Regime Switching in Stock Market Returns

Simon van Norden
International Department
Bank of Canada

Huntley Schaller
Economics Department
Carleton University

November 18, 1993

Fields: financial markets, time-series econometrics, regime-switching, asset pricing

We thank participants at the International Conference on Financial Market Dynamics and Forecasting (Paris, Sept. 1993), especially Clive Granger, Jurgen Kaehler, and Neil Shephard, for useful comments. Professor Schaller thanks the Social Sciences and Humanities Research Council of Canada for financial support and Boston University for providing a convenient location for work on this project. Any errors are our own. We are solely responsible for the views expressed here; no responsibility for them should be attributed to the Bank of Canada.

Address for correspondence: Huntley Schaller, Economics Department, Carleton University, Ottawa, Ontario, Canada K1S 5B6; telephone: (613) 788-2600 ext. 3751; email: schaller@ccs.carleton.ca; fax: (613) 788-3906.

Abstract

In this paper, we use an extension of Hamilton's (1989) Markov switching techniques to describe and analyze stock market returns. Using new tests, we find very strong evidence of switching behaviour. A major innovation of our work is to use a multivariate specification which allows us to examine whether the price-dividend ratio has marginal predictive power for stock market returns after accounting for state-dependent switching. We find strong evidence of predictability. The response of returns to the past price-dividend ratio is strongly asymmetric - about four times larger in the low-return state than in the high-return state. A second innovation in our work is to allow the probability of transitions from one regime to another to depend on economic variables. Here again, we find an asymmetric response to the past price-dividend ratio.

I. Introduction

In an influential paper, Hamilton (1989) has suggested Markov switching techniques as a method for modelling non-stationary time series. In the Hamilton (1989) approach, the parameters are viewed as the outcome of a discrete-state Markov process. For example, expected returns in the stock market may be subject to occasional, discrete shifts. In this paper, we use an extension of Hamilton's approach to describe and analyze stock market returns.

The Markov switching technique allows us to pose a variety of interesting new questions. Can we distinguish distinct regimes in stock market returns? How do the regimes differ? How frequent are regime switches and when do they occur? Are returns predictable, even after accounting for regime switches? Are regime switches predictable? The answers to these questions give us a new set of stylized facts about stock market returns.

Three previous papers have used the techniques proposed by Hamilton (1989) to examine stock market returns. Schwert (1989) considers a model in which returns may have either a high or a low variance and switches between these return distributions are determined by a two-state Markov process. Turner, Startz, and Nelson (1989) consider a Markov switching model in which either the mean, the variance, or both may differ between two regimes. Using S&P monthly index data for the period 1946-1989, they consider univariate specifications with constant transition probabilities. Hamilton and Susmel (1993) propose a model with sudden discrete changes in the process which governs volatility. They find that a Markov-switching model provides a better statistical fit to the data than ARCH models without switching.[1]

---

[1]There is also an interesting paper by McQueen and Thorley (1991) which uses Markov chains to test the random walk model of stock market returns. Because the authors distill the information in stock market returns into a discrete state variable, their work does not fit into the

Our paper is closest to that of Turner, Startz, and Nelson (1989), but we extend their work in several directions. First, we consider the interesting period including the 1929 crash, the Great Depression, and World War II, along with the post-war period. Second, we examine whether stock market returns are predictable even after accounting for Markov switching behaviour. Third, we examine whether the transition probabilities of the Markov chain vary over time in response to changes in economic variables. Finally, our paper uses the new tests of Hansen (1992,1993) and Garcia (1992) to determine whether the evidence of switching is statistically significant.

Our paper is organized as follows. Section II examines whether there are distinct regimes in stock market returns. Section III estimates diffferent specifications of switching behaviour. Section IV tests whether economic variables (specifically, the price-dividend ratio) have marginal predictive power for stock market returns. Section V considers time variation in transition probabilities. Section VI summarizes some of the main empirical results.

II. Are there regimes in stock market returns?

In this section, we examine whether there is evidence of distinct regimes in stock market returns. In a formal sense, regime-switching econometric models refer to a situation in which stock market returns are drawn from two different distributions, with some well-defined stochastic process determining the likelihood that each return is drawn from a given distribution.

Regime switching in returns could arise in a number of ways. One example is provided by Cecchetti, Lam, and Mark (1990), who consider a Lucas asset pricing model in which the

---

Hamilton (1989) framework, which treats stock market returns as a continuous variable.

economy's endowment switches between high economic growth and low economic growth. They show that such switching in fundamentals accounts for a number of features of stock market returns, such as leptokurtosis and mean reversion. A second example is the Blanchard-Watson (1982) model of stochastic bubbles. In each period, a bubble may either survive or collapse; in such a world, returns would be drawn from one of two distributions - surviving bubbles or collapsing bubbles.[2]

Previous papers which have estimated Markov switching models on stock market returns have faced a problem in testing the null hypothesis of no switching against the alternative hypothesis of switching. The problem is that the transition probabilities are not identified under the null hypothesis of no switching. Under these circumstances, the asymptotic distributions of likelihood ratio, Lagrange multiplier, and Wald tests are non-standard. Recent econometric work by Hansen (1992,1993) has suggested a method for calculating the non-standard asymptotic distribution. Garcia (1992) shows how Hansen's work can be applied to the problem of testing for Markov switching by treating the transition probabilities as nuisance parameters.

To determine whether there is switching in stock market returns, we consider four specifications. In the first, stock market returns R are drawn from a single Gaussian distribution with mean α_0 and variance σ_0:

$$R_t = \alpha_0 + \sigma_0 \epsilon_t \tag{1}$$

and ε_t is a standard Gaussian variable. This is the specification for the null hypothesis of no switching.

[2]See, for example, Hall and Sola (1993) or van Norden and Schaller (1993a).

We consider three alternative hypotheses. In the first, returns are drawn from two distributions with different means (α_0 and α_1):

$$R_t = \alpha_0(1-S_t) + \alpha_1 S_t + \sigma_0 \epsilon_t \tag{2}$$

where S_t is a binary state variable which follows a first-order Markov chain:

$$\begin{aligned} Pr(S_t = 0 | S_{t-1} = 0) &= q \\ Pr(S_t = 1 | S_{t-1} = 1) &= p \end{aligned} \tag{3}$$

In words, (2) means that the probability that a given state will occur this period depends only on the state last period. The probability that state 0 (1) will persist from one period to the next is q (p).

The second alternative hypothesis is that returns are drawn from two distributions with the same mean but different variances (σ_0 and σ_1):

$$R_t = \alpha_0 + [\sigma_0(1-S_t) + \sigma_1 S_t]\epsilon_t \tag{4}$$

The third alternative hypothesis allows for different means and variances:

$$R_t = \alpha_0(1-S_t) + \alpha_1 S_t + [\sigma_0(1-S_t) + \sigma_1 S_t]\epsilon_t \tag{5}$$

Under each of the alternative hypotheses, the distribution from which stock market returns are drawn is determined by the state variable S_t.

The data we use are the CRSP value-weighted monthly stock market returns (including dividends) for the period January 1929 to December 1989. All of our results are based on excess returns, which are constructed by subtracting the monthly rate of return on a 90-day T-bill from the CRSP returns.

Table 1 presents the log likelihood for the null hypothesis and each alternative.[3] As the second column of the table shows, when we test (2) against the null of (1), the likelihood ratio statistic is 95.1. Garcia (1992) shows that the 5% critical value for the likelihood ratio statistic is 10.34 and the 1% critical value is 13.81 for this case. Our results therefore imply very strong rejection of the null hypothesis of no switching. The results in columns three and four imply even stronger rejections of the null hypothesis of no switching.

The results in the upper panel of Table I parallel the findings of Turner, Startz, and Nelson (1989), who examined U.S. stock market returns from January 1946 to December 1987. Since the work of Hansen (1992,1993) and Garcia (1992) appeared after 1989, Turner, Startz, and Nelson (1989) were not able to use their test. In the lower panel of Table 1, we show that an application of the Hansen-Garcia test to the log likelihoods reported by Turner, Startz, and Nelson (1989) leads to a rejection of the null hypothesis of no switching at the 1% level in all specifications.

The conclusion is clear. There is strong evidence of regime switching in U.S. stock market returns. The evidence for switching is robust to different specifications of the nature of

[3] Additional information on the estimation techniques used may be found in the appendix.

switching. The conclusion does not appear to be sensitive to the estimation period, since the results in Table 1 show even stronger evidence of switching when the period from the late 1920's through the end of World War II is included.

III. Different Specifications of Switching

The nature of the switching which one finds in the data will depend on the economic forces which give rise to switching behaviour. Suppose, for example, that the source is time variation in the uncertainty of excess stock returns. If this is variation in the diversifiable component of returns, then mean returns might be the same across regimes while their volatility differs as in Schwert (1989). However, if the undiversifiable risk component is the source of this variation, then we might expect the high variance state to have higher mean returns. On the other hand, Black (1976) and Christie (1982) have suggested that the leverage effect might cause higher variances to be associated with lower average returns. These differences make it interesting to consider how the behaviour of mean returns and their variances are related across regimes.

A. Switching in Means

The first specification we examine is one in which stock market returns are drawn from two distributions which differ only in their means, as in equation (2). Empirical estimates of the specification with switching in means are presented in Table 2. The most striking feature of the estimates is the enormous difference in mean returns between the two regimes. In state 0, monthly mean excess returns are .0082, implying annual excess returns of about 10%. When state 1 occurs, stock prices drop by about 17% in a single month.[4] The estimates of the transition probabilities imply that q is about .985; thus, the probability of remaining in state 0 is very high. The situation is very different for state 1; here the probability that the state will persist for one more period is only about .264.

Figure 1 provides a visual interpretation of the results, showing how the probability of being in state 0 evolves over the sample. This probability is calculated using the two-sided filter described by Hamilton (1989). This is an ex post probability in the sense that it uses all available information up to the end of the sample in determining the classification probabilities at time t. The probability of being in state 0 is very close to 1 for most of the sample period. When the probability of being in state 0 deviates from 1, it typically does so for a very short period of time. This is refected in sharp spikes at irregular intervals. Several of these spikes correspond to well-known market crashes, such as October 1929 or October 1987. These spikes are particularly frequent in the early 1930's and again from about 1937 to 1940. There are only very small blips from 1940 until 1960, with spikes reappearing in the 1960's and 1970's and then again in 1987.

To summarize, the parameter estimates and the graph of ex post probabilities provide complementary pictures of the nature of regimes. There is a very persistent state in which excess returns are positive and a rare state characterized by major stock market crashes.

B. Switching in Variances

---

[4]In estimating this specification for the period 1946-87, Turner, Startz, and Nelson (1989) obtain similar estimates: monthly excess returns of .0064 in state 0 and -.2169 in state 1.

The second specification we examine is one in which stock market returns are drawn from two distributions which differ only in their variances, as in equation (4). Empirical estimates of specification (4) are presented in the second column of Table 2. The estimates of σ_0 and σ_1 show that the variance of returns is about three times as high in state 1 as in state 0. The estimates of the transition probabilities show that the low variance state is extremely persistent; the implied value of q is .991. In contrast to the specification with switching in means, state 1 is also quite persistent; the implied value of p is .941.

Figure 2 illustrates the probability of state 0. The picture is both similar to and different from Figure 1. As in Figure 1, the probability of state 0 is close to one for most of the sample period and deviates little from one from 1940 through 1960. Like the model with switching in means, most of the variation occurs in the early 1930's, the late 1930's, the 1970's, and around 1987. The periods in which the probability of state 0 is small look much less like rare spikes. In fact, for most of the early 1930's, the probability of state 0 is very small. Even in the mid 1970's and around 1987, the probability of state 0 shows less tendancy to jump from one to zero; visually, the "spikes" are more like "wedges."

To summarize, specification (4) yields a picture of regimes with sharply different variances. Both the low-variance and the high-variance states are persistent. There are two extended periods of high volatility (1929-33 and 1937-40) followed by isolated bursts of high volatility since then. There is a substantial overlap in the periods which are marked by regime shifts between the specifications with switching in means and switching in variances.

C. Switching in Means and Variances

The third specification we examine is one in which stock market returns are drawn from distributions which differ in both their means and variances, as in equation (5). Empirical results for this specification are presented in the third column of Table 2. State 1 is characterized by a variance which is about three times as large as the variance in state 0. The estimates of the transition probabilites are very similar to the specification with switching only in the variances. The mean return in state 1 is negative; if state 1 were to persist for a year, it would reduce stock prices by about 14%.

These results imply that the stock market is characterized by a state in which risk is relatively low and investors earn more than they would by holding T-bills and a state in which risk is substantially higher and investors lose money. Our result parallels recent findings by Brock, Lakonishok, and LeBaron (1992). They find that when certain trading rules give a "buy" signal, subsequent returns are positive on average, while when the trading rules give a "sell" signal, subsequent returns are negative on average.[5] As in our results, the negative returns are associated with a higher variance than the positive returns.

If investors believed that stock market returns were characterized by equation (5) and they were able to observe the current state, it would be hard to understand why they held equity in state 1. Turner, Startz, and Nelson (1989) offer an explanation for results of this type which is based on the idea that investors are unable to observe the current state and therefore must learn about the current state over time. It would be interesting to see whether their story could account for the results obtained by Brock, Lakonishok, and LeBaron (1992).

D. Diagnostic Tests

[5] Pesaran and Potter (1993) consider the types of unobservable stochastic discount factors and economic structures which are compatible with predictions of negative excess returns.

All of the foregoing specifications are based on the assumption that ε_t is an i.i.d. Gaussian variable. White (1987) suggests a series of diagnostic tests which can be used to assess the validity of this assumption.[6] Table 2a presents individual tests for serial correlation, ARCH(1), higher order Markov effects, and a joint test for all of these forms of misspecification. There is no evidence of serial correlation in any of the specifications.

The diagnostic tests show no evidence of omitted ARCH effects. This is particularly interesting because stock market returns show strong ARCH effects, as has been widely documented. This is consistent with the results of Hamilton and Susmel (1993), who find that Markov switching models do a better job of capturing stock market returns than previous ARCH or GARCH specifications.[7] A question for further research is how simple models of Markov switching (without autoregressive conditional heteroscedasticity) compare to ARCH models.

The specifications in this paper assume that stock market returns can adequately be described by a first order Markov chain.[8] In other words, they assume that the current state is a function only of the previous state and not a function of the state two periods or more ago. The test for higher order Markov effects is designed to check this assumption. Table 2a suggests that there is little evidence of higher order Markov effects.

In the next two sections, we consider a variety of other specifications. For reasons of

[6]See the appendix for a more technical description of the diagnostic tests.

[7]If the underlying data generating process reflects regime switching, this would induce a non-linearity in stock market returns which could show up in ARCH tests, since these can detect non-linear structure in time series.

[8]Note that a nth-order Markov model with two states can always be reparametrized as first-order Markov model with 2^n states. Therefore, these tests for higher-order Markovian effects can also be interpreted as tests for the appropriate number of states.

space, we do not report diagnostic tests for Tables 3 and 4, but the results are qualitatively similar. In no case do we find evidence of serial correlation, ARCH effects, or higher order Markov effects which is statistically significant at conventional levels. The joint tests also fail to reveal any evidence of misspecification.

IV. Multivariate Specifications

A number of recent studies have found evidence that stock market returns can be predicted using macroeconomic variables. These studies include Keim and Stambaugh (1986), Campbell and Shiller (1988), Fama and French (1988a), and Cutler, Poterba, and Summers (1991).[9] Several other studies have provided evidence that stock market returns do not follow a random walk and, more specifically, that future returns can be predicted on the basis of returns over the previous several years. These include Fama and French (1988b), Poterba and Summers (1988), and Jegadeesh (1991).[10] A natural question is whether, after controlling for switching, there is still evidence that stock market returns can be predicted using macroeconomic variables.

Fads models provide one possible economic motivation for Markov switching processes in which stock market returns are predictable. In a world with fads, Cutler, Poterba, and Summers (1991) show that returns will be correlated with lagged values of the price-dividend

---

[9]Several recent studies raise questions about the statistical strength of the evidence that returns are predictable. See, for example, Elliott and Stock (1992), Goetzmann and Jorion (1993), and Nelson and Kim (1993).

[10]Again, subsequent studies have examined the strength of the inferences that can be drawn from existing data. See, for example, Richardson and Stock (1989), Kim, Nelson, and Startz (1991), McQueen (1992), and Jog and Schaller (1993).

ratio.[11] If allowance is made for state-dependent heteroscedasticity, this yields a switching model in which the price-dividend ratio predicts returns.[12] The predictability of returns can also come from time variation in the risk premium in an efficient market.[13]

In the first column of Table 3, we present estimates of the following specification:

$$R_t = \alpha_0 + \beta_0 d_{t-1} + [\sigma_0(1-S_t) + \sigma_1 S_t]\epsilon_t \qquad \textbf{(6)}$$

where d_{t-1} is the log price-dividend ratio.[14] Equation (6) allows for switching in variances and the predictability of mean returns. To the best of our knowledge, we are the first to estimate specifications of this type on stock market returns.[15]

The empirical estimates provide stong evidence of predictability. The t-statistic for β_0 is close to three. In economic terms, a one-standard deviation increase in last period's log price-dividend ratio implies that stock market prices would fall about 5% per year. This is illustrated graphically in Figure 3. Expected monthly returns vary over a wide range.[16] In the depth of the Great Depression, when the price-dividend ratio was very low, expected excess returns were

---

[11]In their model, the price-dividend ratio is a proxy for the deviation of the actual price from the fundamental price.

[12]See van Norden and Schaller (1993).

[13]See Fama and French (1988) and Fama (1991) for further discussion.

[14]To remove the seasonal component from dividends, we follow Fama and French (1988a) in using the sum of dividends over the twelve months ending in period t-1. In addition, we normalize the log price-dividend ratio to have a mean of 0 and unit variance.

[15]See also Curcio and Shaw (1993).

[16]By expected returns, we mean $E[R_t|d_{t-1}]$ based on the parameter estimates in Table 3; i.e., $\alpha_0 + \beta_0 d_{t-1}$.

almost 2% per month (27% per year). At the height of the bull market in the late 1960's, when the price-dividend ratio was unusually high, expected excess returns were actually slightly negative.

A second economic motivation for a Markov switching process in which returns are predictable is regime switching in endowments. As noted above, such models have been proposed by Cecchetti, Lam, and Mark (1990) and Kandel and Stambaugh (1990) as a way of accounting for stock market anomalies such as mean reversion. van Norden and Schaller (1992) provide simulation evidence that regime switching in endowments in a Lucas asset pricing model will lead to regime switching in returns. The non-linear predictability of returns could also arise in a switching specification as a result of stochastic bubbles of the kind proposed by Blanchard and Watson (1982).

In the second column of Table 3, we present estimates of the following specification:

$$\begin{aligned} R_t = {} & \alpha_0(1-S_t) + \alpha_1 S_t + \beta_0 d_{t-1}(1-S_t) + \beta_1 d_{t-1} S_t \\ & + [\sigma_0(1-S_t) + \sigma_1 S_t]\epsilon_t \end{aligned} \qquad \textbf{(7)}$$

Equation (7) allows for switching in both means and variances. Unlike (6), it allows the effect of the price-dividend ratio on returns to be asymmetric.

The point estimates of β_0 and β_1 suggest economically important differences in the influence of the price-dividend ratio between the two regimes. In regime 1, a one-standard-deviation increase in the log price-dividend ratio implies a fall in stock market prices of about 21% per year. In regime 0, a similar increase in the price-dividend ratio implies a fall of only about 5% per year. The differences between α_0 and α_1 are also economically significant. Over one year, they imply stock market prices would be about 45% lower in state 1 than in state 0.

The statistical significance of the difference between α_0 and α_1 and β_0 and β_1 is weaker: a likelihood ratio test would reject the restrictions which (6) imposes on (7) at the .10 level.[17]

The point estimates in Table 3 suggest a highly asymmetric response to the lagged price-dividend ratio; the effect is four times larger in state 1 than state 0. A simple calculation shows that the price-dividend ratio must be extremely low (more than two standard deviations below its mean) before a risk-neutral investor would be indifferent between the states. Even then, state 1 is much riskier: the variance of ε is about three times higher in state 1. Another way of thinking about this is the following. State 1 is riskier and, for more than 97% of the values of the lagged price-dividend ratio, the expected return (conditional on d_{t-1} and S_t=1) is lower than the expected return in state 0.

The behaviour of expected monthly returns in the two regimes is plotted in Figure 4. This illustrates the point that expected returns in state 0 are always positive and almost always greater than expected returns in state 1, typically by economically large amounts. If the states were observable, this would be very strange. We noted above that Turner, Startz, and Nelson (1989) consider a model in which agents learn about the current state. It would be very interesting to explore whether an extension of their model could account for the results reported in Table 3.

Figure 4 has other interesting economic implications. In the depths of the Great Depression, the price-dividend ratio fell to extremely low levels. The parameter estimates in Table 3 therefore imply extremely high expected returns (almost 10% in a single month). In fact,

---

[17] We also estimated (7) with d lagged an additional period, as reported in the third column of Table 3. The point estimates of the parameters are similar, but the estimates are more precise. The likelihood ratio statistic for the restrictions $\alpha_0=\alpha_1$ and $\beta_0=\beta_1$ is 8.15, implying rejection at the .05 level.

the spike shown in 1932 was followed by one of the largest rallies in our sample.

V. Time Variation in Transition Probabilities

In this section, we examine whether the transition probabilities vary over time. In particular, we look at whether the transition probabilities are influenced by the price-dividend ratio. To the best of our knowledge, this paper is the first to examine whether economic variables influence the transition probabilities in a Markov switching specification of stock market returns.[18]

In addition to econometric novelty, there are two motivations for allowing transition probabilities to depend on the price-dividend ratio. First, the results presented above show that the price-dividend ratio appears to influence future returns. It is possible that this apparent predictive power arises because the price-dividend ratio is useful in determining the state in the next period, but that once this effect is taken into account, d_{t-1} has no influence on expected returns (conditional on the state). Second, the results recorded above show that regime 1 is a "bad" state for investors, since it involves lower returns and greater risk. It would be highly desirable for investors if they were able to predict that regime 1 would occur in the following period, based on currently available information.

A. Spurious Evidence of Returns Predictability?

The first question we examine is the true econometric role of the price-dividend ratio. Consider the following data generating process:

---

[18]Diebold, Lee, and Weinbach (1993) discuss the econometrics of time-varying transition probabilities. Filardo (1992) provides an application to data on industrial production.

$$R_t = \alpha_0(1-S_t) + \alpha_1 S_t + \beta_0 d_{t-1}(1-S_t) + \beta_1 d_{t-1} S_t + [\sigma_0(1-S_t) + \sigma_1 S_t]\epsilon_t \quad \textbf{(8)}$$

$$Pr(S_t=0|S_{t-1}=0) = q(d_{t-1})$$
$$Pr(S_t=1|S_{t-1}=1) = p(d_{t-1}) \quad \textbf{(9)}$$

$$q = \Phi(\gamma_{q0} + \gamma_{qd} d_{t-1})$$
$$p = \Phi(\gamma_{p0} + \gamma_{pd} d_{t-1}) \quad \textbf{(10)}$$

Now suppose that the true data generating process was (8)-(10) with $\beta_0=\beta_1=0$, but that we estimated the model with γ_{qd} and γ_{pd} constrained to zero.[19] In a complex non-linear specification, it is possible that by omitting d_{t-1} in the transition probabilities, we might induce correlation between R_t and d_{t-1} and thus find spurious evidence that β_0 and β_1 were non-zero. In other words, the price-dividend ratio might have predictive ability for the state but not for returns (conditional on the state).

To test whether the apparent predictability of returns is spurious, we estimate (8)-(10) first with $\beta_0=\beta_1=0$ and then with β_0 and β_1 unconstrained. We can then use a standard likelihood ratio test to determine whether the price-dividend ratio affects returns, even after allowing it to affect the transition probabilities. Columns one and two of Table 4 report results for $\beta_0=\beta_1=0$ and β_0 and β_1 unconstrained, respectively. The likelihood ratio statistic is 10.6, implying rejection of the hypothesis that $\beta_0=\beta_1=0$ at the .01 level. Thus, the predictive ability of the price-dividend ratio for returns cannot be attributed to the assumption of constant transition probabilities.

---

[19]Note that this is precisely the specification of equation (7) where we found evidence of predictability of returns, even after accounting for state-dependent switching.

Column three of Table 4 reports a specification which allows predictability of mean returns but restricts if to be symmetric. This corresponds to equation (6), combined with equations (9) and (10). It is interesting that allowing for time variation in the transition probabilities has little effect on the estimate of β_0. Without time variation in transition probabilities (Table 3, column one), the estimate of β_0 is -.0046; with time variation in transition probabilities (Table 4, column three), the estimate of β_0 is also -.0046. In both cases, the t-statistic is close to three. This provides further evidence that allowing the price-dividend ratio to influence the transition probabilities has little impact on the evidence that the price-dividend ratio has predictive ability for returns.

B. Is the "Bad" State Predictable?

We illustrate the effects of time-varying transition probabilities in Figure 5, which is based on the parameters in equations (8)-(10) as reported in the second column of Table 4. The probability of remaining in state 0 (shown by the solid line) is high throughout the sample. There is considerable variation in the persistence of state 1. The dashed line (which shows the time path of p) varies from a peak of nearly 1 in the depths of the Great Depression to a low of only about .5 in the early 1970's. There is considerable variation throughout the period from 1927 to 1989.

The probability of going from the "good" state to the "bad" state is 1-q. Since q is quite high throughout the sample, the scale of Figure 5 conceals considerable relative variation in the probability of a transition to the bad state. Figure 6 presents 1-q, again based on the estimates in column three of Table 4. In periods when the price-dividend ratio is high relative to its values in the previous several years (such as 1929 and 1987), the probability of a transition into the "bad" state is high relative to its values in the previous several years (e.g., 1927-28 and 1985-86). Conversely, when the price-dividend ratio is low (e.g., in the late 1970's), the probability of a transition into the "bad" state is relatively low.

It is noticeable that there are local peaks in the probability of a transition into the "bad" state just before the 1929 and 1987 crashes. To investigate this further, we compiled a list of the largest three-month losses in our sample.[20] Of the 10 distinct crashes we identified, 7 are preceded by rises in the probability of a transition into the "bad" state. These include crashes in 1929, 1937, 1946, 1962, 1970, 1975, and 1987.

We can rule out at least one possible explanation of why rises in the probability of a transition into the "bad" state might precede stock market crashes. The dips in the probabality of being in state 0 shown in Figures 1 and 2 might occur slightly in advance of actual crashes since they are produced by a 2-sided smoother.[21] This means that probability of being in regime 1 at any point in time is a function of all past and all future returns in the sample. Since regimes tend to be persistent, a crash at time *t* tends to raise *(ceteris paribus)* the probability of being in the crash regime at *t-1* and *t+1*.

This explanation does not apply to Figure 5 and 6 since the probabilities graphed there are not produced by any form of smoothing. The difference in construction reflects a difference in definition. The series in Figures 1 and 2 are $Pr(S_t=0 | (R_1, ..., R_T))$, whereas those in Figure

---

[20]We focussed on three-month losses (rather than one-month losses) to exclude transitory losses that are almost immediately offset by subsequent price increases, as well as to capture more gradual (but large) price decreases. We identified the 20 largest three-month losses in our sample, but this yielded only 10 distinct crashes, since many of the three-month losses either overlapped or occurred within the same year.

[21]See the Appendix for a description of the two-sided smoother.

5 are $Pr(S_t=0|S_{t-1}=0,d_{t-1})$ and $Pr(S_t=1|S_{t-1}=1,d_{t-1})$. In other words, the series in Figures 1 and 2 are the probabilities of being in a given state while those in Figure 5 are the probabilities of remaining in a state based on what the state was in the previous period. The latter are calculated using equation (13), and therefore depend only on information available at *t-1* (i.e., d_{t-1}) and the estimated parameters.

While the effect of the price-dividend ratio on the transition probabilities is provocative, it is not very precisely measured. The t-statistics for γ_{qd} and γ_{pd} are both less than two. Statistically speaking, the "bad" state is therefore hard to predict, at least using the price-dividend ratio. Whether other variables are better able to predict transitions into the "bad" state is a question we leave to future research.

VI. Summary and Conclusion

In this section we summarize some of our main empirical results. First, we find very strong evidence of switching behaviour in stock market returns. This evidence is robust to a variety of different specifications - switching in means, switching in variances, or switching in both means and variances. By applying new tests to the results previously reported by Turner, Startz, and Nelson (1989), we show that the hypothesis of no switching can be rejected in the period since World War II. The evidence of switching is much stronger, however, when we include the pre-war period.

When we allow for switching in means, we find two distinct states in the data. In state 0, there are excess returns of 82 basis points per month (about 10% per year). This state is highly persistent. In state 1, the stock market crashes, suffering losses of about 17% in a single month. This state is very transitory.

When we allow for switching in variances, we find that the variance of returns is about three times higher in the high-variance state than in the low-variance state. Both states are highly persistent, with the probability of remaining in each state well above 90%.

Figures 1 and 2 illustrate why it may be hard to distinguish between specifications with switching in means and switching in variances. The periods in which the probability of the high-return state are low include the early 1930's, the late 1930's, and 1987. Smaller dips in the probability of the high-return state occur in the 1960's and 1970's. To a substantial extent, these are the same periods in which the probability of the low-variance state is low.

When we allow for switching in both means and variances, we find that in one state, excess returns are high and the variance is low; in the other state, excess returns are low (in fact, negative) and the variance is high. This parallels a similarly surprising result documented by Brock, Lakonishok, and LeBaron (1992).[22]

The results described so far reinforce the earlier work of Turner, Startz, and Nelson (1989). A major innovation of our work is to use a multivariate specification for the Markov switching model. This allows us to examine whether the price-dividend ratio has marginal predictive power for stock market returns after accounting for state-dependent switching.

We find strong evidence of predictability. In the specification with Markov switching in variances, the coefficient on the lagged price-dividend ratio (in a regression where the dependent variable is current excess returns) has a t-statistic of about three. When we graph expected

---

[22]They find that certain trading rules predict two states. In the first (preceded by a "buy" signal), returns are positive on average and the variance is relatively low; in the second (preceded by a "sell" signal), returns are negative on average and the variance is relatively high.

returns (conditional on the past price-dividend ratio), we find that they vary from about 200 basis points per month during the Great Depression to a slightly negative value in the late 1960's.

The response of returns to the past price-dividend ratio is strongly asymmetric. In a specification where we allow for switching in both means and variances, we find that the effect of the price-dividend ratio is about four times larger in the low-return state than in the high-return state. This means that there is even greater variability in expected returns (conditional on both the price-dividend ratio and state 1); they vary from a gain of almost 1000 basis points per month in the Great Depression to a loss of about 500 basis points per month in the early 1970's. State 1 is riskier and, for more than 97% of the values of the lagged price-dividend ratio, the expected return is lower than the expected return in state 0.

A second important innovation in our work is to allow the probability of transition from one regime to another to depend on economic variables. Here again, we find an asymmetric response to the past price-dividend ratio. The price-dividend ratio has a qualitatively small effect on the probability of remaining in state 0 (the high-return, low-variance state).[23] The probability of remaining in state 1 (the low-return, high-variance state) varies substantially with the price-dividend ratio - from almost one during the Great Depression to about one-half in the early 1970's.

Stock market returns have been so extensively studied that it often seems difficult to add to our existing knowledge. By extending the techniques and results of Hamilton (1989) and Turner, Startz, and Nelson (1989), this paper highlights a number of new features of stock market

[23]The ex ante probability of switching from this state to the bad state shows an economically interesting pattern, however. It tends to rise shortly before several major stock market crashes.

returns. We invite empirical researchers to join in using these exciting new methods and invite asset pricing theorists to attempt to account for the stylized facts that are emerging from this work.

Appendix

This section gives additional information on the methodology used to estimate the models discussed in the body of the paper. All of the models fit within the basic framework of a Markov switching regression proposed by Goldfeld and Quandt (1973) and therefore can be written as special cases of the general model:

$$Y_t = X_{0t}\cdot\beta_0 + \epsilon_{0t} \Leftrightarrow S_t=0 \qquad \epsilon_{0t}\sim N(0,\sigma_0) \tag{A}$$

$$Y_t = X_{1t}\cdot\beta_1 + \epsilon_{1t} \Leftrightarrow S_t=1 \qquad \epsilon_{1t}\sim N(0,\sigma_1) \tag{B}$$

$$Pr(S_t=0|S_{t-1}=0) = \Phi(-X_{qt}\cdot\gamma_q) \tag{C}$$

$$Pr(S_t=1|S_{t-1}=1) = \Phi(-X_{pt}\cdot\gamma_p) \tag{D}$$

where Φ(x) is the standard Gaussian c.d.f. and X_{it} is a vector of explanatory variables. For example, all the univariate models discussed in sections II and III set $X_{0t} = X_{1t} = X_{qt} = X_{pt} = 1$. Aside from being multivariate in nature, these univariate models differ from that considered by Hamilton (1989) in two respects. First, Hamilton uses a logit function in the transition equations while we use a probit function. Second, Hamilton allows for additional autoregressive dynamics. The above framework can, of course, capture state-dependent persistence through the addition of lagged values of Y_t to the vectors of explanatory variables X_{it}. However, our diagnostic tests find little evidence of serial correlation in the residuals.

We jointly estimate the parameters of this model by maximizing the conditional log-likelihood function llf(β_1, β_2, γ_q, γ_p, σ_1, σ_2). The likelihood function itself is evaluated recursively using an updating formula analogous to that in Hamilton (1993). Specifically, we define

$$\phi_{it} = \frac{\phi\left(\frac{Y_t - X_{it}\beta_i}{\sigma_i}\right)}{\sigma_i} \qquad \textbf{(E)}$$

for i=0,1 where $\phi(x)$ is the standard Gaussian p.d.f.,

$$\Phi_{jt} = \Phi(-X_{jt} \cdot \gamma_j) \qquad \textbf{(F)}$$

for j=p,q and

$$\pi_t(\theta,S_0) = Pr(S_t=0|Y_t,\ldots,Y_1,X_t,\ldots X_1,\theta,S_0) \qquad \textbf{(G)}$$

where

$$\theta = [\beta_0',\beta_1',\gamma_p',\gamma_q',\sigma_0,\sigma_1]' \qquad \textbf{(H)}$$

The π_t's may be calculated using the recursive formula

$$\pi_t(\theta,S_0) = \frac{\lambda_t \cdot \phi_{0t}}{\gamma_t \cdot \phi_{0t} + (1-\gamma_t) \cdot \phi_{1t}} \qquad \textbf{(I)}$$

where

$$\lambda_t = \pi_{t-1}(\theta,S_0) \cdot \Phi_{pt-1} + (1-\pi_{t-1}(\theta,S_0)) \cdot (1-\Phi_{qt-1}) \qquad \textbf{(J)}$$

The probability density at each point in time (conditional on θ and S_0) is therefore

$$\omega_t(S_0,\theta) = \pi_t(\theta,S_0)\cdot\phi_{0t}+(1-\pi_t(\theta,S_0))\cdot\phi_{1t} \tag{K}$$

and that of the sample 1,...,T is

$$\prod_{j=1}^{T}\omega_j(S_0,\theta) \tag{L}$$

The log likelihood function is just the log of Equation (L) and is in general conditional on S_0. As noted by Hamilton (1993), the model can therefore be estimated by treating $\pi_0(\theta,S_0)$ as an additional parameter over which to maximize, or by setting it to some predetermined value, such as the unconditional probability of being in state 0. We take the latter approach, setting

$$\pi_0(\theta,S_0) = \frac{1-\Phi_q}{2-\Phi_p-\Phi_q} \tag{M}$$

where Φ_i is the mean of Φ_{it} over t=1,...,T.

The maximization is performed using numerical gradients and the *maxlik* procedure in GAUSS, which uses a variety of algorithms. We found that results were insensitive to the optimization algorithm chosen. The covariance matrix of the parameters was always calculated by a final iteration with a Newton algorithm. We do not report the robust standard-errors available in *maxlik* since they are not generally valid for this class of models. (Validity requires that the model be consistently estimated in the presence of heteroscedasticity, which is not typically the case.) However, an inspection of these estimates found that they were generally similar to the estimates that we report.

We also calculate the full-sample smoothed probabilities of being in regime 0. In contrast to $\pi_t(\theta, S_0)$, which conditions only on information over the interval [1,t], the full-sample smoothed probabilities use information over the interval [1,T], so may be defined as

$$\bar{\pi}_t(\theta,S_0) = Pr(S_t=0|Y_T,...,Y_1,X_T,...,X_1,\theta,S_0) \qquad \textbf{(N)}$$

It is calculated using the formula

$$\bar{\pi}_t(\theta,S_0) = \pi_t(\theta,S_0) \cdot \prod_{\tau=t+1}^{T} F_\tau \qquad \textbf{(O)}$$

where

$$F_\tau \equiv \frac{\pi_\tau(\theta,S_t) \cdot \phi_{0\tau} + (1-\pi_\tau(\theta,S_t)) \cdot \phi_{1\tau}}{\pi_\tau(\theta,S_0) \cdot \phi_{0\tau} + (1-\pi_\tau(\theta,S_0)) \cdot \phi_{1\tau}} \qquad \textbf{(P)}$$

and we assume that S_t in Equation (P) is equal to 0. Calculating the smoothed probability at any time t therefore requires that we work out the conditional density for each observation τ from t to T and then go back and re-evaluate it for each τ under the assumption that $S_t = 0$.

Table 2a reports the results of various score-based diagnostic statistics for this class of models. These are based on our adaptation of tests suggested by White (1987) and Hamilton (1990) to the switching regression context. Let $llf_t(\theta)$ denote the log-likelihood function of a given observation Y_t. The score $h_t(\theta)$ is simply the gradient of $llf_t(\theta)$ with respect to θ. Subject to regularity conditions, it can be shown that at the true parameter estimates θ_0, $h_t(\theta_0)$ should be unforecastable on the basis of any information available at t-1, which includes $h_{t-1}(\theta_0)$. Since $h_t(\theta)$ is a vector with the same dimensions as θ (say, mx1), this implies that the absence of any first-order serial correlation in $h_t(\theta_0)$ alone gives us mxm testable restrictions.

While in theory, we could test all of these conditions, or test even higher order restrictions on its serial correlation, in practice, most of these restrictions are very difficult to interpret. In what follows, attention will be limited to those restrictions that are most easily understood.

White (1987) constructs the general test by listing those ℓ elements of the mxm matrix $h_t(\theta)\cdot h_t(\theta)'$ that we wish to test in the ℓx1 vector $c_t(\theta)$. He then lets $\hat{\theta}$ denote our maximum-likelihood estimate of θ, and lets $\hat{A}$ be the (2,2) sub-block of the *inverse* of the partitioned matrix

$$T^{-1}\cdot\begin{bmatrix}\sum_{t=1}^{T} h_t(\hat{\theta})\cdot h_t(\hat{\theta})' & \sum_{t=1}^{T} h_t(\hat{\theta})\cdot c_t(\hat{\theta})' \\ \sum_{t=1}^{T} c_t(\hat{\theta})\cdot h_t(\hat{\theta})' & \sum_{t=1}^{T} c_t(\hat{\theta})\cdot c_t(\hat{\theta})'\end{bmatrix} \qquad \textbf{(Q)}$$

where T is the sample size. In this case, he shows that if the model is correctly specified, the matrix product

$$T^{-1}\cdot\left[\sum_{t=1}^{T} c_t(\hat{\theta})\right]'\cdot\hat{A}\cdot\left[\sum_{t=1}^{T} c_t(\hat{\theta})\right] \qquad \textbf{(R)}$$

will have a $\chi^2(l)$ asymptotic distribution.

If we assume that X_{0t} and X_{1t} each contain a constant, then testing whether each of these gradients are autocorrelated amounts to testing whether there is omitted serial correlation in ε_{t1} or ε_{t2}. The intuition here is that serial correlation in these gradients implies that we tend to find "runs" where the constant should be higher or lower, which in this context implies persistence in the residuals, or serial correlation. Testing the gradient of the constant in X_{pt} or X_{qt} gives a test for neglected Markov-like switching effects, since serial correlation here implies that regimes seem to be more persistent than indicated. Finally, tests of the gradients of σ_S and σ_C amount to tests for first-order regime-specific ARCH effects, since persistence here implies that the volatility in each regime seems to vary over time in a way captured by a first-order autoregression. Note that if we restrict either the intercepts or the variances to be the same in both regimes, then there will be only a single gradient vector. Its single corresponding diagnostic test therefore tests both regimes simultaneously for evidence of either serial correlation or ARCH effects.

References

Black, F. (1976), "Studies of Stock Market Volatility Changes," 1976 Proceedings of the American Statistical Association, Business and Economic Statistics Section, 177-181.

Blanchard, Olivier J. and Mark Watson (1982), "Bubbles, Rational Expectations and Financial Markets," in P. Watchel (ed.), Crises in the Economic and Financial Structure, Lexington Books, Lexington, MA.

Brock, William, Josef Lakonishok, and Blake LeBaron (1992), "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," Journal of Finance, Vol. 47, No. 5, 1731-1764.

Campbell, John Y. and Robert J. Shiller (1988), "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," Review of Financial Studies, Vol. 1, 195-228.

Cecchetti, Stephen G., Pok-Sang Lam, and Nelson C. Mark (1990), "Mean Reversion in Equilibrium Asset Prices," American Economic Review, Vol. 80, No. 3, 398-418.

Christie, A.A. (1982), "The Stochastic Behavior of Common Stock Variances: Value, Leverage, and Interest Rate Effects," Journal of Financial Economics, 10, 407-432.

Curcio, Riccardo and Mushtaq Shah (1993), "Business Conditions and Stock Returns: A Markov Switching Regression Model for the U.S. Stock Market," London School of Economics mimeo.

Cutler, David, James M. Poterba, and Lawrence H. Summers (1991), "Speculative Dynamics," Review of Economic Studies, 58, 529-546.

Diebold, F.X., Joon-Haeng Lee, and Gretchen C. Weinbach (1993), "Regime Switching with Time-Varying Transition Probabilities," Nonstationary Time Series Analysis and Cointegration, ed. C. Hargreaves, (Ser. Advanced Texts in Econometrics, eds. C.W.J. Granger and G. Mizon), Oxford: Oxford University Press.

Elliott, Graham and James H. Stock (1992), "Inference in Time Series Regression when the Order of Integration of a Regressor is Unknown," NBER Technical Working Paper No. 122.

Fama, Eugene F. (1991), "Efficient Capital Markets: II," The Journal of Finance, Vol. XLVI, No. 5, December 1991, 1575-1617.

Fama, Eugene F. and Kenneth R. French (1988a), "Dividend Yields and Expected Stock Returns," Journal of Finacial Economics, Vol. 22, No. 1, 3-25.

Fama, Eugene F. and Kenneth R. French (1988b), "Permanent and Temporary Components of Stock Prices," Journal of Political Economy, April 1988, 98, 246-74.

Filardo, Andrew J. (1992), "Business Cycle Phases and Their Transitions," Federal Reserve Bank of Kansas City.

Garcia, Rene (1992), "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," University of Montreal.

Goetzmann, William N. and Philippe Jorion (1993), "Testing the Predictive Power of Dividend Yields," Journal of Finance, Vol. 48, No. 2, 663-680.

Hall, Stephen, and Martin Sola (1993), "Testing for Collapsing Bubbles: An Endogenous Switching ADF Test," London Business School Discussion Paper No. DP 15-93.

Hamilton, James D. (1993) Time Series Analysis: Chapter 22 - Modeling Time Series with Changes in Regime, University of California at San Diego mimeo.

Hamilton, James D. (1990), "Specification Testing in Markov-Switching Time Series Models," University of Virginia Department of Economics Discussion Paper, No. 209.

Hamilton, James B. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," Econometrica, Vol. 57, No. 2, 357-384.

Hansen, B. E. (1992), "The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP," Journal of Applied Econometrics, Vol. 7,S61-S82.

Hansen, Bruce E. (1993), "Inference when a Nuisance Parameter is not Identified under the Null Hypothesis," University of Rochester, mimeo.

Jegadeesh, Narasimhan (1991), "Seasonality in Stock Price Mean Reversion: Evidence from the U.S. and the U. K.", The Journal of Finance, Vol. XLVI, No. 4, September 1991, 1427-1444

Jog, Vijay and Huntley Schaller (1993), "A Panel Data Test for Mean Reversion Using Randomization", Paper presented at the North American Winter Meetings of the Econometric Society, Anaheim, January, 1993.

Kandel, S. and R.F. Stambaugh (1990), "Expectations and Volatility of Consumption and Asset Returns," Review of Financial Studies, 3, 207-232.

Keim, Donald B. and Robert F. Stambaugh (1986), "Predicting Returns in the Stock and Bond Markets," Journal of Financial Economics 17, 357-390.

Kim, Myung Jig, Charles R. Nelson, and Richard Startz (1991), "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence," Review of Economic Studies, 515-528, May.

McQueen, Grant (1992), "Long-Horizon Mean-Reverting Stock Prices Revisited," Journal of Financial and Quantitative Analysis, Vol. 27, No. 1, March, 1992, 1-17.

Nelson, Charles R. and Myung J. Kim (1993), "Predictable Stock Returns: The Role of Small Sample Bias," The Journal of Finance, Vol. 48, No. 2, 641-662.

van Norden, Simon, and Huntley Schaller (1992), "Speculative Behaviour, Regime-Switching, and Stock Market Fundamentals," paper presented to the North American Summer Meetings of the Econometric Society, Seattle.

van Norden, Simon, and Huntley Schaller (1993a), "The Predictability of Stock Market Regime: Evidence From the Toronto Stock Exchange," Review of Economics and Statistics, forthcoming.

van Norden, Simon, and Huntley Schaller (1993b), "Fads or Bubbles?" paper presented to the SEDC meetings, Anaheim.

Poterba, James M., and Lawrence H. Summers (1988), "Mean Reversion in Stock Prices: Evidence and Implications," Journal of Financial Economics, Vol. 22, No. 1, 27-59.

Richardson, Matthew and James H. Stock (1989), "Drawing Inferences From Statistics Based on Multi-year Asset Returns," Journal of Financial Economics, 25, 323-348.

Schwert, G. William (1989), "Business Cycles, Financial Crises, and Stock Volatility," Carnegie-Rochester Conference Series on Public Policy, Vol. 31, 83-126.

Turner, Christopher M., Richard Startz, and Charles R. Nelson (1989), "A Markov Model of Heteroscedasticity, Riskm and Learning in the Stock Market," Journal of Financial Economics, Vol. 25, 3-22.

White, Halbert (1987), "Specification Testing in Dynamic Models," in Advances in Econometrics, Fifth World Congress, Volume II, Truman F. Bewley (Ed.) Cambridge, Cambridge Press.

Table 1

Tests for Regime-Switching in Stock Market Returns

	No Switching	Switching in Means (2)	Switching in Variances (4)	Switching in Both Means and Variances (5)
	1927-89			
Log Likelihood	1072.218	1119.772	1228.545	1230.080
Likelihood Ratio		95.108	312.654	315.724
	1946-87			
Likelihood Ratio		15.8	28.18	30.84

This table reports tests of the null hypothesis of no switching in stock market returns against three alternative specifications which involve switching in stock market returns. The numbers in parentheses at the top of each column refer to the equation numbers for these specifications. The 5% and 1% critical values for this non-standard distribution are 10.34 and 13.81 (for switching in means) and 13.52 and 17.67 (for switching in both means and variances), as tabulated in Garcia (1992). The data for the upper panel are CRSP value-weighted monthly returns over the period January 1927 to December 1989; these are expressed as excess returns by subtracting the rate of return on T-bills. The lower panel shows comparable statistics for the post-war period, as reported in Turner, Startz, and Nelson (1989).

Table 2

Univariate Specifications of Regime-Switching

	Switching in Means (2)	Switching in Variances (4)	Switching in Both Means and Variances (5)
α_0	0.0082 (3.71)	0.0071 (4.54)	0.0077 (4.75)
α_1	-0.1705 (-5.05)		-0.0129 (-1.15)
$q=\Phi(\ )$	2.1872 (10.61)	2.3573 (12.18)	2.3371 (11.75)
$p=\Phi(\ )$	-0.5980 (-1.30)	1.5642 (5.98)	1.5494 (5.82)
σ_0	0.0515 (31.50)	0.0392 (33.18)	0.0390 (32.17)
σ_1		0.1180 (12.44)	0.1155 (12.39)

This table reports parameter estimates from three univariate specifications of switching in stock market returns - equations (2), (4), and (5), respectively. The mean return is α_0 in regime 0 and α_1 in regime 1. The standard deviation of returns is σ_0 in regime 0 and σ_1 in regime 1. The transition probabilities are reported in rows three and four as the number which gives q or p when used as the argument of a standard Gausian cumulative distribution function. The figures in parentheses are t-ratios. Results are based on CRSP value-weighted monthly returns over the period January 1927 to December 1989; these are expressed as excess returns by subtracting the rate of return on T-bills.

Table 2a

Diagnostic Tests of the Markov Switching Specification

	Switching in Means (2)	Switching in Variances (4)	Switching in Means and Variances (5)
AR(1)	1.61 (0.205)	1.98 (0.160)	1.69 (0.193)
AR(1)	2.14 (0.143)	1.68 (0.195)	1.43 (0.231)
ARCH(1)	2.93 (0.087)	1.88 (0.170)	2.04 (0.153)
ARCH(1)		0.92 (0.336)	1.19 (0.275)
Higher Order Markov Effects	0.00 (0.971)	0.16 (0.691)	0.11 (0.742)
Higher Order Markov Effects	0.02 (0.898)	1.87 (0.172)	1.61 (0.204)
Joint	5.65 (0.342)	8.28 (0.218)	7.72 (0.260)

The "AR(1)" statistic tests for serial correlation in the residuals, "ARCH(1)" for serial correlation in volatility, and "higher-order Markov effects" for evidence that a first-order Markov chain is insufficient to capture the stochastic process for returns. (Technical details are described in the appendix.) Where there is a pair of tests, the upper row refers to regime 0 and the lower row to regime 1. Marginal significance levels are shown in parentheses. All of the test statistics are distributed χ^2 with one degree of freedom except the joint test which is distributed χ^2 with five degrees of freedom. There is only one ARCH(1) test for the switching in means specification because there is only one variance parameter.

Table 3

The Price-Dividend Ratio and Conditional Returns

	Switching in Variances (6)	Switching in Both Means and Variances (d_{t-1}) (7)	Switching in Both Means and Variances (d_{t-2}) (7)
α_0	0.0074 (4.762)	0.0080 (5.043)	0.0080 (5.055)
β_0	-0.0046 (-2.996)	-0.0046 (-2.917)	-0.0046 (-2.643)
α_1		-0.0232 (-1.657)	-0.0250 (-1.936)
β_1		-0.0190 (-1.475)	-0.0209 (-2.417)
γ_{q0}	2.3614 (12.305)	2.3399 (11.791)	2.3413 (11.795)
γ_{p0}	1.5581 (5.913)	1.5354 (5.641)	1.5398 (5.672)
σ_o	0.0390 (33.756)	0.0389 (32.875)	0.0389 (32.851)
σ_1	0.1189 (12.318)	0.1156 (12.336)	0.1133 (12.473)
Log Likelihood Function	1233.006	1235.523	1234.425

In this table, we allow returns to be influenced by the log price-dividend ratio (d). The specification for column 1 is equation (6), in which all returns have the same mean (conditional on the price-dividend ratio in the previous period). The specification for columns 2 and 3 is equation (7), in which returns may have different means (conditional on d_{t-1} or d_{t-2}, respectively). In column two, the expected return (conditional on d_{t-1}) is $\alpha_0+\beta_0 d_{t-1}$ in regime 0 and $\alpha_1+\beta_1 d_{t-1}$ in regime 1. Expected returns (conditional on d_{t-2}) are defined similarly for column three. The standard deviation of returns is σ_0 in regime 0 and σ_1 in regime 1. The transition probabilities are reported in rows three and four as the number which gives q or p when used as the arguement of a standard Gausian cumulative distribution function. The figures in parentheses are t-ratios.

Table 4

The Price-Dividend Ratio and Transition Probabilities

	Univariate Switching in Means & Variances (5),(9),(10)	Predictable Returns with Switching in Means (8)-(10)	Predictable Returns (6),(9),(10)
α_0	0.0078 (4.816)	0.0081 (5.059)	0.0074 (4.767)
β_0		-0.0044 (-2.772)	-0.0046 (-2.937)
α_1	-0.0144 (-1.227)	-0.0257 (-1.634)	
β_1		-0.0207 (-1.488)	
γ_{q0}	2.2629 (9.784)	2.2697 (9.959)	2.3043 (10.678)
γ_{qd}	-0.2357 (-1.310)	-0.1919 (-1.070)	-0.1617 (-0.992)
γ_{p0}	1.2245 (3.498)	1.1978 (3.233)	1.2680 (3.777)
γ_{pd}	-0.5277 (-1.652)	-0.5268 (-1.597)	-0.5063 (-1.648)
σ_0	0.0388 (30.679)	0.0388 (31.718)	0.0389 (33.027)
σ_1	0.1166 (11.542)	0.1172 (11.516)	0.1205 (11.802)
Log Likelihood Function	1231.970	1237.254	1234.578

In this table, we allow the transition probabilities to depend on the price-dividend ratio (d). The specification for column 1 is equations (8)-(10). The specification for column 2 is equations (11)-(13); this allows expected returns to depend on this price-dividend ratio. The specification for column 3 is equations (14)-(16); this allows expected returns to depend asymmetrically on d_{t-1}. The figures in parentheses are t-ratios. Results are based on CRSP value-weighted monthly returns over the period January 1927 to December 1989; these are expressed as excess returns by subtracting the rate of return on T-bills.