (see Harvey 1989), ie a_t is solution of $Max\ p(A_t|Y_t)=p(A_t,Y_t)/p(Y_t)$. The updating process can then be understood as follows: before observing data the prior density of A_t is $p(A_{t-1})$ and once data are observed the posterior density is $p(A_t|Y_t)$.

One potential problem is the dependency of the posterior estimates on prior information. In practice, initial values will very often be assigned using a fixed coefficients estimation technique when estimating time-varying coefficients, or by applying other filtering methods (such as the HP filter) when estimating unobserved variables.

2.3.3 OLS, recursive OLS and the Kalman filter

A key variable in the estimation of state-space models is the relative smoothness of the unobserved variable, which is governed by the relative size of the error variances in measurement and state equations. As an example, in the univariate case, the lower the ratio H_t/Q_t (referred to as the signal to noise ratio), the greater the explanatory power accorded to the unobserved variable and the better the fit of the measurement equation. For very large values of Q_t, the unobserved variable may absorb all of the residual variation in the measurement equation. Alternatively, if Q_t is zero

and $T_t=Id$, the filtered estimate will correspond to recursive OLS and the smoothed estimate will be the standard full sample OLS estimates.

In practice, most studies fix the signal-to-noise ratio so that the estimated unobserved variable is relatively smooth, with fluctuations which are judged to be reasonable from one period to the next.

2.3.4 Dealing with non-linearities

The only major constraint of the Kalman filter is that it imposes a linear structure. If the relations of interest are believed to be non-linear, a useful approach is to linearise models with a Taylor series approximation. Let the state-space model be:

$$Y_t = h(A_t, \varepsilon_t) \tag{10}$$

$$X_t = f(X_{t-1}, \nu_t) \tag{11}$$

In practise the residuals are unknown, but we can refer (with the subscript ~) to the measurement and state variables as the values without any noise: $\tilde{a}_{t|t-1} = f(a_{t-1}, 0)$ and $\tilde{Y}_t = h(\tilde{a}_{t|t-1}, 0)$. $\tilde{a}_{t|t-1}$ is the a posteriori estimate of the state variable, from the previous time step $t-1$. The first order approximation of the model in the neighbourhood of a_{t-1} is then:

$$Y_t = \tilde{Y}_t + J_{h,x}(A_t - \tilde{a}_t) + J_{h,\varepsilon}\varepsilon_t \tag{12}$$

$$A_t = \tilde{a}_{t|t-1} + J_{f,x}(A_{t-1} - a_{t-1}) + J_{f,\nu}\nu_t \tag{13}$$

where $J_{\phi,R}$ is the Jacobian of ϕ with respect to R: $J_{\phi R} = \frac{\partial \phi}{\partial R}(\tilde{a}_{t|t-1})$. For a given value of a_{t-1} (and therefore $\tilde{a}_{t|t-1}$ as well), this model is linear in the state A_t, and the standard formulae of the Kalman filter can be applied.

It should be noted that the distributions of error terms are no longer normal after undergoing their respective non-linear transformations.

It is simply an ad hoc state estimator, which nevertheless still has a justification as a minimum squared error estimator.

2.4 Conclusion

The power of the Kalman filter comes from the fact that these estimation formulae work for any model that can be put into state space form. Once we have formulated a state space version of a model, the actual estimation is easy. The real art of using the Kalman filter does not lie in the formulae used to estimate parameters but rather in learning to write problems in state space form.

Kalman filtering can be viewed as belonging to the Bayesian class of estimators. Before starting the estimation process, one has to specify the prior knowledge available, ie a vector of prior values for the parameters and hyper-parameters ($a_{t|t-1}$ and $P_{t|t-1}$). The real problem is then to ask what our priors are and to address their relevance. Modelling in a state-space form makes clear what our priors are, ie not only our priors on the structure of the model but also on the prior values of coefficients. The issue of the relevance of priors can then be addressed using the various tests available, as well as investigating alternative specifications.

3 Modelling unobserved trends

In this section we investigate some applications of state-space modelling in modelling trends. To begin with, we explore the modelling of trends and cycles, starting with the standard HP filter reformulated in a state-space model (see section 3.1). An advantage of this approach is that it is an encompassing one that permits a large degree of flexibility. Although the first part of this section is mostly related to standard time series analysis, we also explore how to explicitly integrate structural modelling, namely by integrating a production function to estimate a capital stock index (see section 3.2).

3.1 Time series models and state-space models

It is quite a standard practice in economic modelling to distinguish the long-run properties of a system from its fluctuations around the long-run trend. There are various approaches to do this, but what is interesting is that most of these (eg the HP filter, common trends, the multivariate HP filter) can be conveniently represented within a state-space model. We first start with the single variable approach, and then turn to systems of equations and how to derive common trends.

3.1.1 The Hodrick-Prescott filter

The HP filter gives an estimate of the unobserved variable as the solution to the following minimisation problem:

$$\underset{\{y_t^*\}}{Min} \quad \sum_{t=1}^{T} \frac{1}{\sigma_0^2}\left(y_t - y_t^*\right)^2 + \frac{1}{\sigma_1^2}\left(\Delta^2 y_t^*\right)^2 \qquad (14)$$

Where y is the observed variable, y^* is the unobserved variable being filtered, σ_0^2 is the variance of the cyclical component $y - y^*$ and σ_1^2 is the variance of the growth rate of the trend component. This problem is of course invariant to a homothetic transformation, therefore what matters is the ratio $\lambda_1 = \sigma_0^2 / \sigma_1^2$. Hodrick and Prescott suggest some parameterisation of λ_1 depending on the frequency of data[5]. Following Harvey (1985), the HP filter can be written in a state space form as follows. The measurement equation defines the observed variable as the sum of its trend and fluctuations around the trend:

$$y_t = y_t^* + e_t \tag{15}$$

with $e_t \sim N(0, \sigma_0^2)$. The state equations define the growth rate of the trend that is accumulated to compute the trend itself:

$$y_t^* = g_{t-1} + y_{t-1}^* + \nu_{1,t} \tag{16}$$

$$g_t = g_{t-1} + \nu_{2,t} \tag{17}$$

with $\nu_{1,t}=0$ and $\nu_{2,t} \sim N(0, \sigma_0^2/\lambda_1)$. $\nu_{2,t}$ is the change in the growth rate of the filtered series or trend. In other words, the change in the trend follows a random walk. Then to get the HP filter estimate one has to use the whole set of information to derive y^* (as done in the minimisation problem (14)), ie to take the smoothed estimate provided by the Kalman filter.

3.1.2 The HP Multi-Variate filter and the Kalman filter

The Hodrick Prescott Multi-Variate (HPMV) filter is an alternative way of estimating unobserved variables, developed at the Bank of Canada (see Laxton and Tetlow (1992)). This method stems from use of the standard HP filter (see Hodrick and Prescott (1997)), augmented by relevant economic information. It has been used by the Central Banks of Canada and New Zealand to estimate potential output (see Butler (1996) and Conway and Hunt (1997)) and by OECD (1999) to estimate the NAIRU. Harvey (1985) explains how to reproduce the simple HP filter with the Kalman filter and Boone (2000) extends this to the HPMV filter. This is done in two steps. Firstly, the minimisation problem is written as a state space model.

[5] 100 for annual data, 400 for semi-annual data and 1600 for quarterly data.

Secondly, restrictions are imposed on the variances of the equations of the state space model, to reproduce the balance between the elements of the minimisation programme.

The HPMV filter seeks to estimate the unobserved variable as the solution to the following minimisation problem:

$$\underset{\{y_t^*\}}{Min} \quad \sum_{t=1}^{T} \left(y_t - y_t^*\right)^2 + \lambda_1 \left(\Delta^2 y_t^*\right)^2 + \lambda_2 \zeta_t^2 \tag{18}$$

With λ_1 and λ_2 given. This is a basic HP filter, augmented with the residuals ζ_t taken from an estimated economic relationship:

$$z_t = \beta y_t^* + dX_t + \zeta_t \tag{19}$$

where z is another explanatory variable can be explained by the unobserved variable y^* and X is a matrix of other exogenous variables. The residuals ζ are normal with a covariance matrix H_2.

As for the simple HP filter, the smoothing constants λ_1 and λ_2 reflect the weights attached to different elements of the minimisation problem. The estimated unobserved variable is not only a simple moving average going through the observed series, but is also modelled to give a better fit to the economic relationship. The HPMV filter can also be reproduced by a Kalman filter, following a similar methodology. In essence the problem is simply to add the additional structural equation in the set of measurement equations:

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 1 \\ \beta \end{pmatrix} y_t^* + \begin{pmatrix} 0 \\ d \end{pmatrix} X_t + \begin{pmatrix} e_t \\ \zeta_t \end{pmatrix} \tag{20}$$

with: $(e_t \quad \zeta_t)' \sim N(0, H)$. The covariance matrix H is defined as:

$$H = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_0^2/\lambda_2 \end{bmatrix} \tag{21}$$

The two state-equations are the same as for the standard HP filter (see equations (16) and (17)). The novelty is the term λ_2, which gives a balance between the HP filter and the economic information embodied in the additional equation. A high value for λ_2 corresponds to a better fit of the economic relationship, and an unobserved variable that can depart significantly from the observed variable.

There are two advantages to be gained from reproducing an HPMV filter with the Kalman filter. Firstly, the method is done in one go (while the method proposed by Laxton and Tetlow (1992) is a multi-step procedure), and also allows estimation of the hyper-parameters. Thus, although it is rarely done in practice, it is possible to freely estimate the parameters λ_1 and λ_2 instead of giving these essentially arbitrary values.

3.1.3 Modelling fluctuations

A potential drawback of the methods proposed above is that they do not explicitly model fluctuations around the trend. Fluctuations can be only deduced from the difference between the filtered trend and the actual values. Thus, there is no distinction between cyclical fluctuations and noises, which can substantially bias the filtered trend when using a standard HP filter (see Harvey and Jaeger 1993, Guay and St Amant 1996). To further explore fluctuations it is possible to extend the HPMV filter:

$$y_t = y_t^* + y_t^r + \varepsilon_t \tag{22}$$

$$y_t^* = y_{t-1}^* + g_{t-1} + \eta_{1,t} \tag{23}$$

$$g_t = g_{t-1} + \eta_{2,t} \tag{24}$$

$$y_t^r = B(L)y_{t-1}^r + \eta_{3,t} \tag{25}$$

where y is a vector of variables, y^* the permanent component – or the trend, y^r the transitory component – or the fluctuation around the trend, g the growth rate of the permanent component and ε, η_1, η_2 and η_3 the residuals. The major difference between the standard HPMV filter presented above is that there is an explicit modelling of the fluctuation around the trend. There are various ways of

modelling it. Stock and Watson (1989) and Garratt and Hall (1996) propose a finite auto-regressive form, while Harvey and Jaeger (1993), Scott (2000), Proietti (2000) and Harvey (2001) propose the following specification in the univariate case:

$$y_t^r = \rho\cos(\tau)y_{t-1}^r + \rho\sin(\tau)y_{t-1}^f + \xi_t \tag{26}$$

$$y_t^f = -\rho\sin(\tau)y_{t-1}^r + \rho\cos(\tau)y_{t-1}^f + \xi_t^* \tag{27}$$

where ρ is a parameter representing the magnitude of the cycle, τ is its frequency and ξ and ξ^* the residuals. An obvious advantage with this specification is that it allows us to more clearly identify the magnitude and periods of fluctuations. Harvey (2001) uses such a specification to analyse the similarities in fluctuations of the US GDP and an investment time series. Harvey develops this formulation in a wider perspective, namely to test unit roots in an unobserved components model. Equations (26) and (27) can be used in parametric tests of unit roots where auto-correlation of residuals is captured by y^r, the test of unit root itself being performed on the stochastic nature of y^*, ie on values of $Var(\eta_{1t})$ and $Var(\eta_{2t})$ significantly different from zero. Scott (2000) builds a similar approach to derive estimates of potential output, identified by y^*, and output gap, identified by y^r.

This kind of approach also encompasses other decompositions of series into trends and cycles, such as the Beveridge-Nelson decomposition (see Beveridge and Nelson 1981). Basically this decomposition is applied for integrated time series (possibly augmented by a deterministic drift). Let y^{bn} be the trend obtained by this decomposition. It is defined as follows:

$$y_t^{bn} \equiv \lim_{\tau\to+\infty} E_t\left(y_{t+\tau} - \tau g\right) = y_t + \lim_{\tau\to+\infty} \sum_{i=1}^{\tau} E_t\left(\Delta y_{t+i} - g\right) \tag{28}$$

where g is the growth rate of the deterministic trend. The cycle is then defined as:

$$y_t^c \equiv -\lim_{\tau\to+\infty} \sum_{i=1}^{\tau} E_t(\Delta y_{t+i}-g) \qquad (29)$$

The problem is then to model the growth rate of the variable. To do so, Morley (2002) and Morley et al. (2002) propose to formulate this problem in a state-space form. The growth rate $\Delta y_t - g$ is explained as an autoregressive process like the one given in equation (25).

$$\Delta y_t - g = h' X_t \qquad (30)$$

with

$$X_t = TX_{t-1} + \eta_t \qquad (31)$$

where h is a vector of weights for each element of X_t and T is a matrix with all its eigenvalues within the unit circle. Thus for any $\tau > 0$ the following holds:

$$E_t(\Delta y_{t+\tau} - g) = h' T^{\tau} X_t \qquad (32)$$

Hence the Beveridge-Nelson trend:

$$y_t^{bn} = y_t + h'T(I-T)^{-1} X_t \qquad (33)$$

In essence this approach is very similar to the state-space model given by equations (22) to (25). What Morley et al. (2002) show is that the difference lies in the representation of the disturbance η_1 in equation (23) and the cycle.

A common assumption in the literature is that the trend and the cycle are not correlated. Beveridge and Nelson take exactly the opposite point of view, assuming that the correlation is exactly equal to one. Morley et al. show that integrating a correlation between the cycle and fluctuations in the state-space form allows one to reproduce exactly the Beveridge-Nelson decomposition. The remaining – crucial – problem is to understand why the Beveridge-Nelson decomposition tends to give a totally different picture from a standard unobserved component decomposition. As shown by Enders (1995) and Morley et al (2002), the difference between the

two approaches come from the different view on the correlation between trend and cycle, as Beveridge and Nelson consider a perfect and negative correlation while the unobserved component decomposition assumes no correlation at all. Again, formulating the problem in a state-space form is particularly useful, as it shows the need to estimate the correlation between the trend and the cycle, hence reconciling those two viewpoints.

3.1.4 Identifying common trends

In the multivariate case the model proposed in equations (21) to (24) can easily be adapted to identify the common trends driving the economy[6] (see Stock and Watson 1988, 1989, Garratt and Hall 1996 or Harvey 2001). The model is now written as:

$$y_t = \Pi y_t^* + y_t^r + \varepsilon_t \qquad (34)$$

$$y_t^* = y_{t-1}^* + g_{t-1} + \eta_{1,t} \qquad (35)$$

$$g_t = g_{t-1} + \eta_{2,t} \qquad (36)$$

$$y_t^r = B(L) y_{t-1}^r + \eta_{3,t} \qquad (37)$$

Where Π is a matrix of weights, and testing the rank of Π will provide the number common trends. The very fruitful contribution of this representation is that it encompasses almost any kind of time-series model. This methodology is applied by Pandher (2002), who develops a forecasting model using a common trend approach. The contribution of this paper is to show that integrating accounting identities can improve the efficiency of forecasts. More precisely, the model considered by Pandher introduces some constraints on the matrix Π and the covariance matrix of ε_t in order to introduce long-run identifying relations à la Blanchard and Quah (1989) and also accounting relations. Those are done as follows: for a variable i in the vector y_t, all the coefficients of the matrix Π are fixed to either 1,

[6] Identifying r common trends among a set of n variables is equivalent to identify $n-r$ cointegration relations.

-1 or 0 and ε^i is set equal to zero. Thus the flexibility of the state-space form allows us to build a canonical model were we can integrate a various range of identifications.

3.2 Modelling the capital stock as an unobserved variable

In the previous section it was demonstrated how state-space modelling offers a flexible and encompassing tool to analyse time series. It is also possible to extend the analysis to integrate more structural relations. The example provided here is the one of Hall and Basdevant (2002), who estimate the Russian capital stock as the Russian statistical offices do not report any data on it.

The underlying technology is assumed to be Cobb-Douglas. Output is a measured variable, which is generated by the following measurement equation:

$$y_t = \alpha_0 + \alpha_1 k_t + \alpha_2 n_t + \varepsilon_t \tag{38}$$

where y is GDP, n total employment (both being expressed in natural logarithms) and k is the unobserved capital stock. It is generated by the following state equation:

$$k_t = k_{t-1} + i_t - \delta_{t-1} + \eta_{1t} \tag{39}$$

This is a linearised capital stock accumulation function where $i_t = Ln(1+I_t/K)$ is investment expressed in proportional terms and δ_{t-1} is the unobserved rate of depreciation. It is assumed to be generated by a second state equation:

$$\delta_t = \delta_{t-1} + \eta_{2t} \tag{40}$$

So the depreciation rate follows a random walk which allows it to freely increase to allow for the higher rate of scrapping during the transformation period in Russia. Together, these three equations form a simple state space model and the Kalman Filter may be used to estimate the unobserved effective capital stock and the rate of depreciation.

The parameter α_1 is calibrated on the basis of the approximate share of total income going to capital and labour in the Russian economy, hence $\alpha_1=0.6$ and $\alpha_2=0.4$. Further, the scaling of the unobserved capital stock determines α_0. This parameter may therefore be fixed at any arbitrary level and can simply change the units of measurement of the capital stock and so we set a value of zero for this parameter. Applying the Kalman filter to the specific problem outlined above using monthly data from May 1994 to October 1998 then yields an estimate of the effective Russian capital stock. The results of this procedure are presented in figure 1, which contrasts the effective capital stock (the dotted line), with a scaled time series for the logarithm of GDP (the solid line). The smoothed capital stock shows a sharp decline, especially in 1998. Unfortunately there are no official comparable capital stock figures for this period, although the annual data which is available up to 1996 has, so far, recorded no decline in the capital stock at all.

Figure 1:
Estimate of the effective Russian capital stock

This section has proposed a technique for estimating the stock of capital in an economy, which is undergoing transition and where the official data is either unavailable or believed to be highly inaccurate. This technique has been applied to the case of Russia using monthly data over the period 1994-1998. It suggests that over this period there has been a substantial reduction in the stock of capital of around 15 per cent.

4 Modelling transition

A key aspect of transition is how we can understand a gradual change from one given economic structure to another one (say, a transition from central planning to a market economy). This can be done by reconsidering the framework of cointegration, where some parameters are allowed to vary through time as a result of transition (see section 3.1). Transition can also be viewed as a move from a low level of development to a higher one. Again, a state-space model is a particularly appropriate tool with which to analyse the convergence problem (see section 4.2).

4.1 Change in causality

An important aspect of structural change is its implications on the causal structure of a model. There is a relatively wide literature on how to integrate structural breaks, but the problem investigated here is the impact of structural change on the causal structure linking a set of variables. Barassi et al (2000, 2001) propose the following formulation:

$$\Delta Y_t = B(L)\Delta Y_{t-1} + \Pi_t Y_{t-1} + \varepsilon_t \tag{41}$$

Then some elements of the matrix Π are modelled as time-varying parameters:

$$\exists J / \forall i \in J \quad \pi_{i,t} = \pi_{i,t-1} + \eta_{i,t} \tag{42}$$

When $Var(\eta_{i,t})=0$ the model reduces to a standard cointegrated system à la Johansen (1988, 1991). A traditional decomposition of the matrix Π is to distinguish cointegrating vectors (identified by a

matrix β) and weights for each long-run relation (identified by a matrix α), where $\Pi=\alpha\beta$. Most of the literature has focussed on the case where α is a constant and β is subjected to breaks following the seminal work of Perron (1989) on unit root tests in presence of breaks in deterministic components.[7] In the next two sections we discuss how some studies have used the Kalman filter to either investigate a change in the long-run relation or a change in the causality, ie a change in the weights.

4.1.1 Change in the structure of the cointegrating vectors

In the case of a changing structure in the cointegrating vectors themselves it is possible to formulate the problem in a state-space form. A general presentation would be as follows:

$$\Delta Y_t = B(L)\Delta Y_{t-1} + \alpha\beta_t' X_{t-1} + \varepsilon_t \tag{43}$$

$$\beta_t = \beta_{t-1} + \eta_t \tag{44}$$

The example taken here is the methodology proposed by Haldane and Hall (1991) on the relationship between the pound sterling, the US dollar and the deutschmark (DM).[8] The starting point of their analysis is the dollar-DM polarisation during the seventies and eighties, ie periods of dollar strength concomitant with DM weakness vis-à-vis other European currencies. The inception of the European monetary system (ERM) in 1979 contributed to weaken this polarisation. The model is set up as follows:

$$e_t^{DM} = \mathrm{A}_{1,t} + A_{2,t} e_t^{\$} + \varepsilon_t \tag{45}$$

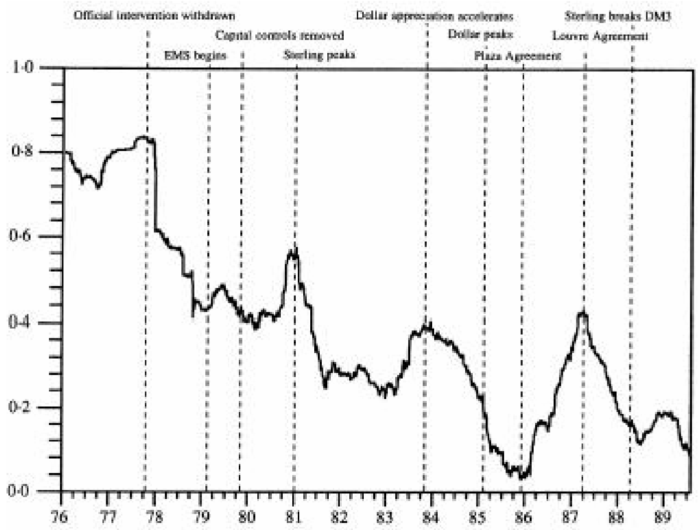
$$\forall i \in \{1,2\} \quad \mathrm{A}_{t,t} = A_{t,t-1} + \eta_{i,t} \tag{46}$$

7 Hansen and Johansen (1999) suggest graphical procedures to evaluate the constancy of the rank of $\alpha\beta'$, and show that it is possible to apply the tests of Ploberger et al. (1989) and Nyblom (1989) to test the constancy of long-run parameters. Seo (1998) extends those results to define an LM test for structural change in α and β separately.

8 See also Manning (2002) for an application on South-East Asian equity markets.

where e^{DM} is the sterling-DM exchange rate and $e^{\$}$ the DM-dollar exchange rate. As the relation weakens one would expect the coefficient A_2 to converge towards zero. The pattern obtained by Haldane and Hall is as follows, using daily data between January 1976 and August 1989:

Figure 2: Coefficient on $e^{\$}$



Source: Haldane and Hall (1991)

It suggests that the relationship between the sterling and the dollar has declined during the period studied.

4.1.2 Change in the structure of the weights of the cointegrating vectors

Barassi et al (2000, 2001) propose not only to apply the existing techniques to shifts in the vector α, but also to determine an optimal way of detecting the presence of multiple breaks. To do so they emphasise that recursive estimations using OLS will not be a suitable technique in this case as the null hypothesis is that

parameters are constant. The model formed by equations (41) and (42) with the restriction that β is constant and can be rewritten as follows:

$$\Delta Y_t = B(L)\Delta Y_{t-1} + \alpha_t \beta' X_{t-1} + \varepsilon_t \tag{47}$$

Then some elements of the matrix Π are modelled as time-varying parameters. Those of the elements of α_t that are time-varying are modelled as follows:

$$\alpha_{i,t} = \alpha_{i,t-1} + \eta_{i,t} \tag{48}$$

It is then tested with simulated data where α is the only time-varying parameter. In the case of multiple breaks the use of a Kalman filter will provide consistent estimates of the time-varying speed of adjustments and will in principle detect quite accurately the timing as well as the magnitude of the breaks. The problem of this method is a relatively wide margin of error, which suggests that the efficiency of the estimates is relatively limited. This could come from the assumption of a fixed variance for the time-varying parameter. More precisely, one could also allow the hyper-parameters to vary, allowing them to be higher during periods of breaks (see Gordon and Smith 1988, 1990 and section 6.1) or to model an ARCH process within a state-space form.

4.2 Modelling convergence

In the previous section we have investigated how the causal structure within a given system may change over time. A related issue is to analyse convergence, which can also be conveniently done in a state-space model as shown by Hall et al. (1997). Consider that we want to investigate the convergence of a certain linear combination of two variables y^1 and y^2 towards zero, say y^1+A+By^2 with A and B two known parameters. We can define the variable ε_t as follows:

$$\varepsilon_t = y_t^1 - A - By_t^2 \tag{49}$$

Naturally we do not assume anything at this stage about the process followed by ε_t, hence there is no reason to consider it a priori as a white noise. The unconditional convergence in probability is then defined as:

$$\underset{t\to+\infty}{p\,lim}\quad \varepsilon_t = 0 \tag{50}$$

A sufficient but non necessary condition for convergence in probability is that the two following conditions hold:

$$\underset{t\to+\infty}{Lim}\quad E\left(y_t^1 - A - By_t^2\right) = 0 \tag{51}$$

$$\underset{t\to+\infty}{Lim}\quad Var\left(y_t^1 - A - By_t^2\right) = 0 \tag{52}$$

The problem is then to identify what is the nature of ε_t. As an example if it follows a Gaussian of parameters $N(0,\sigma^2)$ then convergence occurs in expectations, but not in probability as $\underset{t\to+\infty}{Lim}\ Var\left(y_t^1 - A - By_t^2\right) = \sigma^2 \neq 0$. More generally convergence in probability will occur when for a given general process $N\left(\mu_t, \sigma_t^2\right)$ the parameters converge with zero, ie when $\underset{t\to+\infty}{Lim}\ \mu_t = 0$ and $\underset{t\to+\infty}{Lim}\ \sigma_t^2 = 0$, and if $\underset{t\to+\infty}{Lim}\ \sigma_t^2 = \sigma^2 \neq 0$ convergence in expectations holds. An important feature of convergence is that it involves a changing structure of the dynamic, while cointegration refers to a property over the whole sample. Nevertheless as we saw in the previous section cointegration may evolve along the time. From a statistical point of view the difference would be that cointegration implies a linear combination that is stationary, while convergence in probability requires that the linear combination evolves towards a deterministic constant. More precisely the problem can be reformulated in a state-space form:

$$y_t^1 = A_t + By_t^2 + \varepsilon_t \tag{53}$$

$$A_t = A_{t-1} + \eta_t \tag{54}$$

With $\varepsilon_t \sim N\left(0, \sigma_t^2\right)$, $\eta_t \sim N\left(0, \Sigma_t^2\right)$ and:

$$\sigma_t = \psi\sigma_{t-1} \tag{55}$$

$$\Sigma_t = \phi\Sigma_{t-1} \tag{56}$$

with $\phi > 0$.

Then convergence in expectations will hold if $0 \le \phi < 1$. Convergence in probability would hold if $0 \le \psi < 1$ but given the measurement errors in time series it is very unlikely to hold. Thus, Hall et al. (1997) limit their study to the convergence in probability and impose $\psi = 1$.

Analysing convergence within such a framework encompasses usual approaches. As the concept of β-convergence has been strongly criticised (see Friedman 1992, Quah 1993, Bernard and Durlauf 1995, 1996) most of contributions focus on σ-convergence that investigates if the standard deviation of the logarithm of productivity across economies tend to decrease over time (see Friedman 1992, Lichtenberg 1994, Carree and Klomp 1997). This can be directly related to the methodology proposed here, as a simple application would be to set $B=0$ and y^1 equal to the observed cross-section dispersion. Quah (1990, 1996) proposes a specification close to the one of Hall et al, but addresses only the question of whether some series have converged or not, while the methodology proposed by Hall et al is also able to understand if some variables are converging or not.

5 Modelling expectations and policy changes

With his critique over traditional economic modelling, Lucas (1976) initiated the so-called new classical economics that is built around optimising behaviours and rational expectations. Nevertheless, the assumption of rational expectations has come under attack for assuming too much information on the part of agents, and rarely being supported by available data (see for example Fuhrer 1997 or Roberts 2001). As an alternative approach, a learning process in building expectations may be considered. It is then assumed that agents' expectations are on average correct, but that only a limited set of information is utilised (see Sargent 1993), which is explored in section 5.1.[9]

The Lucas critique involves not only expectations but also how policy changes. More precisely the policy rules that we have to model may change over time, and agents have then to learn how those parameters have changed. Taylor (1975) and Friedman (1979) have criticised rational expectations models, as precisely they do not integrate how agents learn the policy that is implemented. This is further explored in section 5.2.

5.1 Learning process

The question of the nature of expectations is a crucial one when conducting economic policy, especially monetary policy. A particularly important aspect of forward looking behaviours is that they put at the centre of monetary policy issues like credibility, commitment, reputation and time consistency. A crucial point in inflation targeting is not so much that the central bank has to target directly inflation instead of an intermediate objective, but rather that it has to provide all the requirements for being credible in achieving its goal.[10] Therefore, inflation targeting can also be viewed as a monetary policy framework that explicitly integrates a forward looking component in inflation.

If expectations were purely rational, then inflation could be reduced without any cost providing that the central bank is fully credible (see Ball 1994, 1995). As noted by Roberts (1997), reducing inflation is usually costly, mostly because there is some inertia in the inflation process. This inertia may exist either because of wage contracts that are set for several periods (see Taylor 1979, 1980 or Fuhrer and Moore 1992, 1995), or because expectations are not perfectly

9 In this section we mostly discuss applications on inflation expectations, but there have been some similar studies on exchange rate expectations, see Canova and Ito (1985), Garratt and Hall (1995) or Basdevant et al. (2001).

10 Basically the requirements are: adoption of an inflation target as the main objective, independence, technical capability to forecast inflation, high levels of transparency.

rational. What Roberts (1997) also emphasises is that in new Keynesian models, inflation can be reduced at no cost in spite of sticky prices, as long as inflation expectations are rational. Thus it is important to understand where the stickiness in inflation comes from. If it is because inflation is inherently sticky, then reducing inflation would always imply some costs. By contrast, if inflation is sticky because of expectations, then reducing inflation could be costless in the long run (or with substantially reduced costs), providing that agents change their expectation rule to learn a rational expectations equilibrium.

In this section we present how recent contributions analyse inflation using small macro-economic models (see section 5.1.1), and also how learning processes are modelled in practice (see section 5.1.2). These discussions are of particular importance as the optimal policy may change when expectations are no longer rational.

5.1.1 Discussion of economic and policy implications

In recent years many authors have investigated very compact macro-economic models, and have discussed intensely the forward looking nature of expectations.[11] A particular feature of those models (sometimes referred to as new Keynesian models) is that they have micro-foundations (see Roberts 1995, McCallum and Nelson 1999, Woodford 1996, Rotemberg and Woodford 1997) and therefore their parameters are structural and not subject to the Lucas critique (see Lucas 1976).

Basically those models can be represented as follows:

$$\begin{cases} y_t = \alpha y_{t+1}^e + \beta\left(i_t - \pi_{t+1}^e - \bar{r}\right) + \varepsilon_t^y \\ \pi_t = \pi_{t+1}^e + \gamma y_t + \varepsilon_t^\pi \\ i_t = (1-\mu)\left(\bar{r} + \pi_t + \lambda(\pi_t - \bar{\pi}) + \theta y_t\right) + \mu i_{t-1} + \varepsilon_t^i \end{cases} \tag{57}$$

where y_t is the output gap, π_t the inflation rate, i_t the short-term nominal interest rate, $\bar{\pi}$ the inflation target, $\bar{r}$ the neutral real

11 See Lindé (2001b) for a discussion of those different models.

interest rate, y^e_{t+1} and π^e_{t+1} are the expectations of the output gap and inflation rate, and α, β, γ, λ, θ and μ are structural parameters, all but β being positive. The first relation the aggregate demand (AD), the second relation is the aggregate supply (AS) and the third summarises the behaviour of the central bank using a Taylor rule with interest smoothing (μ being the smoother parameter).

In this set-up we do not assume anything regarding the rationality of expectations. They can be rational, adaptive or derived from a learning process. As mentioned before a first step is to investigate whether inflation inertia comes from the structure of the economy itself or from expectations. Roberts (1997) proposes a simple way to address where the inertia in inflation comes from, by using survey expectations in a model of the type of (1):

$$\begin{cases} y_t = \alpha y_{t+1} + \beta\left(i_t - \pi^e_{t+1} - \bar{r}\right) + \varepsilon^y_t \\ \pi_t = \delta\pi_{t-1} + (1-\delta)\pi^e_{t+1} + \gamma^1 y_t + \gamma^2 y_{t-1} + \varepsilon^\pi_t \\ i_t = (1-\mu)\left(\bar{r} + \pi_t + \lambda(\pi_t - \bar{\pi}) + \theta y_t\right) + \mu i_{t-1} + \varepsilon^i_t \end{cases} \tag{58}$$

The two main differences are that in this model π^e refers to survey expectations (therefore expectations are defined as opposed to the general formulation given in (57)) and the coefficient δ measures the inertia in the inflation dynamic. As δ is found not significantly different from *0* he concludes that in the US the inertia in inflation comes only from expectations that are not perfectly rational.

Several contributions propose hybrid models that nest backward and forward looking dynamics, and then try to evaluate how relevant the forward looking dynamic is (see Clarida et al 1999, Roberts 2001 or Rudebusch 2002). Empirical hybrid models give contrasting results. Among others, Fuhrer (1997) or Roberts (2001) find that forward looking behaviours are unimportant, while Gali and Gertler (1999) and Gali et al (2001) find to the contrary that forward looking behaviours are dominant. As pointed out by Gali and Gertler (1999) a possible explanation of those contrasting results is the choice of the tension variable entering in the Phillips curve. If one uses the output gap then the model tends to reject a forward looking nature of inflation, while models based on marginal cost exhibit the opposite

result. Jondeau and Le Bihan (2001) have investigated further those findings, by extending the number of countries studied and by checking systematically the impact of different tension variables. Basically they show that the observed differences are much more dependent on the structure of lags and leads than the choice of the tension variable. They also accept a general hybrid model with three leads and lags, that gives roughly equal weights on backward and forward dynamics. Another argument often put forward in favour of backward looking models is that empirically the Lucas critique does not seem to be relevant as parameters do not exhibit significant instability (see Ericsson and Irons 1995, Rudebusch and Svensson 1999). Nevertheless those results have been criticised by Lindé (2001a) who points out that because of a low power on small samples the instability tests cannot correctly distinguish between changes in the policy from other shocks that affect the economy. Moreover Lindé (2001b) suggests that both forward and backward looking models exhibit instability in parameters,[12] which might be caused by a flawed measure of expectations.

What those findings reveal is that the nature of expectations formation is still under investigation, and the use of hybrid models is much more an ad hoc specification that acknowledges that expectations are neither totally rational nor totally adaptive. Going back to the theoretical foundations, the problem can be understood as follows:

- Adaptive expectations are an unsatisfactory concept mostly because they assume that agents do not react to systematic mistakes they make;

- Rational expectations have come under attack because they assume too much information on the part of agents.

As an alternative, a learning process in modelling expectations may be considered. It is then assumed that agents' expectations are on average correct, but that only a limited set of information is utilised.

---

[12] Thus even forward looking models should be tested for parameter instability (see also Estrella and Fuhrer 1999).

Hence a relatively simple representation of expectations is possible, avoiding systematic errors in a model similar to the one by Feige and Pearce (1976), Caskey (1985) or Hamilton (1985), in which agents are assumed to use a univariate model to form their expectations. While implementing a learning process it should be emphasised that agents adjust the expectation rule when they observe the errors they make, and the weights they assign to the different variables used are changing over time. An interesting feature of learning processes is that under some circumstances they show how a rational expectations equilibrium (REE) may be learnt by agents. There have been a rather abundant literature relating rational expectations and learning process (see Lucas 1986, Woodford 1990, Beeby et al. 2001 or Orphanides and Williams 2002). Most of those contributions use least square estimations to simulate the learning process. Its convergence to REE will depend on the set of prior information that agents will consider to form their expectations[13] (see Marcet and Sargent 1988, 1989a, b, Timmermann 1994 Sargent 1999 or Evans and Honkapohja 2001).

A major criticism addressed to learning is that the choice of the specification for the learning process is arbitrary.[14] Nevertheless, Garratt and Hall (1997) and Beeby at al. (2001) investigate the impact of different learning processes, to find that at least for large macro-economic models it makes little difference. Intuitively the idea is that learning processes extract information well enough so that the precise form of the learning is not crucial. Marcet and Sargent (1989a) also demonstrate that as long as the variables entering in the learning rule are correlated to the variables that explain the dynamic under REE, then the learning process will converge to REE. In a small macro-economic model the problem is of a different nature: providing that the structure is simple enough it is possible for agents to use the variables that are relevant in the REE

[13] Let A_t be the optimal estimate at date t of an unknown vector A. The learning process can be viewed as updating A_t using a simple rule of the type: $A_t=TA_{t-1}$. It will converge towards REE if the true value A is a solution of the updating equation, the initial value chosen for A is close to the true value and if the matrix T has its eigenvalues within the unit circle.

[14] As an example Woodford (1990) considers that agents could adopt a sunspot variable to form their expectations.

and then learn about the weights of each variable (see Orphanides and Williams 2002).

Beyond this criticism the most important issue is that a model with a learning process can exhibit very different properties than a model with another type of expectations rule. Beeby et al (2001) show that if expectations are not fully rational then a model based on a learning process will provide simulation properties much closer to the true model than a model based on rational expectations. Following this perspective Orphanides and Williams (2002) show in a small model of inflation, policies that are efficient under rational expectations are not when agents use a learning process. More precisely they suggest that the optimal monetary policy under a learning process should be more aggressive and also narrowed to inflation stability as it is the main objective of the central bank under an inflation targeting system. These findings are linked to the nature of expectation formation: being aggressive towards inflation and focussed on one objective facilitates the learning process.

5.1.2 Discussion of implementation in practice

Having discussed the economic and policy implications of modelling expectations with a learning process in this section we turn to the description of how expectations are derived in the empirical literature. We also discuss under which conditions a learning process may eventually converge towards rational expectations.

Basically a learning process assumes that inflation expectations can be derived from the following simple rule:

$$\pi_{t+1} = a_{1,t} + a_{2,t}\pi_{t-1} + a_{3,t}z_{t-1} + \varepsilon_t \qquad (59)$$

where π_t is the inflation rate and z_t is a vector of variables that agents use to form their expectations, ie the set of information they find relevant (besides the lagged value of inflation). Let X_t and A_t be the following vectors: $X_t=(1, \pi_{t-1}, z_{t-1})$ and $A_t=(a_{1,t}, a_{2,t}, a_{3,t})'$.

$$A_t = A_{t-1} + \frac{\alpha_t}{t} R_t^{-1} X_t' \left(\pi_{t+1} - X_t A_{t-1}\right) \qquad (60)$$

where $R_t = \frac{1}{t}\sum_{\tau=1}^{t} \alpha_\tau X'_\tau X_\tau$ and α_t a sequence of positive numbers.[15] This formula is actually a version of weighted least squares and if $\alpha_t=1$ the formula above corresponds to recursive least squares. An interesting feature is that this method of updating parameters can be cast into the Kalman filter formulae. Defining $P_t = \frac{1}{t} R_t^{-1}$ and $f_t = X_t P_{t-1} X'_t + \frac{1}{\alpha_t}$ it becomes (see Bullard 1992):

$$A_t = A_{t-1} + P_{t-1} X'_t f_t^{-1} \left(\pi_{t+1} - X_t A_{t-1} \right) \tag{61}$$

$$P_t = P_{t-1} - P_{t-1} X'_t X_t P_{t-1} f_t^{-1} \tag{62}$$

Which corresponds to the following state-space model:

$$\pi_{t+1} = a_{1,t} + a_{2,t}\pi_{t-1} + a_{3,t} z_{t-1} + \varepsilon_t \tag{63}$$

$$\forall i \quad a_{i,t} = a_{i,t-1} + \eta_{i,t} \tag{64}$$

with the hyper-parameters given by:

$$Var(\varepsilon_t) = \frac{1}{\alpha_t} \tag{65}$$

$$Var(\eta_t) = 0 \tag{66}$$

Thus the problem with least square estimations is that the learning process is not optimal in the sense that they assume that coefficients are stable while they estimate time-varying ones. Put in other words, the results of Marcet and Sargent (1989a,b) on the convergence of learning process towards rational expectations hold only when the law of motions of parameters is viewed as invariant. Ljung and Söderstöm (1983) and Bullard (1992) show that when this assumption is relaxed, the convergence property no longer holds. Intuitively the reason is straightforward: if $Var(\varepsilon_t)\neq 0$ then P_t does not

15 R_t can also be derived according to the following formula: $R_t = R_{t-1} + \frac{\alpha_t}{t}\left(X'_t X_t - R_{t-1}\right)$, which is used by Orphanides and Williams (2002) and Honkapohja and Mitra (2002) for most recent contributions.

converge towards 0, and thus learning does not converge to rational expectations.

More precisely, in a more general state-space form the coefficients would be derived as follows:

$$A_t = A_{t-1} + P_{t-1} X'_t f_t^{-1} \left(\pi_{t+1} - X_t A_{t-1} \right) \tag{67}$$

$$P_t = P_{t-1} + Q_t - P_{t-1} X'_t X_t P_{t-1} f_t^{-1} \tag{68}$$

with $Var(\varepsilon_t)=H_t$, $Var(\eta_t)=Q_t$ and $f_t = X_t P_{t-1} X'_t + H_t$.

The set of equations (67)and (68) should be viewed as follows. First agents form their expectations for the value of inflation for the next period, before they observe its current value. Once it is known they use this information to revise their belief, in order to avoid systematic mistakes. Expectations are thus computed as the predicted estimate for π_{t+1}:

$$\hat{\pi}_{t+1} = a_{1,t|t-1} + a_{2,t|t-1}\pi_{t-1} + a_{3,t|t-1} z_{t-1} \tag{69}$$

Sargent (1999) Evans and Honkapohja (1999, 2001) or Orphanides and Williams (2002) propose a simpler version of permanent learning, using the algorithm reproducing weighted least squares given in (60). Their methodology consists of setting a geometrical pattern for the weights: $\alpha_t=\kappa t$, where κ is set to an arbitrary small value. The advantage of using the Kalman filter in empirical contributions is that it will give the optimal gain that agents apply when updating their parameters, and can also allow to test if the variance of the state variables is significantly different from zero or not, ie to test if the learning is perpetual or if it converges towards rational expectations.

5.2 Modelling policy change

Modelling expectations using a learning process is nevertheless insufficient to address completely the Lucas critique, as policy rules are also subjected to revisions (see McNelis and Neftci 1982, Mitchell 1982 or Plantier and Scrimgeour 2002).

A possible formulation is as follows:

$$y_t = A_{1,t} + A_{2,t}\left(X_{t-1} - X_{t-1}^{*}\right) + \varepsilon_t \tag{70}$$

$$\forall i \in \{1,2\} \quad A_{i,t} = A_{i,t-1} + \eta_{i,t} \tag{71}$$

with $\varepsilon_t \sim N\left(0,\sigma_t^2\right)$ and $\eta_{i,t} \sim N\left(0,\xi_i^2\right)$. y is the instrument set by monetary authorities, x is the variable of interest, and x^* is the targeted value of this variable. An important feature pointed by Kim (1993) is that there is no reason to consider that the variance on the measurement equation is constant. More precisely what agents have to do to understand the rule that is implemented is to extract from the observed changes in x_t what is due to policy and what is due to shocks. The case considered by Kim (1993) is the one where σ_t^2 follows a Markov-switching model, and apply this to the Fed policy in terms of money growth targeting from early sixties to late eighties. Kim finds evidence that uncertainty coming from time-varying parameters has a significant effect on economic activity, while uncertainty due to heteroscedastic disturbances has little impact. This result is consistent with the literature on irreversible investments:[16] when the uncertainty on policy parameters increases there is more incentive to delay irreversible decisions.

Another crucial issue is the credibility and time-consistency of monetary authorities. Again, modelling a time-varying policy rule is an appropriate method. Hardouvelis and Barnhart (1989) define credibility as the response of commodity prices to weekly announcements of the Fed on M1. This is done within a standard state-space model:

$$P_t = a + A_t M_t \varepsilon_t \tag{72}$$

$$A_t = A_{t-1} + \eta_t \tag{73}$$

[16] See Bernanke (1983), Cukierman (1980), Henry (1974) or Pindyck (1988).

Where P is the commodity price change and M is the unanticipated percentage change in M1 (measured as the difference between the observed growth rate of money and the forecasted one). If the Fed policy is credible then the parameter A should converge to a negative value: when M1 grows faster then its expected value agents anticipate that monetary authorities are going to raise interest rates which induces a decrease in inflation expectations and therefore a decrease in current inflation. What they show is an improvement of the credibility during the eighties.

6 Forecasting

Forecasting properties of models can also be improved by using state-space modelling, when errors exhibit some patterns (see section 5.1) or when data are subjected to revisions, which can be of a particular importance (see section 5.2).

6.1 Consistent intercept adjustments

Forecasting under structural change is a particularly difficult exercise. When the crucial assumption of stability in the data generating process is relaxed, then standard econometrics cannot account for an evolving structure. As pointed by Clements and Hendry (1996, 2001) Hall (1993) or Greenslade and Hall (1996) econometric modelling can still be useful if it accounts explicitly for the structural change that occurred. In this section we briefly presents the contribution of Clements and Hendry (1996) who demonstrate how consistent correction in intercepts may capture a wide range of structural change, and then an application of this paper to a macro-economic model developed for the Russian economy (see Basdevant 2000).

Clements and Hendry (1998, 1999) decompose the sources of forecasting errors between the following categories:

- Shifts in the coefficients of deterministic components (either the intercept or the trend);
- Shifts in the coefficients of stochastic components;
- Mis-specifications;

- Mis-estimations;
- Mis-measurement of data;
- Changes in the variance of errors;
- Errors cumulating over the forecast horizon.

An important contribution of their work is to show that the first source of errors is the most damaging error, and can be the source of systematic forecast failures. Hendry and Clements (1994a, b), Clements and Hendry (1996) show that VAR models in differences will usually dominate standard VECM models in terms of forecasting errors, in the presence of structural change in the deterministic components of long-run relations. It is worth noting that the term ECM has been renamed by Hendry and Clements as 'Equilibrium Correction Model' to emphasise that systematic forecasting errors usually come from the fact that the ECM keeps adjusting to a long-run relation while it has shifted because of structural change. Hence, the VAR model will tend to dominate the VECM simply because it does not force variables to come back to a relation that may no longer be accurate. As a result, intercept corrections can capture a wide range of structural changes in the long-run relation. More precisely, adding the residuals of the current period to the next period's forecast will improve the forecast performance in terms of mean forecasting square errors when the autocorrelation of forecasting errors is sufficiently high.

In the case of the Russian model developed by Basdevant (2000), the challenge was to provide a model that could reflect the specificities of the Russian economy and also provide sensible forecasts. The major problem was that the Russian economy was still going through a difficult transition towards a market economy. The model was developed just in the aftermath of the 1998 crisis, where the future path of the economy was likely to look substantially different from its pattern since the beginning of the transition, which had mostly been characterised by a decline in GDP. It was reasonable to believe that the Russian economy could recover, and a growth recovery would eventually take place.

The strategy adopted was to reconcile a) that the model should reflect that the economy would eventually behave as basic theory predicts with b) the need to have a model that would provide

sensible forecasts. A lot of emphasis was therefore made on the long-run relations of the ECM: in the long-term the Russian economy should conform to economic theory, hence the calibration of many long-term parameters. Structural change was then captured by a change in the deterministic components of the model, which was eventually implemented in a state-space model.

Hence the standard ECM format was changed in order to specify the rule for a time-varying intercept:

$$\Delta y_t = A(L)\Delta y_{t-1} + B(L)\Delta x_t - \lambda\left(y_{t-1} - bx_{t-1} + c_t\right) + \varepsilon_t \tag{74}$$

$$c_t = c_{t-1} + \eta_t \tag{75}$$

where ε_t and η_t are white noises. The adjustments in c_t allowed the model to fit to data despite calibrations, and therefore provided a reasonable forecasting model. As an example, economic theory predicts that consumption is related to disposable income. Nevertheless, disposable incomes declined sharply over the transition period, while consumption remained constant. The stochastic intercept then suggests that transition has led to an increasing consumption to income ratio, which has been almost exactly offset in terms of actual consumption by the fall in incomes.

It is worth noting that specifying the state equation as a random walk is a common method adopted by practitioners (see Belongia et al 1988, Hafer and Sheehan 1990). Nevertheless using such a method may create some bias when the variance of η_t is assumed to be constant. The problem that may arise is if changes are abrupt and of a rather large magnitude then the variance over the sample will tend to be over-estimated during the periods of stability, which can lead to exaggerate the variability of the time-varying parameters (see Gamble and Le Sage 1993). This can be accommodated by allowing a time-varying structure for the hyper-parameters (see Gordon and Smith 1988, 1990). The method adopted by Gordon and Smith is to adopt switching models, but it could be also done with a standard ARCH process that can be estimated with a Kalman filter.

Another possible extension is to deepen the analysis of the time-varying parameters. In the above analyses the parameters have always been assumed to follow a random walk process, this has the disadvantage that the forecast from the model will always be that structural change has ceased. If we have an idea as to the general speed or form of parameter change we can build this into the state equations in the following way.

Let the measurement equation be:

$$Y_t = A_t X_t^1 + \varepsilon_t \tag{76}$$

and the state equations be:

$$A_t = A_{t-1} + \phi X_t^2 + \eta_t \tag{77}$$

Here the variable X^2 will have a cumulative effect on the state variable and as long as it is forecasted to be non zero the parameter will be forecasted to continue to change.

6.2 Dealing with real time data

The period of high inflation and low growth during the seventies has lead to various interpretations regarding monetary policy. A general view on monetary policy is that it should keep inflation low while maintaining economic stability. This has been formalised by the Taylor rule (see Taylor 1993). A fundamental issue in the failure of the Fed (as well as many other central banks in the world) to control inflation during the seventies is to know whether this was due to an inappropriate policy or a misguided policy. The former would suggest that if monetary authorities aim at over controlling the business cycle, then they might worsen the situation and thus the period of 'stagflation' could be easily solved in principle by implementing a different monetary policy more focused on controlling inflation. The latter has much more implications in the design of monetary policy: basically it would suggest that even an appropriate monetary policy, like an inflation targeting one, could lead to undesirable results if the tools used by policymakers give inappropriate guidance. This is related to two major points: the Lucas critique, that involves the structure of the models used, and data used, that involves measurement problems. In this section we

discuss problems related to data measurement. This is of particular importance, as Orphanides et al (1999), Orphanides and Van Norden (1999) and Orphanides (2000a, b, 2001, 2002) show that the period of stagflation can to a large extent be explained by mis-measurement problems in the output gap, which in turn led the Fed to undertake an over-activist policy in terms of stabilising the economic activity. The problem those authors point out is that policy-makers have to use real-time data, which are subjected to revision and therefore also to mis-measurement.

In these circumstances, which are quite normal, two different problems can occur. Firstly, estimates based on real-time data can substantially differ from those based on revised data. Secondly, using the HP filter to derive potential output, which is often the case among practitioners, has the problem of end-of-sample estimates. Furthermore, alternative methods of computing potential output will usually lead to different estimates. Thus there is a need to robustify the methods used.

Those two problems can partially be addressed by using state-space modelling. The next two sub-sections describe how it is possible to address - at least partly - those issues.

6.2.1 Addressing the sensitivity to detrending methods

A lot of empirical contributions estimate potential output using rather simple techniques that require few data and can be easily implemented. Among these are the split time-trend method (see Giorno et al (1995)), the univariate Hodrick-Prescott filter (see Hodrick and Prescott (1997), Taylor (1999), Clarida et al (2000)) or the quadratic trend method (Taylor (1999), Clarida et al (2000)). Those approaches basically use historical data on output to filter a trend, which is associated with the potential output. The major drawback is that potential output is derived only from actual output, and therefore it makes any forecasting exercise particularly difficult if data are not stable enough to allow the output gap to be derived from past data. In the split time-trend method, forecasting potential output requires ad hoc assumptions on the relative current position in the business cycle. With the Hodrick-Prescott filter the problem comes from the fact that it imposes the properties of the estimated cycle by construction, while the problem is precisely to estimate this cycle (see Harvey and Jaeger (1993), Casanova (1998) or Orphanides and Van Norden (1999)).

Thus, many authors have moved towards a more structural approach, in which potential output estimates will be more usable for forecasting and policy simulations. To do so, some have based the estimation on a production function (see Giorno et al (1995), De Masi (1997) Room (2001) or Dimitz (2001)). The problem is that potential output will usually depend on the level of employment compatible with the NAIRU, which is itself computed using a Hodrick-Prescott filter or any other simple filtering method. Thus, those approaches only partially address the drawbacks of the initial methods, because the problem will mostly be shifted from potential output to NAIRU.

There are also various multivariate approaches that have been developed. As an example the VAR methodology can be used with a structural VAR model (SVAR model) à la Blanchard Quah (1989).[17] Those approaches are rather fruitful as they integrate economic relations, and therefore are less mechanical that the standard HP filter. Nevertheless they do not necessarily provide a more accurate estimate as they will also be subjected to specification and identification problems. A state-space model is usually set-up as a reduced form compatible with various underlying structural forms, thus it may reveal more robust to mis-specifications than estimations based on a chosen structural form. Furthermore, within a state-space form it is possible to combine the two approaches mentioned, and to specify how far we want to rely on other sources of information when estimating the potential output. As an example, Apel and Jansson (1999), Rasi and Viikari (1998), Clark (1989), Gerlach and Smets (1999) or Scott (2000) use models where output and/or employment are decomposed into trend and cycle, which are treated as unobserved variables, and more generally the use of the HPMV filter can be of a particular interest as it allows the user to determine what are the weights set on other sources of information. It has been

[17] See Apel and Jansson (1999) Gerlach and Smets (1999), Camba-Mendez and Rodriguez-Palenzuela (2001).

used by the Central Banks of Canada and New Zealand to estimate potential output (see Butler (1996), Conway and Hunt (1997), Kichian (1999) or Scott (2000)) and by OECD (1999) to estimate the NAIRU.

In a recent study, Richardson et al (2000) discuss how to model the NAIRU in practice. What those authors point is the difficulty in estimating correctly the long-run value of the NAIRU, especially because the Phillips curve may be subjected to structural change. As an example Blanchard and Wolfers (1999) emphasised that labour market institutions may change over time which could explain a time-varying NAIRU. More precisely if there are indications that the NAIRU might have been time-varying then it is rather natural to model it within a state-space form. Within a state-space form it is also possible to estimate simultaneously the NAIRU and the Phillips curve (see also Gordon 1997, 1998). Basically the idea is to set-up a HPMV filter as follows:

$$\underset{\{y_t^*\}}{Min} \quad \sum_{t=1}^{T}\left(u_t - u_t^*\right)^2 + \lambda_1\left(\Delta^2 u_t^*\right)^2 + \lambda_2 \zeta_t^2 \tag{78}$$

With λ_1 and λ_2 given. The residuals ζ_t taken from the Phillips curve:

$$\pi_t = A(L)\pi_{t-1} + B(L)u_t + \beta\left(u_t - u_t^*\right) + \zeta_t \tag{79}$$

Regarding the estimation of the output gap, an interesting approach to integrate more structural relations is the one adopted by the Reserve Bank of New Zealand for its forecasting model (see Conway and Hunt 1997)

$$\underset{\{y_t^*\}}{Min} \quad \sum_{t=1}^{T}\left(y_t - y_t^*\right)^2 + \lambda_1\left(\Delta^2 y_t^*\right)^2 + \lambda_2 \zeta_{1,t}^2 + \lambda_3 \zeta_{2,t}^2 + \lambda_4 \zeta_{3,t}^2 \tag{80}$$

With λ_1 to λ_4 given. The residuals $\zeta_{i,t}$ correspond to three relations: a Phillips curve, an Okun's law and a link between the utilisation rate of capacities survey data and the potential output:

$$\pi_t = \pi_t^e + A(L)\left(y_t - y_t^*\right) + \zeta_{1,t} \tag{81}$$

$$u_{t-4} - \bar{u}_{t-4} = \beta\left(y_t - y_t^*\right) + \zeta_{2,t} \tag{82}$$

$$cu_t - c\bar{u}_t = \gamma\left(y_t - y_t^*\right) + \zeta_{3,t} \tag{83}$$

where π is the inflation rate, π^e is the expected inflation rate (taken from survey data), y is the output and y^* the potential output, u is the unemployment rate and cu the capacity utilisation rate that are taken as differences from their equilibrium value.

6.2.2 Addressing the issue of data revision

Patterson (1995) proposes a model to investigate the presence of some patterns in data revisions,[18] which can be used to capture partly of the Orphanides' critique. To do so the problem of data revision is represented in a state-space form, following the works of Conrad and Corrado (1979), Howrey (1978, 1984), Harvey et al (1983) and Bordignon and Trivellato (1989).

The starting point of the study is to analyse how to forecast accurately a vector of i variables that are published at a given frequency. Those data are revised over time. Basically, in period t there are a set of publications made: the first vintage for the current period, the second vintage for the previous period and so on, up to the final vintage m. It is then important to distinguish two different objects, the data generating process (DGP) and the data measurement process (DMP):

- The DGP involves some economic links between all the final vintages, and corresponds to what econometricians are trying to capture.

[18] Patterson applies his methodology to UK aggregated data on consumption and income at a quarterly frequency, from 1970Q1 to 1992Q1.

- The DMP involves the link between the different intermediate vintages and the final ones, as well as measurement errors between each vintage and the final one.

Let denote y_{t-q}^k the k-th vintage of the vector y for the period $t-q$ (with $0 \le k \le q \le m$). This variable is available at period $t-q+k$. Then for each vintage k a relation with the final vintage m can be estimated:

$$y_t^k = c^k + d^k y_t^m + \eta_t^k \tag{84}$$

Then two problems arise, one is that the vintage k is usually a biased estimator of the final vintage for low values of k and the other one is that errors are correlated. Let us discuss those two points further.

Firstly, at date t the final vintages of all the variables between t and $t-m+1$ are unknown. If $c^k=0$ and $d^k=1$ then the revision $r^k \equiv y^k - y^m$ and the error term are identical. This is particularly important, as the k-th vintage would be an unbiased predictor of the final vintage. The first finding of Patterson is precisely to show that for earlier vintages $c^k \neq 0$ and $d^k \neq 1$. Quite naturally those coefficients get much closer to 0 and 1 respectively when k gets closer to m.

Secondly, errors can exhibit some pattern, which can be modelled and then used to obtain a better estimate of the final vintage. Harvey et al (1983) and Howrey (1984) suggest an AR(1) process, but Patterson investigates two extensions: one by extending the order of the AR process and the other by allowing interrelated errors (either between different variables and different vintages of the same variable). More precisely, defining η_t as the vector of errors of all the vintages: $\eta_t = \left(\eta_t^0 \cdots \eta_t^m\right)'$, then the model proposed is:

$$\eta_t = \Psi(L)\eta_{t-1} + u_t \tag{85}$$

In this very general set-up it is possible to investigate various structures, from a relatively simple one to more complicated ones. What the author shows is that there is a significant pattern in the errors terms (whether the estimation is made on a single-equation or on the whole system).

The DGP can be represented by an auto-regressive distributed lags (ADL) process between all the final estimates:

$$y_t^m = b + B(L)y_t^m + \varepsilon_t \tag{86}$$

the set of equations(84) and (85) can then be viewed as state equations, while equation (86) is the measurement equation: for each period t there are a set of variables for which the final vintage is an unknown variable, which can be estimated by a Kalman filter using the measurement equation, that links the latest vintage available to the final one, the first set or state equations that model the residuals and the final state equation that model the pattern of the state variable given an economic model. As emphasised by the author

> "integrating models of the DGP and DMP in the state space approach keeps a clear distinction between the consistent vintage of data need to build the model and the other vintages of data which are available and are used to model the DMP. If a prediction of a preliminary, rather than the final, vintage is required this is easily obtained through a combination of the transition and measurement equations of the state space model."

7 Conclusion

In this paper we have described various applications of state-space models in macroeconomics. They have proven to be of particular use in estimating various types of time-varying parameters: expectations, causal structures, long-run relations as well as unobserved variables. Their applications are of course much wider than macroeconomics, but what is also particularly important is that they provide not only an simple and flexible analytical tool, but also a tool that can be used by practitioners, as most of the applications discussed here have direct policy implications. Given the flexibility of this modelling methodology, its application should become increasingly popular in the future.

References

Anderson, B D O and J B Moore, (1979), Optimal Filtering, Englewood Cliffs: Prentice-Hall.

Apel, M and P Jansson (1999), "A theory-consistent system approach for estimating potential output and the NAIRU," *Economics Letters* 64(3): 271-275.

Ball, L (1994), "Credible disinflation with staggered price setting," *The American Economic Review,* 84(1): 282-289.

Ball, L (1995), "Disinflation with imperfect credibility," *Journal of Monetary Economics,* 35: 5-23.

Barassi, M R, Caporale G M and S G Hall (2000) "Interest linkages: a Kalman filter approach to detecting structural change," *University of Oxford, Centre for International Macroeconomics Discussion Paper* 2000.09.

Barassi, M R, G M Caporale and S G Hall (2001), "Testing changes in the long-run causal structure of cointegrated vector autoregressions," *South Bank University, Centre for Monetary and Financial Economics Discussion Paper* #06-2001.

Basdevant, O (2000), "An econometric model of the Russian Federation," *Economic Modelling,* 17(2): 305-336.

Basdevant, O, S Brink and R Koekemoer (2001), "Exchange rates and learning – a Rand/US dollar model," *Journal of Propagation in Probability and Statistics,* 2(1), 77-96.

Beeby, M, S G Hall and S B Henry (2001), "Rational expectations and near rational alternatives: how best to form expectations," *European Central Bank Working Paper* #86.

Belongia, M T, R W Hafer and R G Seehan (1988), "On the temporal stability of the interest rate-weekly money relationship," *The Review of Economic and Statistics,* 70(3): 516-520.

Bernanke, B S (1983), "Irreversibility, uncertainty and cyclical investment," *The Quarterly Journal of Economics,* 85-106.

Bernard, A B and S N Durlauf (1995), "Convergence in international output," *Journal of Applied Econometrics,* 10: 97-108.

Bernard, A B and S N Durlauf (1996) "Interpreting tests of the convergence hypothesis," *Journal of Econometrics,* 71: 161-173.

Beveridge, S and C R Nelson "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'", *Journal of Monetary Economics,* 7(2): 151-74.

Blanchard, O and D Quah (1989), "The dynamic effects of aggregate demand and aggregate supply disturbances," *The American Economic Review* 79: 655-673.

Blanchard, O and J Wolfers (1999), "The role of shocks and institutions in the rise of European unemployment: the aggregate evidence," *NBER Working Paper* #7282.

Boone, L (2000), "Comparing semi-structural methods to estimate unobserved variables: the HPMV and Kalman filter approaches," *OECD Working Paper* #240.

Bordignon, S and U Trivellato (1989), "The optimal use of provisional data in forecasting with dynamic models," *Journal of Business and Economics Statistics,* 7: 275-286.

Bullard, J (1992), "Time-varying parameters and non-convergence to rational expectations under least square learning," *Economic Letters* 40: 159-166.

Butler, L (1996), "A semi-structural method to estimate potential output: combining economic theory with a time-series filter," *Bank of Canada Quarterly Projection Model.*

Canova, F and T Ito (1991), "The time-series properties of the risk premium in the yen/dollar exchange market," *Journal of Applied Econometrics,* 6: 125-142.

Carree, M and L Klomp (1997), "Testing the convergence hypothesis: a comment," *Review of Economics and Statistics,* 79: 683-86.

Casanova, F (1998), "Detrending and business cycle facts," *Journal of Monetary Economics,* 41: 475-512.

Caskey, J (1985), "Modeling the formation of price expectations: a Bayesian approach," *The American Economic Review,* 75(4): 768-76.

Clarida, R, J Gali and M Gertler (1999), "The science of monetary policy: a new Keynesian perspective," *Journal of Economic Literature,* 37(4): 1661-1707.

Clarida, R, J Gali and M Gertler (2000), "Monetary policy rules and macroeconomic stability: evidence and some theory," *Quarterly Journal of Economics,* 147-180.

Clark, P (1989), "Trend reversion in real output and unemployment," *Journal of Econometrics,* 40: 15-32.

Clements, M and D Hendry (1996), "Intercept corrections and structural change," *Journal of Applied Economics,* 11: 475-94.

Clements, M and D Hendry (1998), Forecasting Economic Time Series, Cambridge University Press, Cambridge.

Clements, M and D Hendry (1999), Forecasting Non-Stationary Time Series, Cambridge University Press, Cambridge.

Clements, M and D Hendry (2001), "Economic forecasting: some lessons from recent research," *ECB Working Paper* # 82.

Conrad, W and C Corrado (1979), "Application of the Kalman filter to revision in monthly sales estimates," *Journal of Economic Dynamics and Control,* 1: 177-98.

Conway, P and B Hunt (1997), "Estimating potential output: a semi-structural approach," *Reserve Bank of New Zealand Working Paper* D97/9.

Crowder, M J (1976), "Maximum likelihood estimation for dependent observations," *Journal of the Royal Statistical Society*, Series B 38(1): 45-53.

Cukierman, A (1980), "The effects of uncertainty on investment under risk neutrality with endogenous information," *Journal of Political Economy,* 88(3): 462-75.

Cuthbertson, K, S G Hall and M P Taylor (1992), Applied Econometric Techniques, University of Michigan Press, Ann Arbor.

De Masi, P R (1997), "IMF estimates of potential output: theory and practise," *IMF Working Papers* #97-177.

Dimitz, M A (2001), "Output gaps and technological progress in European Monetary Union" *Bank of Finland Discussion Paper,* 20-2001.

Ducan, D B, and S D Horn (1972), "Linear dynamic regression from the viewpoint of regression analysis," *Journal of the American Statistical Association,* 67: 815-21.

Enders, W (1995), Applied Econometric Time Series, John Wiley and Sons, New York.

Ericsson, N R and J S Irons (1995), "The Lucas critique in practice: theory without measurement" in Hoover K D (ed.) Macroeconometrics: Developments, Tensions and Prospects, Kluwer Academic Publishers, Boston.

Estrella, A and J C Fuhrer (1999), "Are 'deep' parameters stable? The Lucas critique as an empirical hypothesis" *Federal Reserve Bank of Boston Working Paper* 99-4.

Evans, G and S Honkapohja (2001), Learning and expectations in macroeconomics, Princeton University Press, Princeton.

Feige, E L and D Pearce (1976), "Economically rational expectations: are innovations in the rate of inflation independent of innovations in measures of monetary and fiscal policy?" *Journal of Political Economy,* 84: 499-522.

Friedman, B M (1979), "Optimal expectations and the extreme information assumptions of rational expectations macromodels," *Journal of Monetary Economics,* 5: 23-41.

Friedman, M (1992), "Do old fallacies ever die?" *Journal of Economic Literature* 30: 2129-132.

Fuhrer, J C (1997), "The (un)importance of forward looking behaviour in price specifications," *Journal of Money, Credit and Banking,* 29(3): 338-350.

Fuhrer, J C and G R Moore (1992), "Monetary policy rules and the indicator properties of asset prices," *Journal of Monetary Economics,* 29(2): 303-36.

Fuhrer, J C and G R Moore (1995), "Inflation persistence," *Quarterly Journal of Economics* 110(1): 127-159.

Gali, J and M Gertler (1999), "Inflation dynamics: a structural econometric analysis," *Journal of Monetary Economics* 44(2): 195-222.

Gali, J, Gertler M and D Lopez-Salido (2001), "European inflation dynamics," *European Economic Review,* 45(7): 1237-1270.

Gamble J A and J P Le Sage (1993) "A monte-carlo comparison of the time varying parameter and muliprocess mixture models in the presence of structural shift and outliers," *The Review of Economics and Statistics,* 75(3): 515-519.

Garratt, A and S G Hall (1995), "Model of consistent learning: the Sterling Deutschmark rate in the London Business School model," *Economic Modelling* 12: 87-96.

Garratt, A and S G Hall (1996), "Measuring underlying economic activity," *Journal of Applied Econometrics* 11: 135-151.

Garratt, A and S G Hall (1997), "E-equilibria and adaptive expectations: output and inflation in the LBS model," *Journal of Economic Dynamics and Control,* 21(7): 1149-1171.

Gerlach, S and F Smets (1999), "Output gaps and monetary policy in the EMU area," *European Economic Review,* 43: 801-812.

Giorno, C , P Richardson, D Roseveare and P Van Den Noord (1995), "Estimating potential output, output gaps and structural budget balances," *OECD Economics Department Working Papers* #152.

Gordon, K and A F M Smith (1988), "Modelling and Monitoring Discontinuous Changes in Time Series", in Bayesian Analysis of Time Series and Dynamic Models (ed. J C Spall), Marcel Dekker, New York.

Gordon, K and A F M Smith (1990), "Modelling and monitoring biomedical time series," *Journal of American Statistical Association* 85: 328-337.

Gordon, R J (1997), "The time-varying NAIRU and its implications for economic policy," *Journal of Economic Perspectives,* 11(1): 11-32.

Gordon, R J (1998), "Foundations of the Goldilocks economy supply shocks and the time-varying NAIRU," *Brookings Papers on Economic Activity,* 2: 297-346.

Greenslade, J V and S G Hall (1996), "Modelling economies subject to structural change: The case of Germany,", *Economic Modelling,* 13: 545-559.

Guay, A and P St Amant (1996), "Do mechanical filters provide a good approximation of business cycle?" *Bank of Canada Technical Report,* 78.

Hafer, R W and R G Sheehan (1990), "The response of interest rates to unexpected weekly money: are policy changes important?" *Southern Economic Journal,* 56: 577-593.

Haldane, A G and S G Hall (1991), "Sterling's relationship with the dollar and the deutschemark," *The Economic Journal* 101: 436-443.

Hall, S G (1993), "Modelling structural change using the Kalman filter," *Economics of Planning* 26, 1-13.

Hall, S G and O Basdevant (2002), " Measuring the capital stock in Russia: An unobserved component model," forthcoming *Economics of Planning*.

Hall, S G, D Robertson and M R Wickens (1997), "Measuring economic convergence," *International Journal of Finance and Economics,* 2: 131-143.

Hamilton, J D (1985), "Uncovering financial market expectations of inflation," *Journal of Political Economy,* 93(6): 1224-1241.

Hansen, B E and S Johansen (1999), "Some tests for parameter constancy in cointegrated VAR models," *Econometrics Journal,* 2: 306-333.

Hardouvelis, G A and S W Barnhart (1989), "The evolution of Federal Reserve credibility: 1978-1984" *The Review of Economics and Statistics,* 385-393.

Harvey, A C (1981) Time Series Models, Deddington, Oxford: Philip Allan; Atlantic Highlands, NJ: Humanities Press.

Harvey, A C (1985), "Trends and cycles in macroeconomic time series" *Journal of Business and Economic Statistics* 3: 216-27.

Harvey, A C (1987) "Applications of the Kalman Filter in Econometrics", in Advances in econometrics: Fifth world congress, (ed T F Bewley), vol 1, Econometric Society Monograph No 13, Cambridge University Press, Cambridge.

Harvey, A C (1989), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge.

Harvey, A C (2001), "Testing in unobserved components models," *Journal of Forecasting* 20: 1-19.

Harvey, A C and A Jaeger (1993), "Detrending, stylised facts and the business cycle," *Journal of Applied Econometrics,* 8: 231-247.

Harvey, A C, C R McKenzie, D P Blake and M J Desai (1983), "Irregular data revisions" in Applied Time Series Analysis of Economics (ed A Zellner), US Bureau of the Census, Washington DC.

Hendry, D and M Clements (1994a), "Can econometrics improve economic forecasting?" *Swiss Journal of Economics and Statistics,* 130: 267-298.

Hendry, D and M Clements (1994b), "On a theory of intercept corrections in macro-economic forecasting" Money Inflation and Unemployment: Essays in Honour of James Ball (ed S Holly), Edward Elgar, Aldershott.

Henry, C (1974), "Investment decision under uncertainty: the irreversibility effect," *The American Economic Review* 64(6): 1006-1012.

Hodrick, R J and E C Prescott (1997), "Post-War US business cycles: an empirical investigation," *Journal of Monetary Credit and Banking,* 29(1): 1-16.

Howrey, E P (1978), "The use of preliminary data in econometric forecasting," *Review of Economics and Statistics,* 60: 193-200.

Howrey, E P (1984), "Data revision, reconstruction and prediction: an application to inventory investment," *Review of Economics and Statistics* 66: 386-393.

Jazwinski, A H (1970), Stochastic Processes and Filtering Theory, Academic Press: New York.

Johansen, S (1988), "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control,* 12: 231-254.

Johansen, S (1991), "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica,* 59: 1551-1580.

Jondeau, E and H Le Bihan (2001), "Testing for a forward looking Phillips curve. Additional evidence from European and US data," *Banque de France Working Paper* #86.

Kalman, R E (1960), "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering, Transaction of the SAME,* 59: 1551-1580.

Kalman, R E (1963), "New methods in Wiener filtering theory," in Proceedings of the First Symposium of Engineering Applications of Random Function Theory and Probability (eds Bogdanoff J L and F Kozin), John Wiley and Sons Inc: New York.

Kim, C J (1993), "Sources of monetary growth uncertainty and economic activity: the time-varying parameter model with heteroskedastic disturbances," *The Review of Economics and Statistics,* 483-492.

Laxton, D and R Tetlow (1992), "A simple multivariate filter for the measurement of potential output," *Bank of Canada Technical Report,* #59.

Lichtenberg, F R (1994), "Testing the convergence hypothesis," *Review of Economics and Statistics,* 76: 576-79.

Lindé, J (2001a), "Testing for the Lucas critique: a quantitative investigation," *The American Economic Review* 91(4): 986-1005.

Lindé, J (2001b), "The empirical relevance of simple forward- and backward- looking models: a view from a dynamic general equilibrium," *Sveriges Riksbank Working Paper Series* #130.

Ljung, L and T Söderstöm (1983), The Theory and Practice of Recursive Identification, MIT Press: Boston.

Lucas, R (1976), "Econometric policy evaluation: a critique" Carnegie-Rochester Conference Series on Public Policy 1: 19-46.

Lucas, R (1986), "Adaptive behaviour and economic theory," *Journal of Business,* 59: 401-426.

Manning, N (2002), "Common trends and convergence? South East Asian equity markets, 1988-1999" *Journal of International Money and Finance,* 21: 183-202.

Marcet, A and T Sargent (1988), "The fate of systems with 'adaptive expectations'," *The American Economic Review,* 78: 168-172.

Marcet, A and T Sargent (1989a), "Convergence of least squares learning mechanism in self-referential linear stochastic model," *Journal of Economic Theory,* 48: 337-368.

Marcet, A and T Sargent (1989b), "Convergence of least-squares learning in environments with hidden satate variables and private information," *Journal of Political Economy,* 97(6): 1306-1322.

McCallum, B T and E Nelson (1999), "Nominal income targeting in an open economy optimizing model," *Journal of Monetary Economics,* 43(2): 553-578.

McNelis, P D and S N Neftci (1982), "Policy-dependent parameters in the presence of optimal learning: an application of Kalman filtering to the Fair and Sargent supply-side equations," *The Review of Economics and Statistics,* 64(2): 296-306.

Mitchell, D W (1982), "Kalman filters and the target values of monetary aggregates," Journal of Money, Credit and Banking 14(3): 365-374

Morley, J C (2002), "A state-space approach to calculating the Beveridge-Nelson decomposition," *Economic Letters,* 75: 123-127.

Morley, J C, C R Nelson and E Zivot (2002), "Why are Beveridge-Nelson and unobserved component decompositions of GDP so different?" Forthcoming *Review of Economics and Statistics*.

Nyblom, J (1989), "Testing for the constancy of parameters over time," *Journal of the American Statistical Association,* 84: 223-230.

OECD (1999), "Making work pay," *OECD Economic Studies* #31.

Orphanides A (2000a), "Activist stabilization policy and inflation: the Taylor rule in the 1970s," *Federal Reserve Board, Finance and Economics Discussion Series* #2000-13.

Orphanides, A (2000b), "The quest for prosperity without inflation," *ECB Working Paper Series,* 15.

Orphanides, A (2001), "Monetary policy rules, macroeconomic stability and inflation: a view from the trenches," *Federal Reserve Board, Finance and Economics Discussion Series* #2001-62.

Orphanides, A (2002), "Monetary-policy rules and the great inflation," *The American Economic Review,* 92(2): 115-120.

Orphanides, A , R D Porter, D Reifschneider, R Tetlow and F Finan (1999), "Errors in the measurement of the output gap and the design of monetary policy," Federal Reserve Board*, Finance and Economics Discussion Series* #1999-45.

Orphanides A and S Van Norden (1999) "The reliability of output gap estimates in real time," *Federal Reserve Board, Finance and Economics Discussion Series* #1999-38.

Orphanides, A and J C Williams (2002), "Imperfect knowledge, inflation expectations and monetary policy," *Federal Reserve Board Finance and Economics Discussion Series* 2002-27.

Pandher, G S (2002), "Forecasting multivariate time series with linear restrictions using constrained structural state-space models," *Journal of Forecasting,* 21: 281-300.

Patterson, K D (1995), "An integrated model of the data measurement and data generation processes with an application to consumers' expenditures," *The Economic Journal,* 105: 54-76.

Perron, P (1989), "The Great Crash, the oil-price shock and the unit root hypothesis," *Econometrica,* 54: 1361-1401.

Plantier, L C and D Scrimgeour (2002), "Estimating a Taylor rule for New Zealand with a time-varying neutral real rate," *Reserve Bank of New Zealand Discussion Paper,* DP2002/06.

Ploberger, W, W Kramer and K Kontrus (1989), "A new test for structural stability in the linear regression model," *Journal of Econometrics,* 40: 307-318.

Proietti, T (2000), "Comparing seasonal components for structural time series models," *International Journal of Forecasting,* 16: 247-260.

Pyndick, R (1975), "Irreversible investment, capacity choice and the value of the firm," *The American Economic Review,* 78: 969-985.

Quah, D T (1990), "International patterns of growth: I. Persistence in cross-country disparities," *MIT mimeo*.

Quah, D T (1993), "Galton's fallacy and tests of the convergence hypothesis, *Scandinavian Journal of Economics,* 95: 427-443

Quah, D T (1996), "Empirics for economic growth and convergence," *European Economic Review,* 40(6): 1353-1375.

Rasi, C M and J M Viikari (1998), "The time-varying NAIRU and potential output in Finland," *Bank of Finland Working Papers* #6/98.

Richardson, P, L Boone, C Giorno, M Meacci, D Rae and D Turner (2000) "The concept, policy use and measurement of structural unemployment: estimating a time varying NAIRU across 21 OECD countries," *OECD Economics Department Working Paper* #250.

Roberts, J M (1995), "New Keynesian economics and the Phillips curve," *Journal of Money, Credit and Banking,* 27(4): 975-984.

Roberts, J M (1997), "Is inflation sticky?" *Journal of Monetary Economics* 39(4): 176-196.

Roberts, J M (2001), "How well does the new Keynesian sticky-price model fit the data*?" Federal Reserve Board Finance and Economics Discussion Series* 2001-13.

Room, M (2001), "Potential output estimates for the central and east European countries using production function method*," Bank of Estonia Working Paper*s #2/2001.

Rotemberg, J J and M Woodford (1999), "Interest rates rules in an estimated sticky price model" in Monetary Policy Rules (ed J B Taylor), University of Chicago Press: Chicago.

Rudebusch, G D and L Svensson (1999), "Policy rules for inflation targeting" in Monetary Policy Rules (ed J B Taylor), University of Chicago Press: Chicago.

Sargent, T J (1993), Bounded Rationality in Macroeconomics, Oxford University Press: Oxford.

Sargent, T J (1999), The Conquest of American Inflation, Princeton University Press: Princeton.

Schweppe, F (1965), "Evaluation of likelihood functions for Gaussian signals," *IEEE Transactions on Information Theory,* 11(1): 61-70.

Scott, A (2000), "A multivariate unobserved components model of cyclical activity," *Reserve Bank of New Zealand Working Paper,* DP2000/04

Seo, B (1998), "Tests for structural change in cointegrated systems," *Econometric Theory* 14: 222-259.

Shumway, R H and D S Stoffer (1982), "An approach to time series smoothing and forecasting using the EM algorithm," *Journal of Time Series Analysis* 3: 253-264.

Stock, J H and M W Watson (1988), "Testing for common trends," *Journal of the American Statistical Association,* 83(404): 1097-1107.

Stock, J H and M W Watson (1989), "New indexes of leading and coincident economic indicators," NBER Macroeconomics Annual 351-394.

Taylor, J B (1975) "Monetary policy during transition to rational expectations," *Journal of Political Economy,* 83: 1009-1021.

Taylor, J B (1979), "Staggered contracts in a macro model," *The American Economic Review,* 69: 108-113.

Taylor, J B (1980), "Aggregate dynamics and staggered contracts," *Journal of Political Economy,* 88: 1-23.

Taylor, J B (1993), "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy,* 39: 195-214.

Timmermann, A (1994), "Can agents learn to form rational expectations? Some results on convergence and stability of learning in the UK stock market," *The Economic Journal,* 104: 777-797.

Woodford, M (1990), "Learning to believe in sunspots," *Econometrica,* 58(2): 277-307.

Woodford, M (1996), "Control of the public debt: a requirement for price stability?" *NBER Working Paper* #5684.