

A Two State Capital Asset Pricing Model

Moshe Fridman *

Institute for Mathematics and its Applications

University of Minnesota

514 Vincent Hall

Minneapolis, Minnesota 55455

March 16, 1994

Abstract

A famous model in financial theory is the Capital Asset Pricing Model (CAPM). In this paper we propose a two state CAPM in which we assume that excess returns for the market and for a particular security are bivariate normally distributed. The parameters of the distribution are determined by the state of an unobserved stationary Markov chain. Two states represent two business regimes that are characterized by low and high volatility. Maximum likelihood estimates for the parameters of the model are obtained via the Baum-Welch algorithm for local maximization of the likelihood function for Markov Regime Models (MRM), also known as Hidden Markov Models (HMM).

We apply the model to monthly return data for three oil industry corporation securities. A comparison of the results with two simpler models, the independent switching regression model and the standard CAPM with one regime, shows a significant improvement in goodness of fit obtained by the proposed two state MRM. Estimates of the periods of high volatility for each corporation depict a general effect of the business cycles on all three processes combined with unique effects stemming from corporation related events.

Key words: Capital Asset Pricing Model, Markov Regime Model, Hidden Markov Model, Baum-Welch algorithm.

AMS 1991 subject classifications. Primary 60J10; Secondary 90A12.

*Research partially supported by ARO Grant DAAL03-91-G-0110

1 Introduction

A famous and widely used model in financial theory is the Capital Asset Pricing Model (CAPM) introduced in Sharpe (1964) and Lintner (1965). The model relates expected returns on capital assets investments with expected market returns. Under the model, the return per period t of an asset R_t is hypothesized to be linearly related to the market return R_{m_t} for that period in the following way:

$$(R_t - R_{f_t}) = \beta(R_{m_t} - R_{f_t}) + \epsilon_t,$$

where ϵ_t are independent and identically distributed as $N(0, \sigma^2)$ and R_{f_t} is the return for period t on a risk-free asset. The quantities $(R_t - R_{f_t})$ and $(R_{m_t} - R_{f_t})$ are referred to as the *excess returns*. The parameter of interest for investors is the *systematic risk* associated with the asset, measured by the regression slope β . The systematic risk of a portfolio is measured by the portfolio β , which is the value-weighted average of the individual asset's systematic risks for the assets held in the portfolio. Portfolio managers are now often charged with the task of achieving a rate of return commensurable with the portfolio's systematic risk. While the traditional CAPM assumes that β is fixed in time, as a practical matter it is not and a deeper understanding of how β changes over time is valuable for appropriate portfolio management. For a recent review of the literature on the time series properties of β see Chang and Weiss (1991).

A natural generalization of the CAPM to a multi-state model is achieved by allowing for model transitions that are governed by a Markov chain on a set of possible models that describe the different states of volatility. This approach has become increasingly popular in econometrics since the introduction of MRM's to the analysis of econometric nonstationary time series by Hamilton (1989,1990). Our purpose is to illustrate the usefulness of MRM's in the study of changes in the β of the CAPM.

Before engaging the analysis, we should note that the CAPM is no longer considered as an accurate model for asset pricing, but due to its simplicity it is popular as an elementary capital asset pricing model.

The proposed model is presented in Section 2, followed by a description of a maximum likelihood estimation procedure and a presentation of smoothing, filtering, and prediction procedures for state identification in Section 3, the empirical results in Section 4, and a summary in Section 5.

2 Model Definition

We begin with a formal definition of the model. We call $\{Y_t\}_{t=1}^T$ the *observed sequence* of an MRM (or HMM) process if there exists a Markov chain $\{Q_t\}_{t=1}^T$ on the state space $\mathcal{S} = \{S_1, \dots, S_N\}$ and cumulative distribution functions F_1, \dots, F_N such that

$$P(Y_1 \leq c_1, \dots, Y_T \leq c_T | Q_t = S_i) = P(Y_1 \leq c_1, \dots, Y_{t-1} \leq c_{t-1}, Q_t = S_i) \cdot F_i(c_t) \cdot P(Y_{t+1} \leq c_{t+1}, \dots, Y_T \leq c_T | Q_t = S_i),$$

for any $1 \leq t \leq T$, $1 \leq i \leq N$, and constants c_1, \dots, c_T . A central hypothesis implied by this definition is that given Q_t the variable Y_t is independent of $\{Y_s, Q_s : s \neq t\}$. What makes the analysis of the process challenging is that the Markov chain $\{Q_t\}$ is *unobserved or hidden*. We denote the transition probability matrix of $\{Q_t\}$ by A and the initial distribution by Π .

Consider a multivariate MRM with a $(p+1)$ -dimensional observed process $\{O_t\}_{t=1}^T$ and an N -state underlying Markov chain $\{Q_t\}_{t=1}^T$. The observed vectors $\{O_t\}_{t=1}^T$ are conditionally independent given $\{Q_t\}_{t=1}^T$. The density of O_t given $Q_t = S_i$ is

$$\phi_i(o_t) = (2\pi)^{-(p+1)/2} |\Sigma_i|^{-1/2} \exp[-(o_t - \mu_i)^T \Sigma_i^{-1} (o_t - \mu_i)/2],$$

for finite μ_i and positive definite matrices Σ_i , $1 \leq i \leq N$. For our regression setting, we let: $O_t = (Y_t, X_t)^T$

$$\mu_i = \begin{pmatrix} \mu_{Y,i} \\ \mu_{X,i} \end{pmatrix} \quad \Sigma_i = \begin{pmatrix} \sigma_{Y,i}^2 & \sigma_{YX,i}^T \\ \sigma_{YX,i} & \Sigma_{XX,i} \end{pmatrix},$$

where Y_t is the univariate dependent variable and X_t is a p dimensional random vector of covariates.

The conditional distribution of Y_t given $X_t = x_t$ and $Q_t = S_i$ is normal with mean

$$\mu_{Y,i} - (\Sigma_{XX,i}^{-1} \sigma_{YX,i})^T \mu_{X,i} + (\Sigma_{XX,i}^{-1} \sigma_{YX,i}) x_t, \quad (1)$$

and variance

$$\sigma_{Y,i}^2 - \sigma_{YX,i}^T \Sigma_{XX,i}^{-1} \sigma_{YX,i}. \quad (2)$$

The parameters

$$\begin{aligned} \alpha_i &= \mu_{Y,i} - (\Sigma_{XX,i}^{-1} \sigma_{YX,i})^T \mu_{X,i}, \\ \beta_i &= (\Sigma_{XX,i}^{-1} \sigma_{YX,i}), \\ \sigma_i^2 &= \sigma_{Y,i}^2 - \sigma_{YX,i}^T \Sigma_{XX,i}^{-1} \sigma_{YX,i}, \end{aligned}$$

correspond to the intercepts, slopes and variances of the CAPM regression planes under regime i , $1 \leq i \leq N$, which along with the transition matrix A are the parameters of interest in the model.

3 ML Estimation and State Identification

3.1 ML Estimation

A maximum likelihood estimation method for the parameters of univariate MRM's with log-concave observed sequence densities was first developed by Baum *et al.* (1970). The numerical procedure for local maximization of the likelihood function is known as the *Baum-Welch* (BW) algorithm, and in its general form was latter called the EM algorithm. The exponential (in T) order of computations required for one evaluation of the likelihood function is reduced to linear by employing a recursive evaluation procedure called the *forward-backward* procedure (see Baum, L.E. (1972)). Liporace (1982) developed a BW type algorithm for multivariate MRM's with elliptically symmetric density functions. His results apply directly to $\{O_t\}$, the multivariate normal MRM observed sequence, offering a numerical algorithm for the local maximization of the likelihood with respect to the parameters $\{\mu_i, \Sigma_i\}_{i=1}^N$ and A . In turn, we can substitute the parameters μ_i, Σ_i in the expressions (1)-(2) by their ML estimates to obtain ML estimates of the parameters α_i, β_i , and σ_i .

We shall hereby present the reestimation formulas for the parameter of the multivariate normal MRM that will be denoted by λ . The proof of the fact that successive applications of the formulas to an initial parameter $\hat{\lambda}$ in the parameter space Λ produces a sequence of increasing likelihood function values, and that a fixed point solution for the reestimation formulas corresponds to a critical point of the likelihood function, follows from Liporace's results.

Forward-Backward Recursion

$$\alpha_t(i) \triangleq P_\lambda(O_1, \dots, O_t, Q_t = S_i) = \left[\sum_{j=1}^N \alpha_{t-1}(j) a_{ji} \right] \phi_i(o_t),$$

$$\beta_t(i) \triangleq P_\lambda(O_{t+1}, \dots, O_T | Q_t = S_i) = \sum_{j=1}^N a_{ij} \phi_j(o_{t+1}) \beta_{t+1}(j),$$

with $\alpha_1(i) = \pi_i \phi_i(o_1)$ and $\beta_T(i) = 1$ for all i .

Reestimation Formulas

$$\begin{aligned}\hat{\mu}_i &= \frac{\sum_{t=1}^T \alpha_t(i) \beta_t(i) o_t}{\sum_{t=1}^T \alpha_t(i) \beta_t(i)} \\ \hat{\Sigma}_i &= \frac{\sum_{t=1}^T \alpha_t(i) \beta_t(i) (o_t - \hat{\mu}_i)(o_t - \hat{\mu}_i)^T}{\sum_{t=1}^T \alpha_t(i) \beta_t(i)} \\ \hat{a}_{ij} &= \frac{\sum_{t=1}^{T-1} \alpha_t(i) a_{ij} \phi_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{t=1}^{T-1} \alpha_t(i) \beta_t(i)} \\ \hat{\pi}_i &= \frac{\alpha_1(i) \beta_1(i)}{\sum_{j=1}^N \alpha_T(j)}\end{aligned}$$

The forward-backward recursion and a scaling scheme that keeps the calculations within the numerical precision of the computer are essential for the implementation of an iterative algorithm to find a fixed point solution of these formulas. An iterative algorithm to solve the reestimation equations can be stated as follows:

Initialization:

Choose $\hat{\lambda} \in \Lambda$;

Choose a level of precision $\epsilon > 0$;

Iteration:

$\lambda \leftarrow \hat{\lambda}$;

Calculate $\alpha_t(i), \beta_t(i)$ for $1 \leq i \leq N, 1 \leq t \leq T$,

for the parameter λ , using the forward-backward algorithm;

Calculate new parameter values $\hat{\lambda}$, according to the reestimation formulas as follows:

Based on λ calculate $\hat{a}_{ij}, \hat{\pi}_i, \hat{\mu}_i, \hat{\Sigma}_i$;

Repeat until the distance $|\hat{\lambda} - \lambda| < \epsilon$;

Substitution:

Using a matrix factorization method such as QR decomposition

calculate $\hat{\Sigma}_{XX,i}^{-1}, 1 \leq i \leq N$;

Substitute $\{\hat{\mu}_i, \hat{\Sigma}_{XX,i}^{-1}, \hat{\sigma}_{YX,i}, \hat{\sigma}_{Y,i}^2\}$ in equations (1)-(2)

to obtain ML estimates of $\{\alpha_i, \beta_i, \sigma_i^2\}$;

As with any numerical procedure for local maximization the solution is sensitive to the initial conditions. Therefore, several runs of the procedure

with initializations at various regions of the parameter space are necessary to find a good candidate for the global maximum. It is not hard to see that the feasibility of the estimates of the probabilities a_{ij}, π_i and the positive definiteness of the estimates of the variance covariance matrices Σ_i are satisfied automatically by the algorithm, as long as the initial point is in the interior of the parameter space Λ .

3.2 State Identification

Identification of the state sequence adds to the understanding of the process, since it enables the analyst to relate historical events to the state process. The state estimation method we shall adopt here is the Maximal A posteriori Probability (MAP) method, by which we estimate Q_t by the state that maximizes the marginal a posteriori probability $P_\lambda(Q_t|Y^s, X^s)$, $1 \leq s \leq T$, where λ is substituted by its MLE. For convenience we use the superscript notation defined as

$$Y^t \triangleq (Y_1, \dots, Y_t) \quad \text{and} \quad X^t \triangleq (X_1, \dots, X_t).$$

We will express the posterior probabilities in term of the forward and backward variables that are readily available from the parameter estimation process.

A different state sequence estimator is the global maximizer estimator given by

$$(\hat{q}_1, \dots, \hat{q}_T) = \arg \max_{(i_1, \dots, i_T)} P_\lambda(Q_1 = S_{i_1}, \dots, Q_T = S_{i_T} | Y^s, X^s).$$

Simulation studies carried out by Fredkin and Rice (1992) compare the marginal estimate with the global maximizer estimate obtained through the Viterbi (1967) algorithm. Fredkin and Rice conclude that there is little difference in the performance of the marginal and global estimates, and that the marginal estimate is slightly less ‘sticky’, and thus more responsive to rapid transitions.

Smoothed State Estimates We begin by considering the smoothed state estimates based on the entire observed data and given by

$$\hat{q}(t|T) = \arg \max_j P_\lambda(Q_t = S_j | Y^T, X^T), \quad 1 \leq t \leq T.$$

The posterior state probabilities are then

$$\begin{aligned} P_\lambda(Q_t = S_j | Y^T, X^T) &= P_\lambda(Y^T, X^T, Q_t = S_j) / P_\lambda(Y^T, X^T) \\ &= \alpha_t(j) \beta_t(j) / P_\lambda(Y^T, X^T) = \alpha_t(j) \beta_t(j) / \sum_k \alpha_T(k). \end{aligned}$$

Maximizing over $1 \leq j \leq N$ we obtain the marginal state estimates.

Filtered State Estimates When we wish to consider only past and present information on the observable sequence, we should use the filtered state estimates,

$$\hat{q}(t|t) = \arg \max_j P_\lambda(Q_t = S_j | Y^t, X^t), \quad 1 \leq t \leq T.$$

In this case we need to maximize over the posterior probabilities

$$\begin{aligned} P_\lambda(Q_t = S_j | Y^t, X^t) &= \frac{P_\lambda(Y_t, X_t | Q_t = S_j, Y^{t-1}, X^{t-1}) P_\lambda(Q_t = S_j | Y^{t-1}, X^{t-1})}{P_\lambda(Y^t, X^t | Y^{t-1}, X^{t-1})} \\ &= P_\lambda(Y_t, X_t | Q_t = S_j) P_\lambda(Q_t = S_j, Y^{t-1}, X^{t-1}) / P_\lambda(Y^t, X^t) \\ &= \alpha_t(j) / \sum_k \alpha_t(k). \end{aligned}$$

State Prediction Clearly, the most useful state estimates from an investors point of view is the one that enables the forecast of the next state Q_{t+1} based on Y^t, X^t . These are given by,

$$\hat{q}(t+1|t) = \arg \max_j P_\lambda(Q_{t+1} = S_j | Y^t, X^t), \quad 1 \leq t \leq T-1.$$

The posterior probabilities maximized by $\hat{q}(t+1|t)$ are given by,

$$\begin{aligned} P_\lambda(Q_{t+1} = S_j | Y^t, X^t) &= \sum_\ell P_\lambda(Q_{t+1} = S_j, Q_t = S_\ell | Y^t, X^t) \\ &= \sum_\ell P_\lambda(Y^t, X^t | Q_{t+1} = S_j, Q_t = S_\ell) P_\lambda(Q_{t+1} = S_j, Q_t = S_\ell) / P_\lambda(Y^t, X^t) \\ &= \sum_\ell P_\lambda(Y^t, X^t, Q_t = S_\ell) a_{\ell j} / P_\lambda(Y^t, X^t) = \sum_\ell \alpha_t(\ell) a_{\ell j} / \sum_k \alpha_t(k). \end{aligned}$$

One can similarly derive m -step state predictions, for example a two-step state prediction,

$$\hat{q}(t+2|t) = \arg \max_j P_\lambda(Q_{t+2} = S_j | Y^t, X^t), \quad 1 \leq t \leq T-2,$$

is obtained by maximizing over $1 \leq j \leq N$ in

$$P_\lambda(Q_{t+2} = S_j | Y^t, X^t) = \sum_r \sum_\ell \alpha_t(r) a_{r\ell} a_{\ell j} / \sum_k \alpha_t(k).$$

4 Empirical Results

4.1 The Data

We apply our model in the analysis of the systematic risk associated with three major oil industry corporation securities, using complete monthly return data for the 144 months starting from January 1980 to December 1991. The data were collected from the Center for Research in Security Prices (CRSP) tapes.

We chose the 30-day Treasury bill rate and monthly returns on the Standard and Poor's 500 Stock Index (S&P500) as the respective proxies for the riskless and market return rates. The oil corporation securities we analyze are Exxon, Texaco, and Chevron. The 30-day Treasury bill rate is subtracted from the monthly return rate for each of the three securities, and from the monthly return rate for the S&P500, to obtain the excess returns.

Figure 1 presents the price time series for the three securities and the S&P500 index. It is evident that during the periods 80-81 and 86-87 all series exhibit a relative high volatility. Naturally, there are also corporation specific periods of high volatility that are related to particular events that may have caused uncertainty amongst the corporation stockholders. These will be discussed later.

4.2 Parameter Estimation and Model Testing

We fitted to each of the three oil corporations security data three models, as follows:

1. The MRM (as described in Section 1) with $N = 2$.
2. The MRM with the constraint $a_{11} = 1 - a_{22}$. We will refer to this model as the Independent Switching Model (ISM).
3. The MRM with a single state referred to as the Single State Model (SSM). This model is equivalent to the MRM with the constraints

$$\mu_{X,1} = \mu_{X,2}, \mu_{Y,1} = \mu_{Y,2}, \sigma_{X,1} = \sigma_{X,2}, \sigma_{Y,1} = \sigma_{Y,2}, \sigma_{YX,1} = \sigma_{YX,2}.$$

Figure 1: Monthly prices for three oil corporation securities and the S&P500 index for the period January 1980 to December 1991. Scale for the S&P500 price is on the right margin.

Table 1: Parameter estimates for the models fitted to the three oil corporation securities.

parameter	Exxon			Texaco			Chevron		
	MRM	ISM	SSM	MRM	ISM	SSM	MRM	ISM	SSM
a_{11}	.480	.264	–	.819	.066	–	.823	.067	–
a_{22}	.882	.736	–	.941	.934	–	.916	.933	–
α_1	.018	.045	.007	.011	.038	.005	.010	.109	.004
α_2	.006	–.009	–	.004	.005	–	.009	–.005	–
β_1	.976	.988	.743	1.03	1.21	.714	1.24	1.27	.923
β_2	.529	.598	–	.456	.526	–	.419	.856	–
σ_1	.046	.030	.040	.101	.140	.069	.092	.103	.069
σ_2	.036	.032	–	.042	.056	–	.047	.055	–

The maximum likelihood estimators of the parameters of the three models for each one of the oil corporations are given in Table 1. In all three securities, the MRM maximum likelihood estimates associate the high volatility state S_1 with a lower self-transition probability and a higher slope and intercept.

The difference in parameter estimate values between the two states is most noticeable in the parameter estimates for Texaco and Chevron, and to a lesser extent in those for Exxon. For the Texaco and Chevron securities, the size of the estimated systematic risk and the error term standard deviation under state S_1 are roughly double the size of the corresponding parameters under S_2 . For both corporations, the average duration of stay in S_1 is about 5.5 months versus 13 months in S_2 . The differences in slope and error term standard deviation between the two states is considerably smaller for the Exxon security. Also, the self-transition probabilities have a different structure for the Exxon security, exhibiting a closer relation to an independent switching model.

The CAPM predicts an intercept of zero for the regression line, but the MLE's for the intercepts of the MRM are all positive, with the high volatility state intercepts α_1 always bigger than those for the low volatility state intercepts α_2 . In fact, we note that the intercepts are all around 0.5%, even for the SSM which is equivalent to the conditional OLS model given X_t . Previous empirical CAPM studies have encountered the same obstacle.

Table 2: Goodness of fit tests to compare the MRM with the constrained ISM and SSM.

Model	Exxon		Texaco		Chevron	
	Loglik.	P	Loglik.	P	Loglik.	P
MRM	500.1	–	436.99	–	428.8	–
ISM	497.0	.016	430.92	.0005	425.6	.0114
SSM	492.2	.0015	415.76	.0000	415.3	.0000

A possible explanation for this result is that our proxies for the risk-free and market returns for the estimating period 1/80 through 12/91 are biased, and since for all the models we used the same proxies, we see a similar phenomenon in all of the estimates. We encounter the non-zero intercept problem also in several previous empirical studies of the CAPM.

The MLE's for the S_2 self-transition probability, slopes and error term standard deviations of the ISM parameters are closely related to those of the MRM. However, we note that all the slope estimates have higher values (especially for Chevron), compensating for the lower self-transition probability (constrained to be the complement to one of a_{22}) for the high volatility state S_1 . The MLE's for the SSM parameters are, as one could expect, a compromise between the MLE values for the two states in the MRM. This result can be seen clearly in the plots of Figure 2, in which the regression lines for the MRM and SSM are plotted for each corporation.

We compare the goodness of fit of the MRM with that of the other two constrained models by assuming an asymptotic χ^2 distribution for the likelihood ratio test statistic. A χ^2 distribution with one degree of freedom is used for the test of the MRM versus the ISM (only one constraint), and a χ^2 distribution with five degrees of freedom is used for the test of the MRM versus the SSM. The results for the tests of hypotheses are given in Table 2. Based on these tests, the MRM is a significant improvement compared to the constrained models especially for the Texaco and Chevron corporations. The slightly less significant results for the Exxon corporation data could be expected because of the similarity between the MRM and ISM parameter estimates.

Figure 2: Scatter plots of excess returns on the S&P500 versus the three oil corporation securities. Solid lines are the regression lines for the MRM; dotted line is the SSM regression line.

4.3 Identification the State Sequence

For each oil corporation, we calculated several sequences of posterior probabilities as described in Section 3.2, and examined the resulting regime switches in light of actual events. The inference is based on the MRM with parameter values given by the MLE.

The results are presented in Figures 3-5, each consisting of four panels. The top panel presents the observed excess return Y_t , $1 \leq t \leq 144$; the second panel shows the smoothed state S_1 probabilities $P(Q_t = S_1 | Y^T, X^T)$, $1 \leq t \leq 144$; the third panel shows the filtered state S_1 probabilities $P(Q_t = S_1 | Y^t, X^t)$, $1 \leq t \leq 144$; and the bottom panel shows the predicted state S_1 probabilities $P(Q_{t+1} = S_1 | Y^t, X^t)$, $2 \leq t \leq 144$. We next discuss the relationships between the observed return sequences and the various estimates presented. In some instances we point out actual events that are possibly related to regime switches as portrayed by these estimates.

Excess Returns All three sequences of excess returns shown in the top panels of Figures 3-5 generally follow the same trends, but differ in their volatility level. Exxon's excess returns are relatively stable compared to the distinctively more volatile Chevron's and Texaco's excess returns. This difference in volatility is most strongly manifested in the first period 80-87. The latter period 88-91 is more stable for all three securities. According to the National Bureau for Economic Research, there were three business cycle turning points during our estimation period, with trough points on 7/80, 11/82, and 3/91. The sharp peaks and troughs in excess returns during the years 80-82 are possible a reflection of the rapid cycles of the economy during those years. Such extreme swings in excess returns are not observed during the months adjacent to the single turning point in 3/91. Another noticeable feature in all three sequences is the sharp trough in 10/87, the month of the stock exchange market crash.

Smoothed and Filtered High Volatility State Probability It is evident from the first and second panels in Figures 3-5 that the months for which we would have concluded that the process is in the high volatility state S_1 (a smoothed probability greater than 0.5) coincide with the visible periods of high volatility and instability in the excess returns. The filtered state sequence estimates showed in the third panels closely resemble the smoothed sequences, except for being rougher, as one would expect. Table 3 summarizes the regime switching information obtained from the smoothed state

Figure 3: (a) Exxon's excess returns for 1/80-12/91; (b) Smoothed probability of state S_1 , $P(Q_t = S_1|Y^T, X^T)$ for 1/80-12/91; (c) Filtered probability of state S_1 , $P(Q_t = S_1|Y^t, X^t)$ for 1/80-12/91; (d) Predicted probability of state S_1 , $P(Q_{t+1} = S_1|Y^t, X^t)$ for 2/80-12/91.

Figure 4: (a) Texaco's excess returns for 1/80-12/91; (b) Smoothed probability of state S_1 , $P(Q_t = S_1|Y^T, X^T)$ for 1/80-12/91; (c) Filtered probability of state S_1 , $P(Q_t = S_1|Y^t, X^t)$ for 1/80-12/91; (d) Predicted probability of state S_1 , $P(Q_{t+1} = S_1|Y^t, X^t)$ for 2/80-12/91.

Figure 5: (a) Chevron's excess returns for 1/80-12/91; (b) Smoothed probability of state S_1 , $P(Q_t = S_1|Y^T, X^T)$ for 1/80-12/91; (c) Filtered probability of state S_1 , $P(Q_t = S_1|Y^t, X^t)$ for 1/80-12/91; (d) Predicted probability of state S_1 , $P(Q_{t+1} = S_1|Y^t, X^t)$ for 2/80-12/91.

sequence estimates. These results are naturally in close agreement with the average values one can obtain from the MLE values for the self-transition probabilities.

Despite the fact that the estimated visits of the Exxon's excess return process to state S_1 are short and frequent compared to those of the Texaco and Chevron processes, we find overlap among the majority of the three corporation's estimated high volatility periods. Specifically, all the processes were estimated to be in S_1 on the 10 months given by the periods 1-3/80, 10-12/80, 8/82, 1/87, and 10-11/87. The Exxon and Chevron S_1 estimates agree also on the two months of 8-9/86, and the Texaco and Chevron estimates agree on 10 additional months.

A distinctive feature of the Texaco excess return sequence is the sharp trough on 10-11/85. It is related to a court ruling against Texaco, awarding Pennzoil damages of about \$11 billion, after finding that Texaco's takeover of Getty Oil corporation was improper. This behavior is indeed captured by the smoothed state sequence estimate in that only the Texaco estimates indicate an S_1 period during those months. Another unique switch to S_1 is that of Exxon during 4/83, possibly related to a court ruling in 3/83 against Exxon, ordering the corporation to pay \$1.5 billion to refineries for overcharging them over the period 75 to 81. Also, it is interesting to note that there is no significant visible effect of the 3/89 Exxon Valdez oil spill disaster on Exxon's monthly excess returns, nor do the smoothed state sequence estimates indicated such an effect.

Table 3: Summary statistics for the smoothed state sequence estimates.

	<u>Exxon</u>	<u>Texaco</u>	<u>Chevron</u>
Months in S_1	14	32	45
# of transits to S_1	7	5	4
Average stay in S_1	1.75	5.33	9

Predicted High Volatility State Probability The predicted probabilities for state S_1 shown in the bottom panels of Figures 3-5 follow the same pattern as the corresponding smoothed and filtered probabilities, but two differences emerge from the graphs. The first is that the prediction probabilities tend to be closer to 0.5, since there is less information and more uncertainty regarding the state. The second, and more crucial difference,

is that the prediction probabilities are shifted to the right by one or two positions relative to the smoothed probabilities, or in other words the detection of switches is delayed by one or two periods. This is not obvious from the graphs, but the correlations between the smoothed and the shifted predicted probability estimate sequences in Table 4 help to clarify the point. To construct the table, we dropped the first smoothed probability estimate, since it has no corresponding predicted probability estimate, and then calculated correlations shifting the predicted probability sequence to the right, one position at a time.

Table 4: Correlations between the estimated smoothed and prediction probability sequences.

<u>Positions shifted</u>	<u>T</u>	<u>Exxon</u>	<u>Texaco</u>	<u>Chevron</u>
0	143	.378	.704	.661
1	142	.966	.906	.884
2	141	.574	.925	.906
3	140	.176	.809	.799

The correlations in Table 4 suggest that a one period delay for the Exxon process, and one or two periods of delay for the Texaco and Chevron processes can be expected. This difference in the delay of the switch detection can be explained by noting that for the Exxon process the probabilities of transition are the largest and hence it is easier to detect a switch given one observation from the new state.

5 Conclusions

By considering a two state CAPM model, we can improve the assessment of the level of systematic risk associated with a security, as measured by the β of the CAPM. In the three cases we examined we saw that our model offers a significant improvement in fit over the simpler constrained models. The independent switching model tends to overestimate the systematic risks, while the systematic risk obtained from the single state model is a compromise between the two systematic risks in the two state MRM. The better fit of the two state MRM over the one state model supports the idea that β is not fixed in time. Furthermore, the proposed model is a simple

tool to measure this instability that need to be taken into account in event studies and studies of market efficiency. Clearly, the understanding of the behavior of β is the key to a more accurate portfolio risk management.

Based on the parameter and state sequence estimates that we obtained for the two state MRM we learn that: (1) the high volatility state β can be more than double the size of the more stable state β , hence making it a higher risk state, and (2) the duration of stay in the high risk state is typically shorter than the one for the low risk state. By comparing smoothed state estimates for the three oil corporations, we were able to relate periods of high volatility to the overall market status and to some particular corporation specific events that had a sizable effect on stock prices.

From a broader viewpoint, this study illustrates the usefulness of Markov regime switching models in the analysis of processes that exhibit only local homogeneity. Such complex processes can be found in a variety of scientific fields, and we believe the ideas presented here can be successfully applied in many such contexts.

Acknowledgements This work is drawn from the authors Ph.D. dissertation written at the university of Pennsylvania under the direction of Professor J. Michael Steele to whom I am indebted for his invaluable guidance and advice.

References

- Baum, L.E. (1972), "An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," *Inequalities, III: Proceedings of the symposium*, (Shisha, Qved ed.), New York: Academic Press, 1-8.
- Baum, L.E., Petrie, T., Soules, G. and Weiss, N. (1970), "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Statist.*, **41**, 164-171.
- Chang, W.C. and Weiss, D.E. (1991), "An examination of the time series properties of Beta in the market model," *Journal of the American Statistical Association*, **86**, 883-890.
- Fredkin, D.R. and Rice, J.A. (1992), "Bayesian restoration of single channel patch clamp recordings," *Biometrics*, **48**, 427-448.
- Hamilton, J.D. (1989), "A new approach to the economic analysis of non-stationary time series and the business cycle," *Econometrica*, **57**, 357-384.
- (1990), "Analysis of time series subject to changes in regime," *Journal of Econometrics*, **45**, 39-70.
- Lintner, J. (1965), "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," *Rev. Econom. and Statist.*, **47**, 13-37.
- Liporace, L.A. (1982), "Maximum Likelihood estimation for multivariate observations of Markov sources," *IEEE Trans. Inform. Theory*, **IT-28**, 729-734.
- Sharpe, W.F. (1964), "Capital asset prices: A theory of market equilibrium under conditions of risk," *J. Finance*, **19**, 425-442.
- Viterbi, A.D. (1967), "Error bounds for convolutional codes and an asymptotically optimal decoding algorithm," *IEEE Trans. Inform. Theory*, **IT-13**, 260-269.