Estimation and inference in a non-linear state space model: Durable consumption

George Kapetanios and Simon Price
Bank of England
Threadneedle Street
London EC2R 8AH

October 17, 2001

Abstract

Several ways of modelling non-linear state space models have been suggested. The extended Kalman Filter is a tractable way of doing so. One application is to consumer durable demand. Models explaining this flow are normally conditioned on the stock. For the UK, measures of the stock are unavailable. However, it might be estimated from a non-linear state space model. The model is estimated using a linear approximation of the first order conditions for the household's consumption problem and the stock accumulation identity. The results suggest there is very little time variation in depreciation rates over our sample, and that households are close to risk neutrality. Diagnostics suggest further refinement of the model is called for.

1 Introduction

There are several areas in economics where models incorporate stock-flow relationships. In some cases, this creates no special problems, but in others it may be much easier to measure the flow than the stock. In particular, where the stock depreciates at an unknown rate, measurement is particularly difficult. Two examples of this are in the demand for consumer durables and investment. However, it may be possible to estimate the stock within an econometric model. In this paper we survey the techniques, and apply one to a model of consumer durable demand. The preliminary results suggest further refinements are required. In the rest of this paper, we discuss econometric methods for non-linear state space models in Section 2. In the following two sections, we motivate our choice of variable to apply the method to, and set out a model. In Section 5 we present some results. The final section concludes.

2 Methodology

In this section we discuss the methods that may be used to estimate and conduct inference using a non-linear state space model. For reference we present the linear state space model and the linear Kalman filter first. We then discuss an approximation to the linear model that can be estimated with the Kalman filter. By contrast, the other methods proposed in the literature are exact in that no conceptual approximation is involved, but do not provide closed form expressions for the filter equations. Numerical methods are used to provide estimates for the states and consequently the likelihood used for estimating the parameters. All the methods deal with general non-linear, non-gaussian state space models. The focus therefore shifts from conditional means and variances which completely characterise normal distributions to whole densities, since the states are no longer normal.

2.1 The linear state space model and the Kalman filter

A general linear state space model is given by

$$y_t = \boldsymbol{X}_t'\boldsymbol{\beta}_t + \epsilon_t \quad \epsilon \sim i.i.d.(0, \sigma^2) \quad t = 1, \ldots, T \tag{1}$$

$$\boldsymbol{\beta}_t = \boldsymbol{A}_t\boldsymbol{\beta}_{t-1} + \boldsymbol{\varepsilon}_t \quad \boldsymbol{\varepsilon}_t \sim i.i.d.(0, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) \tag{2}$$

The optimality of the Kalman filter crucially depends on assuming that the measurement and transition equation errors are normally distributed. We abstract from issues arising from the estimation of the parameters of the models and concentrate on the estimation of the state variable conditional on the parameters being known. We denote the estimator of $\boldsymbol{\beta_t}$ conditional on the information set up to and including time t by $\boldsymbol{b_t}$. We denote the covariance matrix of $\boldsymbol{b_t}$ by $\boldsymbol{P_t}$. The estimator of $\boldsymbol{\beta_t}$ conditional on the information set up to and including time $t-1$ is denoted by $\boldsymbol{b_{t|t-1}}$. Its covariance matrix is denoted by $\boldsymbol{P_{t|t-1}}$. The Kalman filter comprises sequential application of two sets of equations which recursively deliver the estimates of the state variable and their covariance matrix. The filter is initialised by specifying the estimate of the state and its covariance matrix at the start of the sample. The two sets of equations are given by

$$\boldsymbol{b}_{t|t-1} = \boldsymbol{A}_t\boldsymbol{b}_{t-1} \tag{3}$$

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{A}_t\boldsymbol{P}_{t-1}\boldsymbol{A}_t' + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \tag{4}$$

which are known as the prediction equations, and

$$\boldsymbol{b}_t = \boldsymbol{b}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{x}_t\left(\frac{y_t - \boldsymbol{x}_t'\boldsymbol{b}_{t|t-1}}{f_t}\right) \tag{5}$$

$$\boldsymbol{P}_t = \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{x}_t\left(\frac{1}{f_t}\right)\boldsymbol{x}_t'\boldsymbol{P}_{t|t-1} \tag{6}$$

which are known as the updating equations. f_t is given by $\boldsymbol{X}_t'\boldsymbol{P}_{t|t-1}\boldsymbol{X}_t + \sigma^2$. The log-likelihood of the model may be easily written in terms of the prediction errors. It is given by

$$-\frac{T}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log f_t - \frac{1}{2}v_t^2 \tag{7}$$

where the prediction errors v_t are given by $y_t - \boldsymbol{X}'_t\boldsymbol{b}_{t|t-1}$. The log-likelihood can be used to estimate any unknown parameters.

2.2 Non-linear models

If either the disturbances are non-normal or (as in our case) the system is non-linear, then the Kalman filter cannot be used in its present form. We suggest alternatives for departures from the linear Gaussian model. Let a non-linear state space model be given by

$$y_t = \boldsymbol{X}_t(\boldsymbol{\beta}_t) + \epsilon_t \quad \epsilon \sim i.i.d.(0, \sigma^2) \quad t = 1, \ldots, T \tag{8}$$

$$\boldsymbol{\beta}_t = \boldsymbol{A}_t(\boldsymbol{\beta}_{t-1}) + \boldsymbol{\varepsilon}_t \quad \boldsymbol{\varepsilon}_t \sim i.i.d.(0, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) \tag{9}$$

Now $\boldsymbol{X}_t(.)$ and $\boldsymbol{A}_t(.)$ denote non-linear time varying functions which will be assumed continuous and smooth enough to accept a first order Taylor series expansion.

2.2.1 The extended Kalman Filter

The extended Kalman filter is the easiest extension that can accomodate non-linearity but not non-Gaussianity. It is an approximate filter. Let first order expansions for the functions $\boldsymbol{X}_t(.)$ and $\boldsymbol{A}_t(.)$ around $\boldsymbol{b}_{t|t-1}$ and $\boldsymbol{b}_{t-1}$ be given by

$$\boldsymbol{X}_t(\boldsymbol{\beta}_t) \simeq \boldsymbol{X}_t(\boldsymbol{b}_{t|t-1}) + \hat{\boldsymbol{X}}_t(\boldsymbol{\beta}_t - \boldsymbol{b}_{t|t-1})$$

and

$$\boldsymbol{A}_t(\boldsymbol{\beta}_{t-1}) \simeq \boldsymbol{A}_t(\boldsymbol{b}_{t-1}) + \hat{\boldsymbol{A}}_t(\boldsymbol{\beta}_{t-1} - \boldsymbol{b}_{t-1})$$

where $\hat{\boldsymbol{X}}_t = \frac{\partial \boldsymbol{X}_t(\boldsymbol{\beta}_t)}{\partial \boldsymbol{\beta}_t)}\Big|_{\boldsymbol{\beta}_t = \boldsymbol{b}_{t|t-1}}$ and $\hat{\boldsymbol{A}}_t = \frac{\partial \boldsymbol{A}_t(\boldsymbol{\beta}_{t-1})}{\partial \boldsymbol{\beta}_{t-1})}\Big|_{\boldsymbol{\beta}_{t-1} = \boldsymbol{b}_{t-1}}$ This leads us to express the original non-linear system as an approximate linear model given by

$$y_t = \hat{\boldsymbol{X}}'_t\boldsymbol{\beta}_t + \hat{d}_t + \epsilon_t \tag{10}$$

$$\boldsymbol{\beta}_t = \hat{\boldsymbol{A}}_t\boldsymbol{\beta}_{t-1} + \hat{\boldsymbol{c}}_t + \boldsymbol{\varepsilon}_t \tag{11}$$

$\hat{d}_t$ and $\hat{\boldsymbol{c}}_t$ are the remaining terms in the Taylor expansion and need never be estimated. The states can be estimated by the Kalman filter with the modified prediction and updating equations given by

$$\boldsymbol{b}_{t|t-1} = \boldsymbol{A}_t(\boldsymbol{b}_{t-1})$$

and

$$\boldsymbol{b}_t = \boldsymbol{b}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{x}_t\left(\frac{y_t - \boldsymbol{X}_t(\boldsymbol{b}_{t|t-1})}{f_t}\right)$$

Once prediction errors are obtained the prediction error decomposition may be used to provide expressions for the log-likelihood and therefore a means to estimate any unknown parameters.

2.2.2 Kitagawa (1987)

The method proposed by Kitagawa (1987) is based on the well known prediction and filtering densities for a non-linear and non-Gaussian state space model of the form given in (8). Let the conditional density of $\boldsymbol{\beta}_t$ given observations $(y_1, \ldots, y_m) = \boldsymbol{y}_m$ be denoted by $p(\boldsymbol{\beta}_t|\boldsymbol{y}_m)$. Then, the prediction density is given by

$$p(\boldsymbol{\beta}_t|\boldsymbol{y}_{t-1}) = \int_{-\infty}^{\infty} p(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t-1})p(\boldsymbol{\beta}_{t-1}|\boldsymbol{y}_{t-1})d\boldsymbol{\beta}_{t-1}$$

The filtering density is given by

$$p(\boldsymbol{\beta}_t|\boldsymbol{y}_t) = \frac{p(y_t|\boldsymbol{\beta}_t)p(\boldsymbol{\beta}_t|\boldsymbol{y}_{t-1})}{\int p(y_t|\boldsymbol{\beta}_t)p(\boldsymbol{\beta}_t|\boldsymbol{y}_{t-1})} d\boldsymbol{\beta}_t$$

Therefore, all relevant densities may be obtained by knowledge of the filtering, prediction and the two error densities, assuming the parameters of the model are known. Kitagawa approximates all these densities by piecewise linear functions. Conceptually, this approximation is not problematic if the number of nodes (points of connection between linear pieces) is allowed to go to infinity. This assumption together with the prediction and filtering relations given above provide all the necessary densities over time and from this characteristics of the state and the likelihood is obtained. Once again the likelihood can be used to estimate parameters. From Kitagawa's comments, it appears this method is rather expensive computationally. With modern computing power this is less of an issue.

2.2.3 Gibbs sampler

The Gibbs sampler is an influential technique for carrying out Bayesian analysis belonging to the class of Markov Chain Monte Carlo algorithms. We give a very brief account of Bayesian econometrics to provide a framework for the analysis.

It is well known that one of the distinguishing features of Bayesian econometrics is the treatment of model parameters not as constants but as random variables. Statistical analysis in its simplest form starts with a model whose parameters are assumed to be random variables and for which there exist prior information in the form of a prior probability distribution which these parameters follow. Once the data are observed the prior distributions of the parameters are coupled with the distribution of the data conditional on the parameters and using Bayes theorem a posterior distribution for the parameters may be obtained. This distribution is simply the distribution of the parameters conditional on the data. Then, inference on the parameters may be carried out.

The Gibbs sampler is a powerful technique for carrying out such inference. In particular, say there is a vector of parameters on which inference is to be carried out, denoted by $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)'$. Also assume that the distribution of θ_i conditional on all other parameters and the data is easy to sample from for all $i = 1, \ldots, m$. By contrast, assume that the joint distribution of $\boldsymbol{\theta}$ conditional on the data is not easy to derive analytically or sample from. Obtaining this distribution is the aim of the analysis.

This is common in many situations including non-linear state space models. Then a Markov chain is defined whereby, for some initial value of the parameters, each of θ_i, $i = 1, \ldots, m$ is sampled from its conditional distribution. Once a complete set of θ's are sampled the whole process is repeated, conditioning on the obtained values of $\boldsymbol{\theta}$. After a sufficient number of repetitions the sampled θ's can be used to provide the posterior distribution of the parameters once a proportion of these repetitions has been dropped to remove dependence on initial conditions.

Denote the distribution of θ_i conditional on the data and the rest of the θ's by $p(\theta_i|\boldsymbol{\theta}_{\neq i})$ where $\boldsymbol{\theta}_{\neq i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1} \ldots, \theta_m)'$. Symbolically, the algorithm is as follows

Step 1 Initialise the parameter vector by setting $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \ldots, \theta_m^{(0)})$

Step 2 For each i generate $\theta^{(1)}$ by sampling from $p(.|\boldsymbol{\theta}_{\neq i}^{(0)})$.

Step 3 Repeat Step 2 R times where for $j = 2, \ldots, R$ $\theta^{(j)}$ is generated from $p(.|\boldsymbol{\theta}_{\neq i}^{(j-1)})$. At the end the following set of sampled parameter values will have been obtained: $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(R)}$.

Step 4 Drop the first R_0 of the above R vectors.

Step 5 $\boldsymbol{\theta}^{(R_0+1)}, \ldots, \boldsymbol{\theta}^{(R)}$ follows $p(\boldsymbol{\theta})$ where $p(\boldsymbol{\theta})$ is the posterior distribution of the parameters.

In the context of the non-linear state space model, both the state and the model parameters are considered as random variables. Then the crucial consideration is to specify the model in such a way that it is easy to sample from the conditional distributions of each of the states and parameters. This forms the bulk of the specification task. We will not discuss this here because it is very case dependent. It is sufficient to say that it is more of an art than science.

3 Durable consumption

In many empirical studies of consumer's expenditure durables are either ignored or subsumed into the aggregate. Yet expenditure on consumer durables forms an increasing proportion of total expenditure: see Figure 1. Moreover, while the ratio of consumption to (financial and housing) wealth has fallen since the late 1960s (Figure 2), the ratio of durables to wealth has remained roughly constant (Figure 3), so there is clear evidence that the drivers of behaviour differ. Moreover, the relative price of durables has fallen by about 40% since 1965 (Figure 4). These facts suggest that it would be worthwhile attempting to model durables and non-durables separately. Yet models of the flow will require estimates of the stock, unless we are in a steady state, which is plainly false. There are no official measures of stocks of durables. The stock

Figure 1: Ratio of durable to total consumption

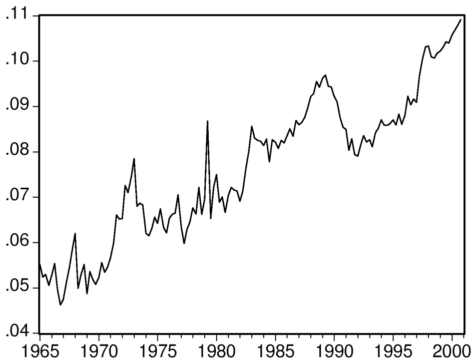
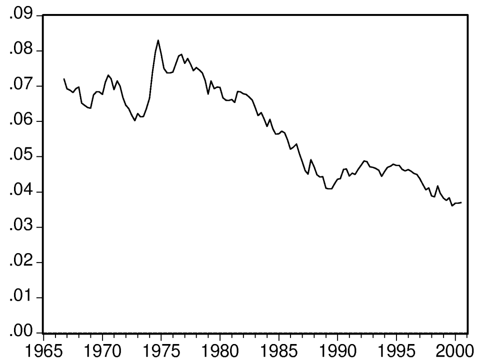


Figure 2: Ratio of total consumption to wealth



of durables may also be large enough to figure significantly in household wealth. For these reasons it would be very useful to have estimates of the stock.

If the stock of durables is K and gross purchases are CD,

$$K_t \equiv K_{t-1}(1 - \delta_{t-1}) + CD_t; \quad s.t. \;\; K_0 = \bar{K}_0 \tag{12}$$

where δ is the depreciation rate. Clearly, if we have an estimate of δ and an initial value K_0, we can infer the stock from the flows. However, the depreciation rate may vary over time for a variety of reasons. For example, the proportion of electronic goods in durables has increased which may raise average depreciation; or the scrapping decision may be cyclical.

Academic research into durables has tended to fall into two categories.[1] Firstly, as an extension of the Euler condition tests of the PI-RE hypotheses, and secondly, on the implications of adjustment costs, especially fixed costs, on the adjustment process. In this paper we take the first approach, and use a model of the relationship between the stock of durables and non-durable expenditure, following Mankiw (1985). The

[1] See Caballero (1994).

Figure 3: Durable consumption to wealth

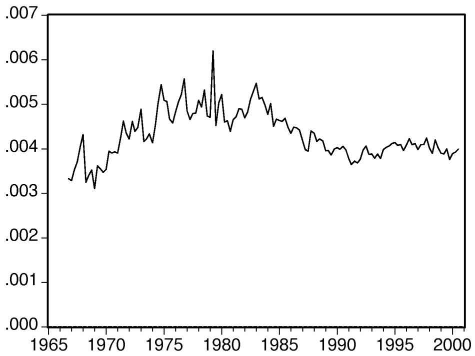
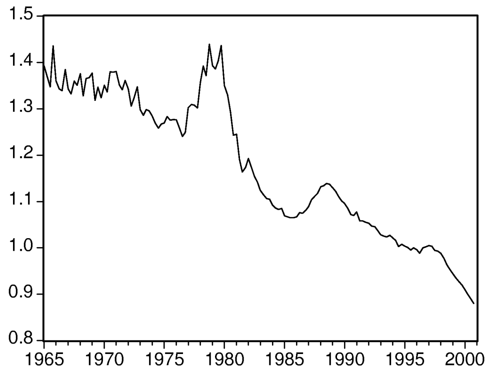


Figure 4: Relative price of durable to non-durable expenditure



model does not attempt to explain the level of consumption. Instead it estimates an expression derived from the FOCs for a simple model.

4 A model of consumption with durable expenditures

Consumers maximise

$$E_t \sum_{s=0}^{\infty} (1+\gamma)^{-s} [U(C_{t+s}) + V(K_{t+s})] \tag{13}$$

subject to an intertemporal budget constraint. Here C is non-durable expenditure and γ is the rate of time preference. $E_t(.)$ denotes the expectation operator conditional on information available at time t. Assume that preferences can be modelled by iso-elastic (CRAA) felicity,

$$U(C) = \frac{C^{1-\alpha}}{1-\alpha}; \quad \alpha > 0 \tag{14}$$

$$V(C) = \frac{\theta K^{1-\beta}}{1-\beta}; \quad \theta,\ \beta > 0. \tag{15}$$

The log-linearised FOCs[2] are

$$\log(C_{t+1}/C_t) = a_0 + 1/\alpha \log(R_{C,t+1}) + \mu_{C,t+1} \tag{16}$$

$$\log(C_{t+1}) = b_0 + 1/\alpha \log(R_{K,t+1}) + \beta/\alpha \log(K_t) + \mu_{K,t+1} \tag{17}$$

where

$$a_0 = \left[\frac{1}{2}\sigma_C^2 - \log(1+\gamma)\right]/\alpha \tag{18}$$

$$b_0 = \left[\frac{1}{2}\sigma_K^2 - \log(\theta+\theta\gamma)\right]/\alpha \tag{19}$$

$$\mu_{j,t+1} = \left[\frac{1}{2}(\epsilon_{j,t+1})^2 - \frac{1}{2}\sigma_j^2 - \epsilon_{j,t+1}\right]/\alpha; \quad j = C, K \tag{20}$$

$$E_t(\mu_{j,t+1}) = 0; \quad j = C, K \tag{21}$$

$$R_{C,t} = 1 + i_t - \pi_{C,t} \tag{22}$$

$$R_{K,t} = (P_{K,t}/P_{C,t})[\delta + i_t - \pi_{K,t}]. \tag{23}$$

$P_{K,t}$ is the durables price level, $P_{C,t}$ is the non-durables price level, δ_t is the unobserved depreciation rate, i_t is the nominal interest rate, $\pi_{K,t}$ is the durables inflation rate, $\pi_{C,t}$ is the non-durables inflation rate, and $R_{C,t}$ and $R_{K,t}$ are the own-price user costs. We assume the level of uncertainty is constant. Here ϵ_j are innovations and $E_t(\epsilon_j) = 0$.

Solving (16) and (17)

$$\log(R_{K,t+1}/R_{C,t+1}) = \alpha(a_0 - b_0) - \beta \log K_t + \alpha \log C_t + v_{t+1} \tag{24}$$

[2]Exact if the error is log-normal.

where

$$v_{t+1} = \alpha(\mu_{C,t+1} - \mu_{K,t+1}). \tag{25}$$

The dependent variable is approximately

$$[\delta_t + (1-\delta_t)(i_t - \pi_{K,t+1})]P_{K,t}/P_{C,t}. \tag{26}$$

The aim is to estimate (24) together with (12) and an equation describing the evolution of δ. We will assume a simple random walk.

5 Results

Arguably, of the methods discussed above, the linearised Kalman filter is the easiest to implement, and we apply it to the non-linear state space model of durable consumption described in Section 4. It is designed to provide an estimate of the unobserved stock of durables using UK data. The key relations are given by (27), (28) and (29),

$$\log(P_{K,t}/P_{C,t}) = a_0 - \log(\delta_t + (1-\delta_t)(i_t + 0.015 - E_t(\pi_{K,t+1})) - \beta \log K_t + \alpha \log C_t + e_t \tag{27}$$

$$\tilde{\delta}_t = \tilde{\delta}_{t-1} + u_{\delta,t} \tag{28}$$

$$K_t = K_{t-1}(1-\delta_t) + CD_t + u_{K,t} \tag{29}$$

where notation is as above. i_t is measured by the London clearing banks' base rate with a 6% per annum retail margin assumed to hold. a_0, β, α are parameters and e_t, $u_{\delta,t}$ and $u_{K,t}$ are error terms which are assumed normal, with covariances σ_e^2, σ_δ^2 and σ_K^2. σ_e^2 and σ_δ^2 are restricted to be strictly positive numbers while σ_K^2 is set to zero, as (29) is an identity. We assume that

$$\delta_t = 0.1exp(-\tilde{\delta}^2) + 0.01 \tag{30}$$

guaranteeing that whereas $\tilde{\delta}_t$ can take values on the whole real line, δ remains bounded between 0.01 and 0.11, which we consider to be a reasonable range. We also restrict the parameters β and α to be positive. To maximise the log-likelihood we use numerical maximisation, and in particular the BFGS algorithm. To initialise the extended Kalman filter we set the initial values of the states to unknown parameters to be estimated and the covariance matrix of the states at time $t = 0$ to zero. We denote these parameters by K_0 and δ_0. We restrict K_0 to lie between 20000 and 100000, a range determined by rough calculations of the likely size of the stock for assumed steady state growth and depreciation rates. The expectation of $\pi_{K,t+1}$ at time t is obtained by using the one-step forecast from an ARDL model where the durable inflation rate is conditioned on past durable and general inflation, GDP growth, nominal interest rates and the ratio of durable to non-durable expenditure. The equation has no significant evidence of autocorrelation ARCH. Details are given in Table 1.

The parameter estimates of our model are given in Table 2. For this model the intial condition estimates go to the corner. Standard errors cannot be calculated as the Hessian is not positive definite, die to the corner solution. The main interest is in the estimates of α and β, the parameters in the felicity functions. A 'standard' case

Table 1: The expectations model

Dependent Variable: $\Delta \log P_{K,t-1}$ Sample: 1980:1 2000:4				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
constant	-0.026991	0.012946	-2.084897	0.0410
$\Delta \log P_{K,t-1}$	0.252518	0.129758	1.946069	0.0559
$\Delta \log P_{K,t-2}$	0.078131	0.108313	0.721345	0.4732
$\Delta \log P_{K,t-3}$	0.168385	0.087243	1.930073	0.0579
$\Delta \log P_{K,t-4}$	-0.104820	0.085927	-1.219869	0.2269
$\Delta \log P_{t-1}$	-0.115274	0.204013	-0.565034	0.5740
$\Delta \log P_{t-2}$	-0.157829	0.199625	-0.790628	0.4320
$\Delta \log P_{t-3}$	-0.308760	0.212005	-1.456379	0.1500
$\Delta \log P_{t-4}$	0.045651	0.221672	0.205938	0.8375
$\Delta \log Y_{t-1}$	-0.266841	0.172915	-1.543190	0.1276
$\Delta \log Y_{t-2}$	0.685876	0.172039	3.986751	0.0002
$\Delta \log Y_{t-3}$	-0.381837	0.156967	-2.432598	0.0177
$\Delta \log Y_{t-4}$	0.160530	0.154239	1.040787	0.3018
$\log CD_{t-1}/C_{t-1}$	-0.011756	0.009359	-1.256150	0.2135
i_{t-1}	0.003304	0.001008	3.278003	0.0017
i_{t-2}	-0.004260	0.001386	-3.072714	0.0031
i_{t-3}	0.001674	0.001427	1.173158	0.2449
i_{t-4}	0.000915	0.000985	0.928048	0.3568

R^2 0.589683
Autocorrelation: LM(8) F: 0.77 (p-value 0.63)
ARCH: F: 0.66 (p-value 0.42)

Table 2: Parameter estimates

K_0	20000
δ_0	0.010
σ_e^2	0.0036310506
σ_δ^2	0.00010000039
a_0	0.056608773
β	0.17887438
α	0.0020950417

Figure 5: Stock of durables: ARDL

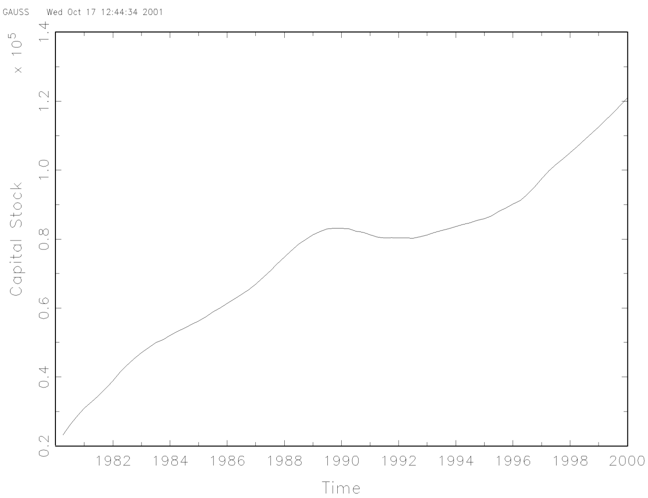
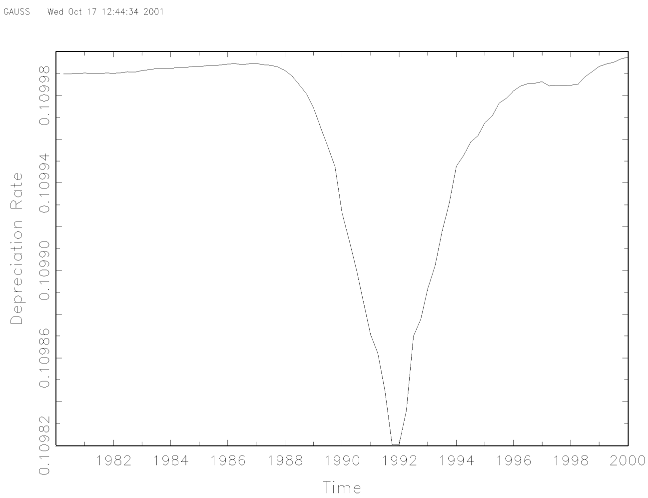


Figure 6: Depreciation rate: ARDL



would be that of homethetic log-linear preferences, so that these are both equal to unity. We are some distance from this. Agents are close to risk neutral. The results also indicate very little time variation in the depreciation rate within sample. The depreciation rate remains close to the upper limit we considered, at nearly 11% per quarter or about 50% per annum. This is higher than that assumed by Mankiw (1985), who took the official US Bureau of Economic Analysis rate of 20%. However, it is not obvious that these rates are implausible. We plot the estimated stock of durables and the depreciation rate in Figures 5 and 6. We also plot durable consumption in Figure 7. With the rapid depreciation, the stock respnds to purchases rapidly, dipping in the early 1990s as households fail to replace depreciated durables.

A by-product of the Kalman filter is an estimate of the 'covariance' of the state estimates around the true state realisations[3]. We have obtained these estimates which provide very narrow bands around the state point estimates and are therefore not plotted. Note that the depreciation rate is modelled as a random walk and therefore has no well defined variance as a result. The narrowness of the bands for the stock of durables estimate reflect the assumption about the transition equation of the stock of durables being an identity, and therefore having an error term with zero variance.

[3]As the state realisations are random variables themselves this estimate is not strictly speaking an estimate of a covariance.

Figure 7: Durable consumption

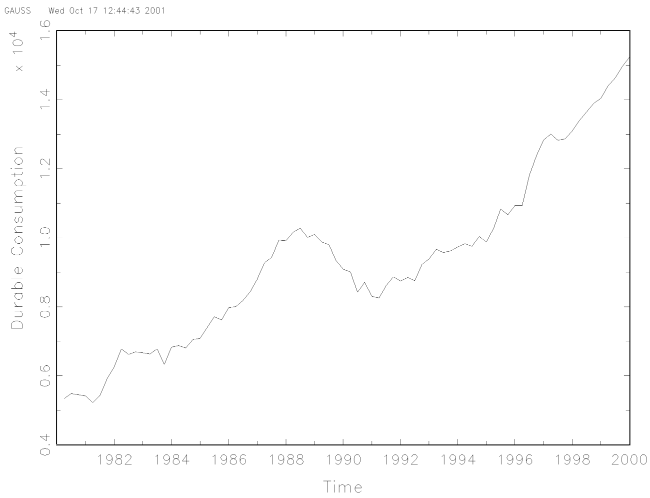


Figure 8: Stock of durables: random walk

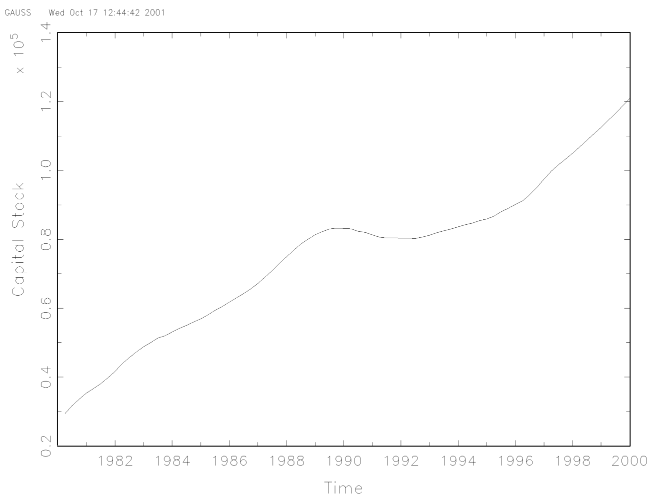


Figure 9: Depreciation rate: random walk

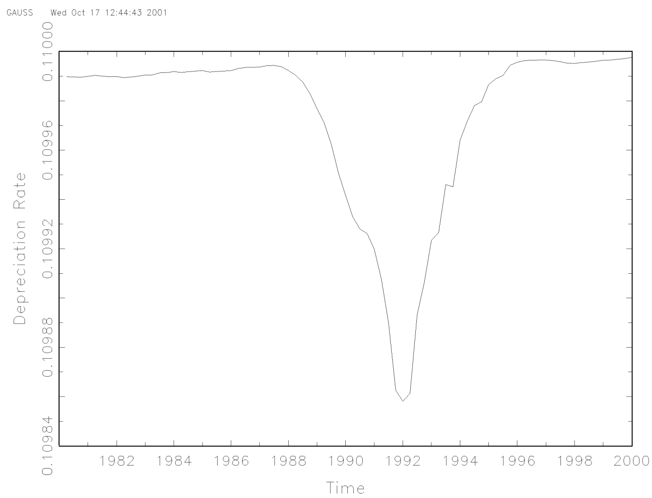


Table 3: Parameter estimates for random walk case

	coeff.	s. error
K_0	26022.775	1153253.84
δ_0	0.01879	0.04979959
σ_e^2	0.00378	0.19514097
σ_δ^2	0.000108	0.00000
a_0	0.1960	1167.59975
β	0.1918	62.8359716
α	0.00264	0.08012490

We also estimate the model using an alternative simpler assumption concerning the expectation of durables inflation. In particular we assume that agents believe inflation to be a random walk and therefore $E_t(\pi_{K,t+1}) = \pi_{K,t}$. Results for this case are presented in Table 3; they hardly differ. However, the initial conditions are not forced to the corner and standard errors can be calculated. The model is clearly very badly determined. We plot the estimated stock of durables and the depreciation rate in Figures 8 and 9.

Neither of these two models have satisfactory statistical features. Apart from the poorly determined standard errors, there is evidence of massive ARCH and serial correlation, although the random walk version does have normal errors. This need not be utterly fatal. Consistent estimation of the coefficients of the model relies firstly on the correct specification of the conditional mean in the measurement equation, and, secondly, given the presence of serial correlation in the error term of the measurement equation, on the independence of the states and the error term of the measurement equation. If these conditions hold then standard results (e.g., in White (1994)) imply consistency of the parameter estimates Nevertheless one would be reluctant to take this model as a good representation of the data.

Table 4: Diagnostics for model residuals (p-values)

	Random walk	ARDL
normality	0.173	0.009
4th order ser. cor.	0.000	0.000
ARCH(4)	0.000	0.000

6 Conclusions

Estimates of the stock of durable goods are arguably essential for meaningful analysis of consumer expenditure; but none exist for the UK. This paper sets out a methodology for estimating such a stock, using an economic model embedded in a non-linear state space model. We briefly discuss three different inference procedures for non-linear state space models. the method of Kitagawa (1987) is the simplest but less widely used. The

Figure 10: Residuals: ARDL

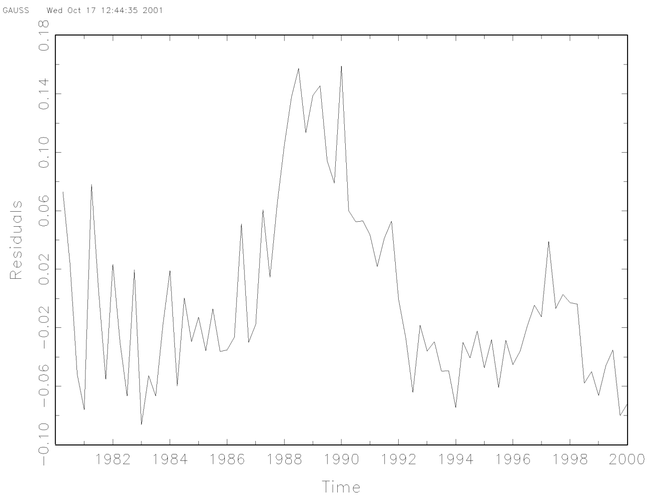
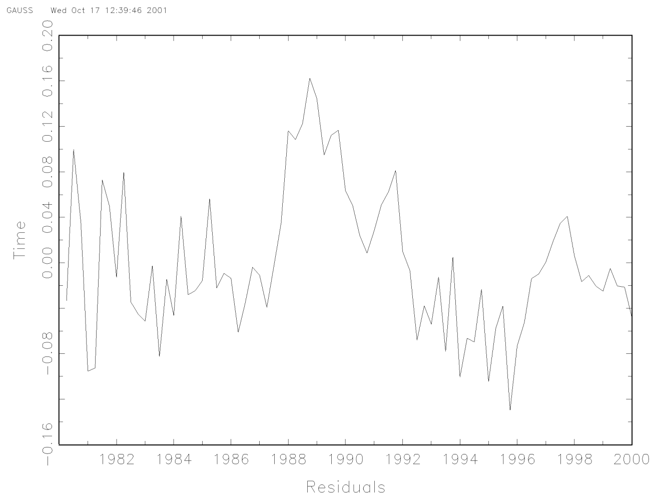


Figure 11: Residuals: random walk



Gibbs sampler is more common, but requires experience in Bayesian analysis to specify the prior and conditional distributions, and is judgemental to an extent. On the other hand the extended Kalman filter, although an approximation, is a promising avenue. As a result we use this to estimate a non-linear state space models derived from intertemporal utility maximisation and provide estimates for the unobserved stock. The model on which they are based assumes no adjustment costs, although this is not an obviously bad assumption. The results indicate very little time variation in the depreciation rate within sample. While the results are not absurd, we do have some reservations. The depreciation rate is close to the upper limit we considered plausible (at nearly 11% per quarter or about 50% per annum), implying a rapid economic depreciation rate. Another aspect of the results is that the parameters on the preference structure imply very little risk aversion. The diagnostics clearly suggest that this is not a well determined statistical model, either.

Several further avenues suggest themselves for future research. Firstly, we could endogenise the depreciation rate, allowing for obsolescence and endogenous scrapping. These issues have been to the fore recently in the context of the role of computers in the capital stock: see Whelan (2000). Secondly we could explore models incorporating adjustments costs. Finally, it should be clear that this methodology might also have applications in the study of investment and the capital stock.

References

[1] Caballero, R J (1994), 'Notes on the Theory and Evidence on Aggregate Purchases of Durable Goods', *Oxford Review of Economic Policy*, 10, 107–117.

[2] Carlin, B.P, N.G. Polson and D.S. Stoffer. (1992), 'A Monte Carlo Approach to Non-normal and Non-linear State-Space Modeling,' *Journal of the American Statistical Association*, 87, 493–500.

[3] Kitagawa, G. (1987), 'Non-Gaussian State-Space Modelling of Nonstationary Time Series', *Journal of the American Statistical Association*, 82, 1033–1063.

[4] Mankiw, N.G. (1985), 'Consumer Durables and the Real Interest Rate', *The Review of Economics and Statistics*, 67, 353–362.

[5] Whelan, K. (2000), 'Computers, Obsolescence and Productivity', *Federal Reserve Board.*

[6] White, H. (1994), *Estimation, Inference and Specification Analysis*, (Cambridge: Cambridge University Press).