Chapter 10

Switching and State Space Models

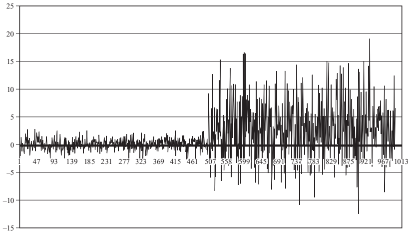
Seasonality and Switching Models

Switching and State Space Models

- Motivation: Episodic nature of economic and financial variables. What might cause these fundamental changes in behaviour?
  - Wars
  - Financial panics
  - Significant changes in government policy
  - Changes in market microstructure - e.g. big bang
  - Changes in market sentiment
  - Market rigidities
- Switches can be one-off single changes or occur frequently back and forth.

Switching Behaviour: A Simple Example for One-off Changes

- Dealing with switching variables



- We could generalise ARMA models (again) to allow the series, y_t to be drawn from two or more different generating processes at different times. e.g.

$$
\begin{aligned}
y_t &= \mu_1 + \phi_1 y_{t-1} + u_{1t} \quad \text{before observation 500} \\
y_t &= \mu_2 + \phi_2 y_{t-1} + u_{2t} \quad \text{after observation 500}
\end{aligned}
$$

How do we Decide where the Switch or Switches take Place?

- It may be obvious from a plot or from knowledge of the history of the series.
- It can be determined using a model.
- It may occur at fixed intervals as a result of seasonalities.
- A number of different approaches are available, and are described below.

Seasonality in Financial Markets

- If we have quarterly or monthly or even daily data, these may have patterns in.
- Seasonal effects in financial markets have been widely observed and are often termed "calendar anomalies".
- Examples include day-of-the-week effects, open- or close-of-market effect, January effects, or bank holiday effects.
- These result in statistically significantly different behaviour during some seasons compared with others.
- Their existence is not necessarily inconsistent with the EMH.

Constructing Dummy Variables for Seasonality

- One way to cope with this is the inclusion of dummy variables- e.g. for quarterly data, we could have 4 dummy variables:

  $D1_t = 1$ in quarter 1 and zero otherwise
  $D2_t = 1$ in quarter 2 and zero otherwise
  $D3_t = 1$ in quarter 3 and zero otherwise
  $D4_t = 1$ in quarter 4 and zero otherwise

- How many dummy variables do we need? We need one less than the "seasonality" of the data. e.g. for quarterly series, consider what happens if we use all 4 dummies

Constructing Quarterly Dummy Variables

		$D1$	$D2$	$D3$	$D4$	Sum
1986	Q1	1	0	0	0	1
	Q2	0	1	0	0	1
	Q3	0	0	1	0	1
	Q4	0	0	0	1	1
1987	Q1	1	0	0	0	1
	Q2	0	1	0	0	1
	Q3	0	0	1	0	1

etc.

- Problem of multicollinearity so $(XX)^{-1}$ does not exist.
- Solution is to just use 3 dummy variables plus the constant or 4 dummies and no constant.

How Does the Dummy Variable Work?

It works by changing the intercept. Consider the folllowing regression:

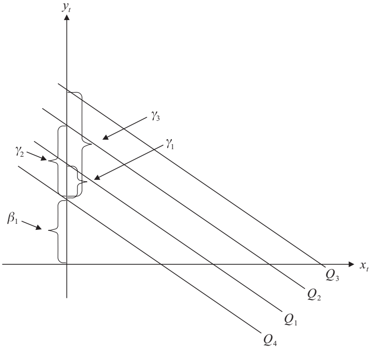
$$y_t = \beta_1 + \gamma_1 D1_t + \gamma_2 D2_t + \gamma_3 D3_t + \beta_2 x_{2t} + \cdots + u_t$$

$\hat{\beta}_1 + \hat{\gamma}_1$ in the first quarter

$\hat{\beta}_1 + \hat{\gamma}_2$ in the second quarter

$\hat{\beta}_1 + \hat{\gamma}_3$ in the third quarter

$\hat{\beta}_1$ in the fourth quarter



Seasonalities in South East Asian Stock Returns

- Brooks and Persand (2001) examine the evidence for a day-of-the-week effect in five Southeast Asian stock markets: South Korea, Malaysia, the Philippines, Taiwan and Thailand.

- The data, are on a daily close-to-close basis for all weekdays (Mondays to Fridays) falling in the period 31 December 1989 to 19 January 1996 (a total of 1581 observations).

- They use daily dummy variables for the day of the week effects in the regression:

$$r_t = \gamma_1 D1_t + \gamma_2 D2_t + \gamma_3 D3_t + \gamma_4 D4_t + \gamma_5 D5_t + u_t$$

- Then the coefficients can be interpreted as the average return on each day of the week.

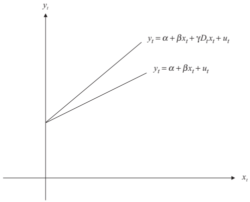
Values and Significances of Day of the Week Effects in South East Asian Stock Markets

	South Korea	Thailand	Malaysia	Taiwan	Philippines
Monday	0.49E-3	0.00322	0.00185	0.56E-3	0.00119
	(0.6740)	(3.9804)**	(2.9304)**	(0.4321)	(1.4369)
Tuesday	−0.45E-3	−0.00179	−0.00175	0.00104	−0.97E-4
	(−0.3692)	(−1.6834)	(−2.1258)**	(0.5955)	(−0.0916)
Wednesday	−0.37E-3	−0.00160	0.31E-3	−0.00264	−0.49E-3
	(−0.5005)	(−1.5912)	(0.4786)	(−2.107)**	(−0.5637)
Thursday	0.40E-3	0.00100	0.00159	−0.00159	0.92E-3
	(0.5468)	(1.0379)	(2.2886)**	(−1.2724)	(0.8908)
Friday	−0.31E-3	0.52E-3	0.40E-4	0.43E-3	0.00151
	(−0.3998)	(0.5036)	(0.0536)	(0.3123)	(1.7123)

Notes: Coefficients are given in each cell followed by *t*-ratios in parentheses;
* and ** denote significance at the 5% and 1% levels, respectively.
Source: Brooks and Persand (2001a).

Slope Dummy Variables

- As well as or instead of intercept dummies, we could also use slope dummies:
- For example, this diagram depicts the use of one dummy – e.g., for bi-annual (twice yearly) or open and close data.



- In the latter case, we could define $D_t = 1$ for open observations and $D_t = 0$ for close.
- Such dummies change the slope but leave the intercept unchanged.
- We could use more slope dummies or both intercept and slope dummies.

Seasonalities in South East Asian Stock Returns Revisited

- It is possible that the different returns on different days of the week could be a result of different levels of risk on different days.
- To allow for this, Brooks and Persand re-estimate the model allowing for different betas on different days of the week using slope dummies:

$$r_t = \left(\sum_{i=1}^{5} \alpha_i D_{it} + \beta_i D_{it} RWM_t \right) + u_t$$

  where D_{it} is the i^{th} dummy variable taking the value 1 for day $t = i$ and zero otherwise, and RWM_t is the return on the world market index
- Now both risk and return are allowed to vary across the days of the week.

Values and Significances of Day of the Week Effects in South East Asian Stock Markets allowing for Time-Varying risks

	Thailand	Malaysia	Taiwan
Monday	0.00322	0.00185	0.544E-3
	(3.3571)**	(2.8025)**	(0.3945)
Tuesday	−0.00114	−0.00122	0.00140
	(−1.1545)	(−1.8172)	(1.0163)
Wednesday	−0.00164	0.25E-3	−0.00263
	(−1.6926)	(0.3711)	(−1.9188)
Thursday	0.00104	0.00157	−0.00166
	(1.0913)	(2.3515)*	(−1.2116)
Friday	0.31E-4	−0.3752	−0.13E-3
	(0.03214)	(−0.5680)	(−0.0976)
Beta-Monday	0.3573	0.5494	0.6330
	(2.1987)*	(4.9284)**	(2.7464)**
Beta-Tuesday	1.0254	0.9822	0.6572
	(8.0035)**	(11.2708)**	(3.7078)**
Beta-Wednesday	0.6040	0.5753	0.3444
	(3.7147)**	(5.1870)**	(1.4856)
Beta-Thursday	0.6662	0.8163	0.6055
	(3.9313)**	(6.9846)**	(2.5146)*
Beta-Friday	0.9124	0.8059	1.0906
	(5.8301)**	(7.4493)**	(4.9294)**

Notes: Coefficients are given in each cell followed by t-ratios in parentheses; * and ** denote significance at the 5% and 1%, levels respectively.
Source: Brooks and Persand (2001a).

Markov Switching Models

- Markov switching models are a generalisation of the simple dummy variables approach described above.
- The universe of possible occurrences is split into m states of the world, called s_t, i=1,..., m.
- Movements of the state variable between regimes are governed by a Markov process.
- This Markov property can be expressed as

$$\mathrm{P}[a < y_t \leq b \,|\, y_1, y_2, \ldots, y_{t-1}] = \mathrm{P}[a < y_t \leq b \,|\, y_{t-1}]$$

Markov Switching Models (Cont'd)

- If a variable follows a Markov process, all we need to forecast the probability that it will be in a given regime during the next period is the current period's probability and a transition probability matrix:

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1m} \\ P_{21} & P_{22} & \dots & P_{2m} \\ \dots & \dots & \dots & \dots \\ P_{m1} & P_{m2} & \dots & P_{mm} \end{bmatrix}$$

  where P_{ij} is the probability of moving from regime i to regime j.

Markov Switching Models – The Transition Probabilities

- Markov switching models can be rather complex, but the simplest form is known as "Hamilton's Filter".
- For example, suppose that m=2. The unobserved state variable, denoted z_t, evolves according to a Markov process with the following probabilities

$$\begin{aligned}
\text{prob}[z_t = 1 | z_{t-1} = 1] &= p_{11} \\
\text{prob}[z_t = 2 | z_{t-1} = 1] &= 1 - p_{11} \\
\text{prob}[z_t = 2 | z_{t-1} = 2] &= p_{22} \\
\text{prob}[z_t = 1 | z_{t-1} = 2] &= 1 - p_{22}
\end{aligned}$$

Markov Switching Models – The Transition Probabilities (Cont'd)

where p_{11} and p_{22} denote the probability of being in regime one, given that the system was in regime one during the previous period, and the probability of being in regime two, given that the system was in regime two during the previous period respectively.

- It must be true that

$$\sum_{j=1}^{m} P_{ij} = 1 \quad \forall \quad i$$

- We then have a vector of current state probabilities, defined as

$$\pi_t = [\ \pi_1 \quad \pi_2 \quad \dots \quad \pi_m\ ]$$

where π_i is the probability that we are currently in state i.

Markov Switching Models – The Transition Probabilities (Cont'd)

- Given π_t and P, we can forecast the probability that we will be in a given regime next period: $\pi_{t+1} = \pi_t P$
- The probabilities for S steps into the future will be given by: $\pi_{t+1} = \pi_t P^s$
- The Markov switching approach is useful when a series is thought to undergo shifts from one type of behaviour to another and back again, but where the "forcing variable" that causes the regime shifts is unobservable.
- The model's parameters can be estimated by maximum likelihood (see Engel and Hamilton, 1990).

An Application of Markov Switching Models to the Real Exchange Rate

- Purchasing power parity (PPP) theory suggests that the law of one price should always apply in the long run such that, after converting it into a common currency, the cost of a representative basket of goods and services is the same wherever it is purchased.

- Under some assumptions, one implication of PPP is that the real exchange rate (that is, the exchange rate divided by a general price index) should be stationary.

- However, a number of studies have failed to reject the unit root null hypothesis in real exchange rates, indicating evidence against PPP theory.

An Application of Markov Switching Models to the Real Exchange Rate (Cont'd)

- It is widely known that the power of unit root tests is low in the presence of structural breaks as the ADF test finds it difficult to distinguish between a stationary process subject to structural breaks and a unit root process.
- In order to investigate this possibility, Bergman and Hansson (2005) estimate a Markov switching model with an AR(1) structure for the real exchange rate, which allows for multiple switches between two regimes.
- The specification they use is

$$y_t = \mu_{s_t} + \phi y_{t-1} + \epsilon_t$$

An Application of Markov Switching Models to the Real Exchange Rate (Cont'd)

where y_t is the real exchange rate, $s_t, (t = 1, 2)$ are the two states and $\epsilon_t \sim N(0, \sigma^2)$.

- The state variable, s_t, is assumed to follow a standard 2-regime Markov process.
- Quarterly observations from 1973Q2 to 1997Q4 (99 data points) are used on the real exchange rate (in units of foreign currency per US dollar) for the UK, France, Germany, Switzerland, Canada and Japan.
- The model is estimated using the first 72 observations (1973Q2–1990Q4) with the remainder retained for out of sample forecast evaluation.

An Application of Markov Switching Models to the Real Exchange Rate (Cont'd)

- The authors use 100 times the log of the real exchange rate, and this is normalised to take a value of one for 1973Q2 for all countries.
- The Markov switching model estimates are obtained using maximum likelihood estimation.

Results

Parameter	UK	France	Germany	Switzerland	Canada	Japan
μ_1	3.554 (0.550)	6.131 (0.604)	6.569 (0.733)	2.390 (0.726)	1.693 (0.230)	−0.370 (0.681)
μ_2	−5.096 (0.549)	−2.845 (0.409)	−2.676 (0.487)	−6.556 (0.775)	−0.306 (0.249)	−8.932 (1.157)
ϕ	0.928 (0.027)	0.904 (0.020)	0.888 (0.023)	0.958 (0.027)	0.922 (0.021)	0.871 (0.027)
σ^2	10.118 (1.698)	7.706 (1.293)	10.719 (1.799)	13.513 (2.268)	1.644 (0.276)	15.879 (2.665)
p_{11}	0.672	0.679	0.682	0.792	0.952	0.911
p_{22}	0.690	0.833	0.830	0.716	0.944	0.817

Notes: Standard errors in parentheses. *Source*: Bergman and Hansson (2005). Reprinted with the permission of Elsevier.

Analysis of Results

- As the table shows, the model is able to separate the real exchange rates into two distinct regimes for each series, with the intercept in regime one (μ_1) being positive for all countries except Japan (resulting from the phenomenal strength of the yen over the sample period), corresponding to a rise in the log of the number of units of the foreign currency per US dollar, i.e. a depreciation of the domestic currency against the dollar.
- μ_2, the intercept in regime 2, is negative for all countries, corresponding to a domestic currency appreciation against the dollar.

Analysis of Results (Cont'd)

- The probabilities of remaining within the same regime during the following period (p_{11} and p_{22}) are fairly low for the UK, France, Germany and Switzerland, indicating fairly frequent switches from one regime to another for those countries' currencies.
- Interestingly, after allowing for the switching intercepts across the regimes, the AR(1) coefficient, ϕ, is a considerable distance below unity, indicating that these real exchange rates are stationary.

So PPP Holds After All?

- Bergman and Hansson simulate data from the stationary Markov switching AR(1) model with the estimated parameters but they assume that the researcher conducts a standard ADF test on the artificial data.
- They find that for none of the cases can the unit root null hypothesis be rejected, even though clearly this null is wrong as the simulated data are stationary.
- It is concluded that a failure to account for time-varying intercepts (i.e. structural breaks) in previous empirical studies on real exchange rates could have been the reason for the finding that the series are unit root processes when the financial theory had suggested that they should be stationary.

Use of the Markov-Switching Model for Forecasting

- Finally, the authors employ their Markov switching AR(1) model for forecasting the remainder of the exchange rates in the sample in comparison with the predictions produced by a random walk and by a Markov switching model with a random walk.

- They find that for all six series, and for forecast horizons up to 4 steps (quarters) ahead, their Markov switching AR model produces predictions with the lowest mean squared errors; these improvements over the pure random walk are statistically significant.

An Application of Markov Switching Models to the Gilt-Equity Yield Ratio

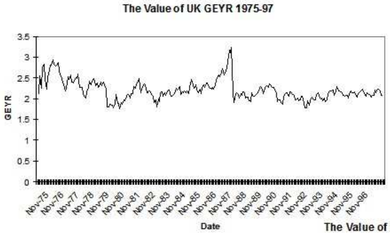
- The gilt-equity yield ratio (GEYR) is defined as the ratio of the income yield on long-term government bonds to the dividend yield on equities.
- It has been suggested that the current value of the GEYR might be a useful tool for investment managers or market analysts
- The GEYR is assumed to have a long-run equilibrium level, deviations from which are taken to signal that equity prices are at an unsustainable level.

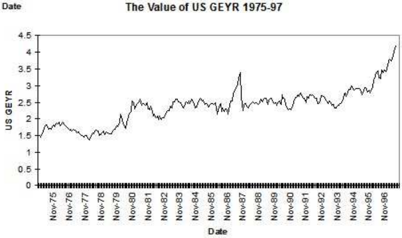
An Application of Markov Switching Models to the Gilt-Equity Yield Ratio (Cont'd)

- Thus, in its crudest form, an equity trading rule based on the GEYR would say, "if the GEYR is low, buy equities; if the GEYR is high, sell equities."

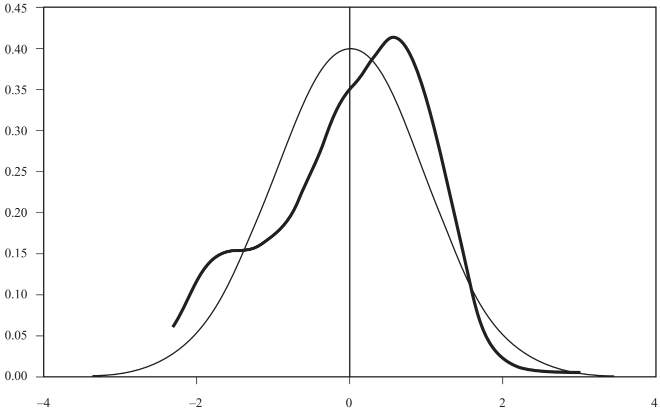
- Brooks and Persand (2001) employ monthly stock index dividend yields and income yields on government bonds covering the period January 1975 until August 1997 (272 observations) for three countries - the UK, the US, and Germany.

Time Series Plots of the GEYR





The Distribution of US GEYR



- The distribution looks as if it could usefully be spilt into two parts

Estimated Parameters for Markov Switching Models

Statistic	μ_1	μ_2	σ_1^2	σ_2^2	p_{11}	p_{22}	N_1	N_2
UK	2.4293	2.0749	0.0624	0.0142	0.9547	0.9719	102	170
	(0.0301)	(0.0367)	(0.0092)	(0.0018)	(0.0726)	(0.0134)		
US	2.4554	2.1218	0.0294	0.0395	0.9717	0.9823	100	172
	(0.0181)	(0.0623)	(0.0604)	(0.0044)	(0.0171)	(0.0106)		
Germany	3.0250	2.1563	0.5510	0.0125	0.9816	0.9328	200	72
	(0.0544)	(0.0154)	(0.0569)	(0.0020)	(0.0107)	(0.0323)		

Notes: Standard errors in parentheses; N_1 and N_2 denote the number of observations deemed to be in regimes 1 and 2, respectively.

Source: Brooks and Persand (2001b).

Analysis of Results

- It is clear that the regime switching model has split the data into two distinct samples - one with a high mean (of 2.43, 2.46 and 3.03 for the UK, US and Germany, respectively) and one with a lower mean (of 2.07, 2.12, and 2.16).
- Also apparent is the fact that the UK and German GEYR are more variable at times when it is in the high mean regime, evidenced by their higher variance (in fact, it is around four and 20 times higher than for the low GEYR state, respectively).

Analysis of Results (Cont'd)

- The number of observations for which the probability that the GEYR is in the high mean state exceeds 0.5 (and thus when the GEYR is actually deemed to be in this state) is 102 for the UK (37.5% of the total), while the figures for the US are 100 (36.8%) and for Germany 200 (73.5%).

- Thus, overall, the GEYR is more likely to be in the low mean regime for the UK and US, while it is likely to be high in Germany.

- The table also shows the probability of staying in state 1 given that the GEYR was in state 1 in the immediately preceding month, and the probability of staying in state 2 given that the GEYR was in state 2 previously.

Analysis of Results (Cont'd)

- The high values of these parameters indicates that the regimes are highly stable with less than a 10% chance of moving from a low GEYR to a high GEYR regime and vice versa for all three series.

Conclusions from the GEYR Application

- The Markov switching approach can be used to model the gilt-equity yield ratio.
- The resulting model can be used to produce forecasts of the probability that the GEYR will be in a particular regime.
- Before transactions costs, a trading rule derived from the model produces a better performance than a buy-and-hold equities strategy, in spite of inferior predictive accuracy as measured statistically.
- Net of transactions costs, rules based on the Markov switching model are not able to beat a passive investment in the index for any of the three countries studied.

Threshold Autoregressive (TAR) Models

- Intuition: a variable is specified to follow different autoregressive processes in different regimes, with movements between regimes governed by an observed variable.

- The model is

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} \textit{ if } s_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} \textit{ if } s_{t-k} \geq r \end{cases}$$

- But what is s_{t-k}? It is the state determining variable and it can be any variable which is thought to make y_t shift from one regime to another.

- If $k = 0$, it is the current value of the state-determining variable that influences the regime that y is in at time t.

Threshold Autoregressive (TAR) Models (Cont'd)

- The simplest case is where $s_{t-k} = y_{t-k}$ we then have a self-exciting TAR, or a SETAR. The model is

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} \text{ if } y_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} \text{ if } y_{t-k} \geq r \end{cases}$$

- We could of course have more than one lag in each regime (and the number of lags in each need not be the same).
- Under the TAR model approach, unlike the Markov switching model, the transitions between regimes are discrete.

Threshold Models: Estimation Issues

- Estimation of parameters in the context of threshold models is complex.
- Quantities to be determined include the number of regimes, the threshold variable, the threshold variable lag, the model order in each regime, the value of the threshold, and the coefficients for each regime.
- We cannot estimate all of these at the same time, so some are usually specified *a priori* based on theory or intuition and the others estimated conditional upon them. E.g., set $k = 1$, $J = 2$, r may not require estimation, etc.

Threshold Models: Estimation Issues (Cont'd)

- The lag length for each regime can be determined using an information criterion conditional upon a specified threshold variable and fixed threshold value. For example Tong (1990) proposes a modified version of AIC:

$$AIC\,(p_1, p_2) = T_1 \ln \hat{\sigma}_1^2 + T_2 \ln \hat{\sigma}_2^2 + 2(p_1 + 1) + 2(p_2 + 1)$$

  where T_1 and T_2 are the number of observations in regimes 1 and 2, respectively, p_1 and p_2 are the lag lengths and $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are the residual variances.

- Estimation of the autoregressive coefficients can then be achieved using nonlinear least squares (NLS).

An Example of a SETAR Model for the French franc / German mark Exchange Rate

- From Chappell *et al.*, 1996, Journal of Forecasting
- The study used daily data from 1/5/90–30/3/92.
- Both the FRF & DEM were then in the ERM which allowed for "managed floating".
- Can use a SETAR to allow for different types of behaviour according to whether the exchange rate is close to the ERM boundary. Currencies are allowed to move up to $^{+}/-2.25\%$ either side of their central parity in the ERM.
- This would suggest the use of the 2-threshold (3-state) SETAR. This did not work as the DEM was never a weak currency then.

An Example of a SETAR Model for the French franc / German mark Exchange Rate (Cont'd)

- The model orders for each regime are determined using AIC
- Ceiling in the ERM corresponded to 5.8376 (log of FRF per 100 DEM).
- The first 450 observations are used for model estimation, with the remaining 50 being retained for out of sample forecasting.
- Forecasts are then produced using the threshold model, the SETAR model with 2 thresholds, a random walk and an AR(2)

Estimated FRF-DEM Regime Switching Model and Out of Sample Forecast Accuracies

Model	For regime	Number of observations
$\hat{E}_t = \underset{(0.0458)}{0.0222} + \underset{(0.0079)}{0.9962} E_{t-1}$	$E_{t-1} < 5.8306$	344
$\hat{E}_t = \underset{(0.2391)}{0.3486} + \underset{(0.0889)}{0.4394} E_{t-1} + \underset{(0.1098)}{0.3057} E_{t-2} + \underset{(0.0866)}{0.1951} E_{t-3}$	$E_{t-1} \geq 5.8306$	103

Source: Chappell *et al.* (1996). Reprinted with permission of John Wiley and Sons.

	Steps ahead				
	1	*2*	*3*	*5*	*10*
Panel A: mean squared forecast error					
Random walk	1.84E-07	3.49E-07	4.33E-07	8.03E-07	1.83E-06
AR(2)	3.96E-07	1.19E-06	2.33E-06	6.15E-06	2.19E-05
One-threshold SETAR	1.80E-07	2.96E-07	3.63E-07	5.41E-07	5.34E-07
Two-threshold SETAR	1.80E-07	2.96E-07	3.63E-07	5.74E-07	5.61E-07
Panel B: Median squared forecast error					
Random walk	7.80E-08	1.04E-07	2.21E-07	2.49E-07	1.00E-06
AR(2)	2.29E-07	9.00E-07	1.77E-06	5.34E-06	1.37E-05
One-threshold SETAR	9.33E-08	1.22E-07	1.57E-07	2.42E-07	2.34E-07
Two-threshold SETAR	1.02E-07	1.22E-07	1.87E-07	2.57E-07	2.45E-07

Source: Chappell *et al.* (1996). Reprinted with permission of John Wiley and Sons.

State Space Models and the Kalman Filter

An Introduction to the State Space Formulation

- State space models were originally used in engineering applications but are now widely used in economics and finance
- They help is to capture relationships between variables that change over time by incorporating for time-varying parameters
- The simplest state space model is a local level model that allows the mean of a series to change over time. It involve two equations:
- A measurement equation that describes how the series of interest moves over time

$$y_t = \mu_t + u_t$$

An Introduction to the State Space Formulation (Cont'd)

- And a transition or state equation that describes how the parameter varies over time

$$\mu_{t+1} = T_t \mu_t + \eta_t$$

- Commonly, T_t is set to one. The two sets of noises u_t and η_t are assumed independent of one another.

Do we Need Time-varying Parameters?

- A test for whether time-varying parameters are required can be conducted by examining the variance of the errors in the transition equation, σ_η^2
- If this is large, it is suggestive that the parameter μ_t is varying over time
- We assess the value of σ_η^2 by taking its ratio to the variance of the disturbances in the measurement equation, σ_η^2/σ_u^2
- σ_η^2 and σ_u^2 are known as *hyperparameters*

A Time-Varying Slope Model

- A simple model for allowing the slope parameter in a regression equation can also be expressed in the state space form
- The measurement equation would be:

$$y_t = \alpha + \beta_t x_t + u_t$$

- And the transition equation:

$$\beta_{t+1} = \beta_t + \eta_t$$

- Such a model could be used to estimate a CAPM specification with a time-varying market risk term.

The Kalman Filter

- The Kalman filter provides a set of recursive formulae used to obtain the state vector (the series of time-varying α_t or β_t) of a state space model
- This works by specifying an arbitrary starting value for the state vector (suppose this is β_1) and then using this with the measurement and transition equations to provide an estimate of the next value of β, which is β_2 and a prediction of y_2
- The actual value of y_2 is compared with the prediction and an adjustment is made to β_2 as a fraction of the prediction error to minimise $\sigma_e ta^2$
- This adjusted value of β_2 is then used to obtain a preliminary estimate of β_3 and so on
- This recursive procedure continues until we have an estimate for the last value of the state variable, β_T

The Kalman Filter (Cont'd)

- Maximum likelihood is then used to estimate appropriate values of the hyperparameters
- Then a final sweep of the Kalman filter to construct final estimates of the state vector