Identifying Bull and Bear Markets in Stock Returns

John M. Maheu
University of Alberta
Dept. of Economics
john.maheu@ualberta.ca

Thomas H. McCurdy
University of Toronto
Rotman School of Management
tmccurdy@mgmt.utoronto.ca

forthcoming
Journal of Business & Economic Statistics
January 2000, vol 18. No.1

Abstract

This paper uses a Markov-switching model which incorporates duration dependence to capture nonlinear structure in both the conditional mean and variance of stock returns. The model sorts returns into a high return stable state and a low return volatile state. We label these as bull and bear markets respectively. The filter identifies all major stock market downturns in over 160 years of monthly data. Bull markets have a declining hazard function although the best market gains come at the start of a bull market. Volatility increases with duration in bear markets. Allowing volatility to vary with duration captures volatility clustering.

Key Words: regime switching, duration dependence, filter, Markov chain

1 INTRODUCTION

Characterizing the dynamics of stock returns has been a particularly active area of research in the past decades. Empirical finance has had considerable success in capturing volatility dependence, while complicated dynamics in the conditional return are often ignored or assumed to be absent. However, asset pricing theory suggests that dependence in expected returns from a time-varying risk premium, stochastic rational bubbles, fads, learning about regimes, and irrational behavior by investors will show up in a nonlinear fashion. Finding this structure in conditional returns has been difficult. The purpose of this paper is to develop a framework to investigate nonlinear dependence jointly in the conditional mean and conditional variance of stock market returns.

A flexible class of models for capturing general nonlinear structure is a discrete mixture of distributions. A significant characteristic of these models is that they allow for a time-varying conditional mean and variance, while the unconditional distribution can have skewness and fat tails. One popular parameterization is the Markov-switching model of Hamilton (1989). This approach sorts data endogenously into regimes. Moreover, in such latent variable models we need not assume that the econometrician's information set coincides with agents' information. This is an attractive aspect of the model and departs from the traditional approach that assumes the information sets of the econometrician and market participants are identical. The Markov switching model has been shown by Rydén, Teräsvirta, and Åsbrink (1998) to be well suited to explaining the temporal and distributional properties of stock returns.

The Markov-switching model has been used extensively in modelling nonlinear structure in time-series data. For example, Turner, Startz, and Nelson (1989) use the model to account for a time-varying risk premium in stock returns, while Schaller and van Norden (1997) use the approach to distinguish between fads and bubbles in stock returns. Hamilton and Lin (1996) use the model to capture the nonlinear dynamics in the stock market and business cycle. Gordon and St-Amour (1999) model risk aversion as a two-state Markov process in their description of the cyclical pattern of asset prices.

Hamilton's (1989) first-order Markov model will not capture duration dependence in states. The latter could be particularly important in explaining volatility clustering, mean reversion and nonlinear cyclical features in returns. Ignoring duration dependence could result in a failure to capture important properties of stock returns. Durland and McCurdy (1994) developed a parsimonious implementation of a higher-order Markov chain which allowed state transition probabilities to be duration dependent. In that model, duration influenced the conditional mean through the hazard functions. That is, duration determined the persistence of the state-specific conditional mean by influencing when we switch states.

In this paper, in order to investigate potential nonlinear structure in stock market cycles, we extend the Durland and McCurdy (1994) model in several ways. In addition to duration-dependent hazards, duration also enters directly as a conditioning variable in both the mean and variance. Now, given persistence in a particular state, the conditional mean and variance can change with duration. This allows us to investigate dynamic behavior for the mean and variance within each state. In addition to revealing some interesting state-specific path dependence, this model also captures ARCH effects. In other words, according to our specification tests, conditional heteroskedasticity remaining in the Durland and McCurdy (1994) parameterization is fully explained by the endogenous duration variable in the conditional variance function.

Estimates of our two-state duration-dependent models classify the states into bull and bear markets. The bull market displays high returns coupled with low volatility, while the bear market has a low return and high volatility. Our empirical results find declining hazard functions (negative duration dependence) in both the bull and bear market states using monthly data from 1834-1995. This means the probability of switching out of the state declines with duration in that state. Despite the declining hazards, the best market gains come at the start of a bull market. That is, returns in the bull-market state are a decreasing function of duration. However, volatility in the bear-market state is an increasing function of duration.

Although the primary objective of this paper is to investigate duration dependence as a source of nonlinearity in stock market cycles, the structure of our econometric implementation is amenable to testing for specific sources of duration dependence by evaluating particular restrictions. Preliminary results suggest that the source of duration dependence in the transition matrix originates from both the conditional mean and conditional variance process. We also discuss several possible explanations of the duration dependence. For example, as the bull market persists, investors could become more optimistic about the future and hence wish to invest more in the stock market. This positive feedback means the probability of switching out of the bull market decreases with duration. This declining hazard could perhaps be interpreted as a momentum effect.

Duration dependence has also been related to the stochastic bubble explanation for nonlinear returns. For example, McQueen and Thorley (1994) and Cochran and Defina (1995) investigate duration dependence in stock market data by sorting the data into regimes, accounting for censoring and estimating a parametric hazard function. McQueen and Thorley (1994) show that a testable implication of stochastic rational bubbles is that high returns will exhibit negative duration dependence (declining hazard). That is, the probability of observing the end of a run of high returns will decline with duration. Our duration-dependent Markov-switching model is different from these approaches in that it sorts returns probabilistically into the alternative states. Although we also find a declining hazard associated with the bull-market state, on balance our results do not appear to fully support a bubble explanation for duration dependence in our sample of monthly stock returns. For example, we find a negative relationship between duration and the conditional return in the bull market which is the opposite of what one would expect for a rational bubble.

This paper is organized as follows: Section 2 outlines our general framework for modeling duration dependence. Section 3 describes the data, while Section 4 presents the model estimates. The sensitivity of these results to the model and various sample periods is considered in Section 5. Section 6 discusses the characteristics of bull and bear markets and possible sources of duration dependence. Conclusions are summarized in Section 7. Finally, the Appendix briefly outlines estimation of the models.

2 MODEL DEVELOPMENT

The nonlinearities we wish to explore in this paper include asymmetric cycles and time variation in the conditional moments of stock returns. A great deal of empirical work in finance focuses on nonlinear structure in the variance. Relatively little attention has been applied to nonlinear effects in the conditional return. To explore both moments in a nonlinear direction, a Markov-switching model is a natural choice. A Markov-switching model can produce skewness and excess kurtosis, as is often found in financial data. It also provides inference concerning the probability of being in a particular unobservable regime or state over time.

In the Markov-switching class of models the unobserved state is governed by a state variable S_t that takes on a finite number of values. Our family of models have two states and allow for a rich set of intra-state dynamics for the conditional mean and variance. Alternatively, more states could be used with simpler dynamics in each state. Our two-state specification accords with a tradition of classifying stock markets into two regimes labelled bull and bear markets.

We begin with a simple AR(l) model allowing two states for the conditional mean and variance. That is,

$$R_t = \mu(S_t) + \sum_{i=1}^{l} \phi_i (R_{t-i} - \mu(S_{t-i})) + \sigma(S_t)\nu_t \quad (1)$$

$$\nu_t \sim NID(0,1) \quad (2)$$

where R_t is the total return, l is the number of lags, and $\mu(S_t)$ is the conditional mean. Conditional on $\{S_{t-1}, S_{t-2}, \ldots, S_1, S_0, \ldots, S_{-\tau+1}\}$, where τ is the memory of duration dependence, S_t is assumed to be independent of $\{R_{t-1}, R_{t-2}, \ldots, R_1\}$. ν_t is an independent standard normal innovation and therefore independent of past R_t and S_t. In the model labelled DDMS-2, both the conditional mean and variance are allowed to take on two values. A simpler constant variance model (DDMS-1), imposes the restriction $\sigma(1) = \sigma(2)$, in equation (1).

Typically the evolution of S_t is governed by a first-order Markov chain. Putting duration dependence in the model results in a higher-order Markov chain. A general model would be intractable due to the loss of degrees of freedom needed to parameterize the transition matrix. Instead, Durland and McCurdy introduce duration using a parsimonious parameterization that puts nonlinear restrictions on a higher-order transition matrix. This model is particularly suited to capturing nonlinear structure and asymmetries from cyclical behavior such as mean reversion.

Duration could be important in capturing volatility clustering, but also could have explanatory power for the conditional return. The Durland and McCurdy duration-dependent Markov-switching (DDMS) model allows duration to affect the transition probabilities. In this section we will extend the model to allow for ARCH and to allow duration to be a conditioning variable in the conditional mean and variance.

Define duration as:

$$D(S_t) = \begin{cases} D(S_{t-1}) + 1 & \text{if } S_t = S_{t-1} \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

In words, duration is the length of a run of consecutive states S_t. Theoretically $D(S_t)$ could grow very large. To make estimation feasible we keep track of duration up to and including τ. As will be discussed further below, the parameter τ was chosen to maximize the loglikelihood starting from $\tau_{min} = l + 1$.

The transition probabilities are parameterized using the logistic function. This ensures that the probabilities are in (0,1). Using i and d to index realizations of states and duration, the transitions probabilities are, for i=1,2 and parameters $\gamma_1(i)$ and $\gamma_2(i)$,

$$P(S_t = i | S_{t-1} = i, D(S_{t-1}) = d) = \frac{\exp(\gamma_1(i) + \gamma_2(i)d)}{1 + \exp(\gamma_1(i) + \gamma_2(i)d)} \tag{4}$$

for $d \leq \tau$ and,

$$P(S_t = i | S_{t-1} = i, D(S_{t-1}) = d) = \frac{\exp(\gamma_1(i) + \gamma_2(i)\tau)}{1 + \exp(\gamma_1(i) + \gamma_2(i)\tau)} \tag{5}$$

for $d > \tau$. That is, duration is allowed to affect the transition probabilities up to τ periods, after which the transition probabilities are constant.

To discuss duration effects it is useful to define the hazard function. The hazard function is the conditional probability of a state change given the state has achieved a duration d. In terms of the transition probabilities the hazard function is,

$$1 - P(S_t = i | S_{t-1} = i, D(S_{t-1}) = d) \tag{6}$$

$$= \frac{1}{1 + \exp(\gamma_1(i) + \gamma_2(i)d)}, \qquad i = 1, 2. \tag{7}$$

A decreasing hazard function is referred to as negative duration dependence while an increasing hazard function is positive duration dependence. The effect of duration on the hazard function is uniquely summarized by the parameters $\gamma_2(i) \quad i = 1, 2$. In particular, for state i, $\gamma_2(i) < 0$ implies positive duration dependence, $\gamma_2(i) = 0$ implies no duration effect and $\gamma_2(i) > 0$ implies negative duration dependence. For example, if state 2 displays negative duration dependence and the stock market persists in state 2, then the probability of staying in state 2 increases.

Construction of the loglikelihood is briefly outlined in the Appendix. Asymptotic theory for Markov-switching models has only recently been developed. Leroux (1992) has proven consistency for hidden Markov models while Bickel, Ritov, and Rydén (1998) have proven asymptotic normality. Krishnamurthy and Rydén (1998) extend the consistency results to non-linear autoregressive models with Markov regime switching. A global maximum does not exist for Markov-switching models. For each model, 100 random starting values were used as startup values and the maximum loglikelihood value was taken as the model estimate. In all cases the model estimates were robust to different starting values. Estimation was carried out using the Fortran quasi-newton optimization routine E04JBF, from the Numerical Algorithms Group (NAG).

A defining feature of the Markov-switching model is the filter. The filter results from constructing the loglikelihood, and provides inference about the unobservable state variable S_t at time t. In a similar fashion, the filter is constructed for the duration-dependent Markov-switching model. However, the filter is now the joint probability of the state S_t and the duration $D(S_t)$ based on time t information.

Due to the latent variable S_t, residuals are unobservable. Using equation (1) the standardized expected residuals can be constructed using the filter as,

$$\sum_{s_t,\ldots,s_{t-l},d} \frac{R_t - E[R_t|s_t,\ldots,s_{t-l},d,Y_{t-1}]}{\sigma(s_t)} P(s_t,\ldots,s_{t-l},d|Y_{t-1}) \tag{8}$$

where $Y_{t-1} = \{R_{t-1}, R_{t-2}\ldots, R_1\}$. In equation (8) and the following, for notational convenience $s_t,\ldots,s_{t-l},d$, denotes realizations of the state variables $S_t,\ldots,S_{t-l},D(S_t)$. The standardized residuals can be used in residual-based diagnostic tests, such as the Ljung and Box (1978) (LB) test. However, since we do not know the asymptotic distribution of the LB statistic using the standardized expected residuals, specification tests should be interpreted with caution.

Financial data frequently displays ARCH effects. Because a two-state Markov model of variance may not be adequate to completely capture conditional heteroskedasticity, we extend the duration-dependent Markov-switching (DDMS) model to allow for ARCH.

Both Hamilton and Susmel (1994) and Cai (1994) extend the first-order Markov-switching model to include ARCH. A GARCH extension is more difficult because the conditional variance depends on the entire past history of S_t, making construction of the loglikelihood a computational burden. Dueker (1997) proposes an approximation to make GARCH feasible. However, Gray (1996) defines a slightly different class of GARCH models that overcomes the path dependence problem. This is the approach followed in this paper to implement ARCH. For parsimony we allowed only the intercept in the conditional variance to be state dependent.

Adding ARCH(K) gives the DDMS-ARCH model:

$$\begin{aligned} R_t &= \mu(S_t) + \sum_{i=1}^{l} \phi_i (R_{t-i} - \mu(S_{t-i})) + \epsilon_t & (9) \\ h_t(S_t) &= \omega(S_t) + \sum_{k=1}^{K} \alpha_k \tilde{\epsilon}_{t-k}^2 & (10) \\ \epsilon_t &= \sqrt{h_t(S_t)}\nu_t & (11) \\ \nu_t &\sim NID(0,1) & (12) \end{aligned}$$

and where

$$\tilde{\epsilon}_{t-k} = R_{t-k} - E_{t-k-1}R_{t-k} \qquad (13)$$

in which,

$$E_{t-k-1}R_{t-k} =$$

$$\sum_{s_{t-k},\ldots,s_{t-k-l},d} E[R_{t-k}|s_{t-k},\ldots,s_{t-k-l},d,Y_{t-k-1}]P(s_{t-k},\ldots,s_{t-k-l},d|Y_{t-k-1})$$

Note that $\tilde{\epsilon}_t$ will in general be different than ϵ_t since the information set Y_t does not imply complete knowledge about ϵ_t. If S_t were observable then ϵ_t could be constructed.

To construct the standardized residuals as in equation (8), simply replace $\sigma(S_t)$ with $\sqrt{h_t(S_t)}$. The expected conditional standard deviation from the DDMS-ARCH model is the square root of

$$h_t = \sum_{s_t,\ldots,s_{t-l},d} h_t(s_t)P(s_t,\ldots,s_{t-l},d|Y_T) \qquad (14)$$

where $P(s_t,\ldots,s_{t-l},d|Y_T)$ indicates the full-sample smoother.

Finally, our most general model (DDMS-DD) allows duration to be a conditioning variable in the conditional mean and variance. Independently, Lam (1997) uses a model that conditions on duration in the conditional mean to explain GNP growth. The DDMS-DD model is:

$$\begin{aligned} R_t = \mu(S_t) + \psi(S_t)D(S_t) &+ \sum_{i=1}^{l} \phi_i (R_{t-i} - \mu(S_{t-i}) - \psi(S_{t-i})D(S_{t-i})) \\ &+ (\sigma(S_t) + \zeta(S_t)D(S_t))^2 \nu_t \qquad (15) \\ \nu_t &\sim NID(0,1) \end{aligned}$$

This model allows for dynamic behavior within each regime. For example, persistence in a state implies $D(S_t)$ is an increasing function, and duration effects in the conditional mean and variance are measured by the coefficients $\psi(S_t)$ and $\zeta(S_t)$, respectively. This model is also capable of capturing complicated correlations between the conditional mean and variance. For example, the correlation between the conditional mean and variance may differ over states. These effects can be captured through the parameters $\psi(S_t)$, and $\zeta(S_t)$.

To construct the expected residuals for the DDMS-DD model replace $\sigma(S_t)$ with $(\sigma(S_t) + \zeta(S_t)D(S_t))^2$ in equation (8). The expected conditional standard deviation for DDMS-DD is the square root of

$$h_t = \sum_{s_t,\ldots,s_{t-l},d} (\sigma(s_t) + \zeta(s_t)d)^4 P(s_t, \ldots, s_{t-l}, d|Y_T). \tag{16}$$

3 DATA

The data were obtained from Schwert (1990). These data are monthly returns including dividends and range from 1802 to 1925. Data from 1926 to 1995:12 are from the monthly CRSP value weighted portfolio for the NYSE/AMEX. The choice of data reflects our objective to capture as many stock market phases as possible in order to obtain precise estimates. We used data from 1834:2-1995:12 due to the apparent structural break in the variance of the series around 1834:1. Data from 1802 to 1833 displayed a substantially lower variance relative to post 1833 data. See Table 6 and Figure 2 of Schwert (1990). Using the full sample 1802-1995 did not change the results but required a third state for 1802-1833. This added little to the model so we used the sample 1834:2-1995:12.

Figure 1 plots the monthly returns for our sample. Summary statistics are reported in Table 1. Standard errors robust to heteroskedasticity (estimated by generalized method of moments) are given in parenthesis. Note that the results indicate that the monthly returns are not normally distributed.

4 ESTIMATION RESULTS

As a comparison for the nonlinear models developed and estimated in this paper, Table 2 presents results from a linear AR(5) model estimated by OLS. Standard errors robust to heteroskedasticity are reported. The linear model gives imprecise estimates of the AR terms; only the first coefficient is significant. However, once the nonlinear effects are accounted for in the conditional mean and variance using the Markov-switching model, persistence in the conditional mean was found up to 5 lags. Therefore, for comparison purposes, 5 AR lags are included in the OLS estimation and in all the nonlinear models.

The BDS test of Brock, Dechert, and Sheinkman (1987) has power against a variety of departures from i.i.d. disturbances, including heteroskedasticity and/or nonlinearities in the conditional mean. Under the null hypothesis the BDS statistic is N(0,1). Note that the BDS test can pick up dependence in higher-order moments. To avoid this, Pagan (1996) suggests the Tsay (1986) test for detecting nonlinear structure in the conditional mean. Table 3 reports the BDS test on returns for various embedding dimensions and the Tsay test for various lags. The Tsay test is made robust to heteroskedasticity by use of the White (1980) covariance estimator. The BDS and Tsay statistics suggest that a linear model may be inadequate in capturing the stochastic properties of returns.

Maximum likelihood estimates for the models summarized in Table 4 are presented in Table 5. DDMS-1 is a Markov-switching model with duration-dependent transition probabilities but only the conditional mean is state dependent. DDMS-2 extends DDMS-1 by allowing the variance to be state dependent. In order to capture remaining conditional heteroskedasticity, DDMS-ARCH extends the DDMS-2 model by including an ARCH specification of the conditional variance. Finally, DDMS-DD extends the DDMS-2 model by allowing duration to enter directly as a conditioning variable in both the mean and variance. The estimates for these four alternative models are presented in the respective columns of Table 5.

Table 6 presents results from some residual-based diagnostic tests to evaluate the statistical adequacy of the models reported in Table 5. In particular, we report the Ljung and Box (1978) portmanteau test statistics (and asymptotic p-values) associated with the standardized expected residuals and their squares.

Figure 1 plots the raw monthly returns over our sample. Figures 2 and 3 plot the filtered probabilities associated with the positive return state for two of the alternative model specifications. Figures 4 and 5 plot the expected conditional standard deviations implied by the DDMS-ARCH versus the DDMS-DD specifications. Figures 6 and 7 plot the duration-dependent transition probabilities for the DDMS-ARCH and the DDMS-DD specifications respectively. Finally, Figures 8 and 9 illustrate the duration dependence of the state-specific conditional return and variance.

Columns 2 and 3 of Table 5 present the estimates for our two simplest nonlinear specifications (DDMS-1 and DDMS-2). In both models, $\mu(1) < \mu(2)$. That is, the states are sorted into low and high return states. When the variance is also allowed to be state dependent (DDMS-2), the low return state is associated with high volatility ($\sigma_1 = .102$) and the high return state has low volatility ($\sigma_2 = .036$). A convenient labelling for these two states is bear versus bull markets. The bull market label refers to the high return, low volatility state; whereas the bear market label refers to the low return, high volatility state of the stock market. This sorting is a robust feature of all models. Although the capital asset pricing model (CAPM) indicates a positive relationship between volatility and return on the market portfolio

(Merton (1980)), this need not be the case for intertemporal asset pricing models. Below we discuss how our model allows the correlation between the expected return and volatility to vary across and within states.

Figure 2 displays the probability of state 2 (bull market), the high return state, using the full sample smoother of Kim (1994). Note that a low probability of being in state 2 implies that we are in state 1. Secondly, in both models $\gamma_2(2) > 0$ which implies a declining hazard associated with a bull market. In other words, the probability of exiting the high return state decreases with duration (negative duration dependence).

Finally, the loglikelihood reported in Table 2 for the linear AR(5) model estimated by OLS was 3105.12. This compares with 3185.26 for the simplest of the nonlinear specifications. Therefore, the LR test statistic for the DDMS-1 two-state model versus the linear single-state model (one degree of freedom) is 160.28. Unfortunately, standard asymptotic theory is not applicable. There are two problems – an identification problem and zero scores under the null. Several approaches are available to obtain a correct p-value: the bound of Davies (1987); the numerical method in Hansen (1992) and Hansen (1996) to put a bound on the p-value; and Garcia (1995) who derives the analytic asymptotic distribution under the null for specific models. Nevertheless, such a large test statistic suggests that we can reject the single-state model.

Column 2 of Table 6 suggests that there is remaining persistence in both the standardized estimated residuals and their squares for the most parsimonious specification (DDMS-1). DDMS-2 allows both the mean and the variance to be state dependent, which improves the specification, but there is still some remaining persistence in the conditional variance.

To account for residual conditional heteroskedasticity we implemented the DDMS-ARCH model. Column 4 of Table 5 presents these estimates with $K = 3$ in equation (10). Recall that for this model the intercept of the conditional variance, $\omega(S_t)$, is state dependent but the $\alpha_i,\ i = 1, 2, 3$, are constant across states. Although not reported in Table 5, each $\alpha_i,\ i = 1, 2, 3$, was highly significant ($\sum \alpha_i = .196$). The estimated σ^2 reported in Table 2 was used as a startup value for the ARCH process. Residual-based diagnostics, reported in Table 6, are acceptable for this model.

The memory, τ, of the Markov chain for this model was determined to be 14. That is, duration is significant in affecting the transition probabilities for a little over a year. Since τ is a discrete valued parameter, $\hat{\tau}$ was determined using a grid search from [5,24] with the loglikelihood value as the criterion. Therefore, the standard errors in Table 5 do not take parameter uncertainty associated with τ into account. However, we always found a maximum in the interior of [5,24].

As in the DDMS-2 case, which has a state-dependent variance but no ARCH effects, the DDMS-ARCH model has a high unconditional variance (.01008) associated with the low return state ($\mu(1)$ is insignificantly different from zero) and a low unconditional variance (.0013) associated with the high return state (the conditional mean return in state 2, $\mu(2)$, is positive with a t-statistic of 8.45). That is, once again the sorting is such that the bull market is a high return, low volatility state of the market whereas the bear market is a low return, high volatility state.

Figure 3 shows the probability of a bull market from the full sample smoother. Most of the sample from 1834:2-1995:12 is sorted into the bull market state. The filter clearly identifies the stock market crashes of 1929 and 1987, and the market downturns of the 1930's. The filter also indicates less notable stock market downturns, such as 1855, 1861, 1860, 1907 and 1974.

Markov-switching models allow us to identify turning points. Although we could discuss several periods, we will focus on the crash of 1929. The filtered probabilities, such as those given in Figure 2 and Figure 3, give the probability of being in a particular state at a particular point in time. Through the 1920's until April 1929 the probability is greater than .5 for the high return, low volatility state. On the other hand, from May 1929 to May 1934 the filter gives a probability greater than .5 to the low return, high volatility state. Thus the DDMS-ARCH model identifies a turning point in the stock market in May 1929 – several months prior to the crash. According to this model, that turning point involved a switch from a high return, low volatility state to one with a low average return and an eight-fold increase in volatility.

Figure 4 shows the expected conditional standard deviation from the DDMS-ARCH model. The most apparent feature is the high volatility associated with the great depression. The volatility clustering around the 1860's corresponds with the U.S. civil war.

The coefficients which estimate the dependence of the transition probabilities on duration, $\gamma_2(1)$ and $\gamma_2(2)$, are both significantly positive which again implies a declining hazard associated with both the bull and bear markets. The implications for the length of a bull and bear market are summarized in Figure 6. This shows how the probability of staying in the state changes as duration increases. For example, when the economy is in the bull market (state 2), the probability of staying in the bull market actually increases with duration. That is, the bull market gains momentum. While the probability of staying in the bear market (state 1) also increases with duration, the probability of staying in the state is less than .5 until after four consecutive occurrences of state 1 occur. That is, the low return high volatility state (bear markets) are not persistent until after we have stayed in them for several months. On average, the stock market spends 90% of the time in a bull market and only 10% in a bear market. These are the unconditional probabilities for each state, $P(S=i) = \sum_{d=1}^{\tau} P(S=i, D=d)$, $i = 1, 2$.

One possibility that the DDMS-ARCH specification ignores is that duration has a direct effect on the conditional mean and standard deviation. Our previous models only allow duration to affect the persistence of a state. This motivates our next model (DDMS-DD) that allows duration to enter as a conditioning variable in the conditional mean and variance.

Estimates of DDMS-DD from equation (15) appear in the last column of Table 5. For this parameterization, the difference in returns across states is more pronounced. The conditional return in state 1, $\mu(1)$, is now significantly negative. The positive return after a switch into state 2 (bull market) is roughly three times larger than in the previous models and is still quite precisely estimated (t-statistic is almost 10).

Allowing intra-state dynamics in the DDMS-DD model results in a sorting of returns into bull and bear markets which is more comprehensive than that from the DDMS-ARCH model. In particular, this structure captures many more market downturns than the previous models. For example, the correction in mid-1990 is clearly indicated. This filter provides more detail in other respects as well. For example, although the previous filters identified the the crash in 1987, the DDMS-DD filter signals the crash several months prior to the event.

Since we find a declining hazard associated with the bull market (Figure 7), we still have a momentum effect in the market – as was the case for the DDMS-ARCH model (Figure 6). However, now both the bull and bear markets are persistent, whereas the bear market in the DDMS-ARCH was not persistent until after 4 consecutive months in the bear market.

This model also allows the conditional mean within each state to display dynamic behavior. For example, the parameter $\psi(2)$ estimates the dependence of the conditional mean in state 2 on duration. Note that $\psi(2)$ is significantly different from 0 with a t-statistic of -3.8. Consider how the conditional return in the bull market changes over time. The first period in the high return state the conditional return is .0278 while if state 2 persists 16 periods the conditional return drops to .0068. Thus, the bull market delivers decreasing positive returns. The best market gains come from the start of the bull market. Figure 8 plots the conditional return in state 2 against duration for the DDMS-DD model compared with the earlier DDMS-ARCH model for which duration only affects the timing of switches between states.

In addition, the DDMS-DD model allows duration to be a conditioning variable for the state-specific standard deviations. The coefficients $\zeta(1)$ and $\zeta(2)$ estimate this effect. Note that both are significant but that they have opposite signs. Figure 9 displays the evolution of the conditional standard deviation in each state as duration increases. Since Figure 9 suggests that the variance in state 2 is almost homoskedastic, this means that sources of heteroskedasticity are mainly from the low return (bear market) state and from switches between the two states. Although $\zeta(2)$ is significantly negative it does not show up clearly in Figure 9 due to the scaling.

The conditional standard deviation for the DDMS-DD model over time is shown in Figure 5. The standard deviation is remarkably like that from the DDMS-ARCH model. This is an interesting result which indicates that, at least for this model, duration explains residual heteroskedasticity without resort to an ARCH specification.

Our DDMS-DD model allows the conditional mean and volatility of returns to be correlated in a highly nonlinear fashion. For example, persistence in state 2 will mean the conditional return and standard deviation both decrease as state 2 persists. This means returns and volatility will be positively correlated in a bull market. However, since $\psi(1)$ is insignificant, returns and volatility will be uncorrelated in a bear market.

5 ADDITIONAL SPECIFICATION TESTS

To explore whether or not our results are sensitive to the sample period, the DDMS-DD model was estimated over subsamples 1834-1925, 1926-1995, and 1947-1995 for $\tau = 16$. Table 7 reports these subsample estimates. Overall, the results are similar to the full sample estimates. The nonlinear structure in the conditional mean and variance in the full sample was also found in each of the subsamples considered. The only noticeable difference was a weaker evidence for duration dependence in the transition matrix. In some cases duration was significant at only the 10% or 5% level. For the 1947-1995 sample we found no evidence of duration dependence in the transition matrix. We should be careful in interpreting duration dependence in the transition probabilities in smaller samples; since fewer bull and bear markets occur this makes identification of duration effects in the transition probabilities difficult. However, duration dependence was still significant in both the conditional mean and variance. For example, $\psi(2)$ is significantly negative in all samples. That is, the finding that returns in the bull market are a declining function of duration is a robust feature of all subsamples.

All the results pertaining to duration dependence in the transition matrix were estimated using the logistic functional form. This functional form may be too restrictive. To consider the importance of the functional form on duration dependence, several alternative functional forms to that in (4) were investigated for the full sample and the subsamples for the DDMS-DD model. The following alternative parameterizations for the transition probabilities, $P(S_t = i|S_{t-1} = i, D(S_{t-1}) = d)$, $i = 1, 2$, were studied:

$$\frac{(\gamma_1(i) + \gamma_2(i)d)^2}{1 + (\gamma_1(i) + \gamma_2(i)d)^2}, \quad \sin(\gamma_1(i) + \gamma_2(i)d)^2, \quad \frac{\exp(\gamma_1(i) + \gamma_2(i)\log(d))}{1 + \exp(\gamma_1(i) + \gamma_2(i)\log(d))}.$$

In most cases the logistic function in (4) gave the best loglikelihood value. In the few cases it did not, the alternative functional form only gave a very small improvement in the loglikelihood. We found that all functional forms suggest negative duration dependence in the transition matrix for the full sample.

6 DISCUSSION

In this section we discuss some possible sources of duration dependence associated with the persistence and the conditional moments of monthly stock returns. In particular, we discuss: duration dependence in fundamentals; positive feedback or momentum effects in markets; and stochastic rational bubbles. Further, to evaluate whether or not duration dependence associated with the transition probabilities (the hazard function or persistence) might be coming from one or the other of the conditional mean and variance, we report results from a decoupled model in which transition probabilities associated with the conditional mean are allowed to be independent from those associated with the conditional variance. Finally, we briefly assess whether our nonlinear model exhibits greater explanatory power for monthly mean returns than a linear model.

For our most general models, we found declining hazards for both the bull and bear markets. For example, in the DDMS-DD model, $\gamma_2(2)$ measures the duration effect on the transition probabilities associated with a bull market. That parameter is positive with a t-statistic of about 3. Therefore, a bull market not only tends to persist but becomes more likely to persist as it continues.

One obvious candidate for duration dependence is the fundamentals themselves. However the quarterly frequency of dividends makes testing this hypothesis problematic. It is not clear what type of micro structure could impart duration dependence in dividends. Possibly corporate reluctance for change. If dividends are positively related to the business cycle, they are likely to display positive duration dependence – see Durland and McCurdy (1994) and Diebold, Rudebusch, and Sichel (1993).

Another possible explanation for declining hazards could be irrational investors, such as noise traders (DeLong, Shleifer, Summers, and Waldman (1990)) or fads. Both models allow stock prices to deviate from fundamental prices.

The declining hazards found in all models could be interpreted as a momentum effect in the market. For example, as a bull market persists investors could become more optimistic about the future and hence wish to invest more in the stock market. This results in a decreasing probability of switching out of the bull market. Similarly, the length of a bear market could be related to the amount of pessimism about future returns by investors. This would lead to a substitution from equity into other expected high return instruments, such as T-bills.

Does the empirical evidence presented in this paper imply that there are bubbles in the stock market? McQueen and Thorley (1994) show that a rational stochastic bubble will display negative duration dependence. They sort data into two regimes, and using traditional duration dependence tests find evidence of negative duration dependence in the high return state.

In contrast, our models endogenously sort states based on maximum likelihood, and also allow for dynamic behavior for the conditional mean and variance in each state. Nevertheless, our estimates of hazard functions also decline with duration. Table 5 shows $\gamma(2)$ to be significantly positive in both the DDMS-ARCH and DDMS-DD models. This is also clearly seen in the upward sloping transition probabilities as duration increases in Figures 6 and 7. This evidence would appear to be consistent with a rational stochastic bubble. However, there are several reasons why this conclusion could be premature.

Firstly, complicating matters is the observational equivalence of a growing bubble and any expected future change in fundamentals. For example, if investors know for certain there will be a future change in tax policy at time $\bar{t}$, then prior to $\bar{t}$ prices will display the same exponential growth as a bubble would. To the econometrician who is unaware of the expected regime change, it will appear that prices contain a bubble.

Secondly, estimates of our DDMS-DD model indicate that the conditional return in the bull market decreases with duration (i.e. $\psi(2) < 0$). This contrasts with intuition that the conditional return associated with a bubble state will be a positive function of duration. To see this, consider the well known bubble model of Blanchard and Watson (1982). In this model both the bubble and fundamental components on average must deliver the required return r. Conditional on the bubble growing, the expected return of the bubble component will be greater than r. That is, as the bubble continues to grow it will become the dominant factor affecting observed returns. Consequently, in the bubble state (state 2) of the DDMS-DD, the conditional return should be an increasing function of the duration of that state, (i.e. $\psi(2) > 0$).

Finally, duration dependence associated with transition probabilities (the hazard function) could originate from the conditional mean or the conditional variance process, or both. For example, two independent latent processes may drive the conditional mean and variance, respectively. It is possible that only the latent process for the conditional variance has a duration-dependent hazard. Therefore, in our two-state models it may be that the mean process is dominated by the variance process.

To further explore this possibility, we estimated the DDMS-DD model as in Maheu (1998) in which the latent process for the conditional mean is decoupled from the latent process for the variance. Thus we build a model with two independent Markov chains, one for the conditional mean and the other for the conditional variance. Both Markov chains allow for duration dependence. In this fashion we can test the origin of duration dependence associated with the hazard functions. This model is much more complicated than those considered in this paper, and so our results are only preliminary. We found a similar negative duration dependence structure in the conditional variance and its associated transition matrix. Moreover, we found a similar regime switching structure in the conditional mean with duration being significant, and we found evidence of negative duration dependence in the transition matrix of the conditional mean for the high return state. Based on these preliminary results, it is probable that the origin of duration dependence in the transition matrix of the two-state DDMS-DD model comes from both the conditional mean and variance process.

How important is the nonlinear structure for explanatory power associated with mean stock returns? Unfortunately, there are some complications in comparing explanatory power between linear and nonlinear models. We use the method suggested by Haessel (1978) which is to use the coefficient of determination (R^2) from the regression,

$$R_t = a + b\hat{R}_t + \text{error} \tag{17}$$

where $\hat{R}_t$ is a forecast of R_t from the model of interest.

Using the full sample smoother to infer the state at time t, the R^2 from equation (17) is .0216 for the DDMS-ARCH model and .279 for the DDMS-DD model. Although this measure is in-sample (uses the full sample for parameter estimates) and conditions on contemporaneous information, it shows that there are important nonlinear components in returns. To compare the amount of in-sample explanatory power the models have conditionally at time $t-1$ for returns at time t, we can use the full sample smoother to infer the state at time $t-1$ and use this with the transition matrix to forecast $\hat{R}_t$. Using this method results in an R^2 of .0093 for DDMS-ARCH and .0453 for DDMS-DD. According to this measure, DDMS-DD explains more than 3 times the variation in returns compared to the linear model. (The R^2 from the benchmark linear AR(5) model is .0147.)

The R^2 from the DDMS-DD model and the significance of $\psi(2)$ are evidence that duration effects are important to the conditional mean, not just the variance. This suggests that optimal investment strategies will be linked to duration. Future work will focus on predicting turning points in the market and investigating profitable investment strategies.

7 CONCLUSION

In this paper we have found evidence of nonlinear behavior in monthly stock returns. We study a general mixtures of distributions model which incorporates duration as in Durland and McCurdy (1994). We extended the model to include ARCH effects and allow duration to be a conditioning variable.

Estimates of the empirical models clearly identify a high return and low return state. We interpret these as bull and bear markets. Associated with the high return state is a low conditional variance while the low return state exhibits a higher conditional variance. These results are robust to several model specifications. The DDMS-ARCH model sorts all major stock market downturns into the low return state. The filter also provides dating of turning points of historical stock market crashes such as 1929. The DDMS-DD model shows an improvement in explaining in-sample returns over the linear model.

According to the residual based diagnostic tests, duration as a conditioning variable is able to explain all residual heteroskedasticity. Plots of the conditional standard deviation from the DDMS-ARCH and DDMS-DD model are remarkably similar.

Our models show negative duration dependence in the transition probabilities in both the bull and bear market. Preliminary results suggest that the source of duration dependence in the transition matrix originates from both the conditional mean and conditional variance process. Furthermore, duration is important as a conditioning variable for both the conditional mean and variance. For example, the best market gains come at the start of a bull market, and volatility increases over the duration of a bear market.

ACKNOWLEDGEMENTS

We thank G. W. Schwert for providing data. Thanks to the editor, associate editor and referee for very helpful suggestions and comments. We have also benefited from comments from Long Chen, Stephen Gordon, Allan Gregory, T. Kamionka, Raymond Kan, Chengjun Li, James MacKinnon, Angelo Melino, Ieuan Morgan, Adrian Pagan, Gregor Smith, Simon van Norden and participants at the North American Summer Meetings of the Econometric Society, the Canadian Economic Association annual meetings, the Northern Finance Association meetings, and workshop participants at University of Toronto, University of Western Ontario, Wilfred Laurier University, University of Alberta, and Queen's University. McCurdy thanks the Social Sciences and Humanities Research Council of Canada for financial support.

APPENDIX

This appendix shows how the duration-dependent Markov model can be collapsed into a first order Markov model. Estimation, and smoothing then follows with the usual techniques. The notation follows closely with Hamilton (1994).

Consider a 2 state l lag model for R_t. Define a new latent variable $\mathcal{S}_t$ as:

$$\begin{aligned}
\mathcal{S}_t &= 1 \quad \text{if} \quad S_t = 1, S_{t-1} = 2, ..., S_{t-l} = 2, D(S_t) = 1 && (18)\\
\mathcal{S}_t &= 2 \quad \text{if} \quad S_t = 1, S_{t-1} = 1, S_{t-2} = 2, ..., S_{t-l} = 2, D(S_t) = 2 && (19)\\
\mathcal{S}_t &= 3 \quad \text{if} \quad S_t = 1, S_{t-1} = 1, S_{t-2} = 2, ..., S_{t-l} = 1, D(S_t) = 2 && (20)\\
&\vdots \qquad \vdots \\
\mathcal{S}_t &= N \quad \text{if} \quad S_t = 2, S_{t-1} = 2, ..., S_{t-l} = 2, D(S_t) = \tau && (21)
\end{aligned}$$

where $N = 2^{l+1} + 2(\tau - l - 1)$. Next define ξ_t as an N vector that contains all zeros except for a 1 at the i-th element when $\mathcal{S}_t = i$, or

$$\xi_t = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \tag{22}$$

Now define the transition matrix for $\mathcal{S}_t$ as:

$$P = \begin{bmatrix} p_{11} & p_{21} & \dots & p_{N1} \\ p_{12} & p_{22} & \dots & p_{N2} \\ \vdots & \vdots & & \vdots \\ p_{1N} & p_{2N} & \dots & p_{NN} \end{bmatrix} \tag{23}$$

where $p_{ij} = P(\mathcal{S}_t = j|\mathcal{S}_{t-1} = i)$. Each p_{ij} is constructed from the primary probabilities in equations (4) and (5).

Let the parameters to be estimated be contained in the vector Ω and $Y_{t-1} = \{R_{t-1}, R_{t-2}, \dots, R_1\}$. Let the conditional density of R_t be denoted $f(R_t|\mathcal{S}_t, Y_{t-1}, \Omega)$, and put in the N row vector

$$\eta_t = \begin{bmatrix} f(R_t|\mathcal{S}_t = 1, Y_{t-1}, \Omega) \\ f(R_t|\mathcal{S}_t = 2, Y_{t-1}, \Omega) \\ \vdots \\ f(R_t|\mathcal{S}_t = N, Y_{t-1}, \Omega) \end{bmatrix} \tag{24}$$

Finally the filter and the likelihood can be easily built using vector notation as:

$$\hat{\xi}_{t|t} = \frac{\hat{\xi}_{t|t-1} \odot \eta_t}{\iota(\hat{\xi}_{t|t-1} \odot \eta_t)} \tag{25}$$

$$\hat{\xi}_{t+1|t} = P\hat{\xi}_{t|t} \tag{26}$$

$$f(R_t|Y_{t-1}, \Omega) = \iota(\hat{\xi}_{t|t-1} \odot \eta_t) \tag{27}$$

where $\hat{\xi}_{t|t} = E_t\xi_t$, $\odot$ is direct matrix product and ι is a row vector of ones of length N. Of course the loglikelihood is then $\sum_{t=1}^{T} \log f(R_t|Y_{t-1}, \Omega)$. One startup method for $\hat{\xi}_{0|0}$ is to solve for the unconditional probabilities. See Hamilton (1994) for other possibilities.

Table 1: Summary Statistics for Returns

	Mean	Standard Deviation	Skewness	Excess Kurtosis
R_t	.008 (.004)	.049 (.003)	.099 (.631)	6.747 (.139)
$\lvert R_t\rvert$	.036 (.765e-3)	.034 (.003)	3.020 (.103)	17.514 (.309)
R_t^2	.002 (.139e-3)	.007 (.001)	11.482 (.680)	182.19 (4.400)

Summary statistics for monthly nominal stock returns from 1834:2-1995:12. Standard errors robust to heteroskedasticity appear in parenthesis. The data is monthly nominal returns from Schwert (1990) merged with the CRSP value-weighted portfolio for the NYSE.

Table 2: OLS Estimates

$$
\begin{array}{llll}
R_t = & .007 & +.092R_{t-1} & -.015R_{t-2} \\
 & (.001) & (.039) & (.036) \\
 & -.053R_{t-3} & +.011R_{t-4} & +.054R_{t-5} \\
 & (.037) & (.032) & (.032)
\end{array}
$$

$\hat{\sigma} = .049, \quad R^2 = .0147, \quad \text{lgl} = 3105.12$

Standard errors robust to heteroskedasticity appear in parenthesis. R^2 is the coefficient of determination, and lgl is the loglikelihood.

Table 3: Tests for Nonlinearity in Returns

embedding dimension	1	2	3	4	5
BDS(ϵ)	6.504 [.000]	8.123 [.000]	9.737 [.000]	11.021 [.000]	12.166 [.000]
Lags	2	4	6	8	10
Tsay(1986)	3.017 [.381]	7.135 [.713]	66.170 [.000]	119.847 [.000]	246.54 [.000]

ϵ is chosen to be the standard deviation of returns. The BDS statistic is distributed as N(0,1) under the null of iid innovations. p-values are in square brackets. The Tsay test has an asymptotic $\chi^2(l(l+1)/2)$ distribution where l is the number of lags. The test has been made robust to heteroskedasticity.

Table 4: Duration-Dependent Markov-Switching Models

DDMS-1, Single Variance Model

$$\begin{aligned} R_t &= \mu(S_t) + \sum_{i=1}^{l} \phi_i (R_{t-i} - \mu(S_{t-i})) + \sigma \nu_t \\ \nu_t &\sim NID(0,1),\ S_t = 1,2 \end{aligned}$$

DDMS-2, State Dependent Variance Model

$$\begin{aligned} R_t &= \mu(S_t) + \sum_{i=1}^{l} \phi_i (R_{t-i} - \mu(S_{t-i})) + \sigma(S_t) \nu_t \\ \nu_t &\sim NID(0,1),\ S_t = 1,2 \end{aligned}$$

DDMS-ARCH

$$\begin{aligned} R_t &= \mu(S_t) + \sum_{i=1}^{l} \phi_i (R_{t-i} - \mu(S_{t-i})) + \epsilon_t \\ h_t(S_t) &= \omega(S_t) + \sum_{k=1}^{K} \alpha_k \tilde{\epsilon}_{t-k}, \qquad \epsilon_t = \sqrt{h_t(S_t)}\nu_t \\ \tilde{\epsilon}_t &= R_t - E_{t-1} R_t \\ \nu_t &\sim NID(0,1),\ S_t = 1,2 \end{aligned}$$

DDMS-DD

$$\begin{aligned} R_t &= \mu(S_t) + \psi(S_t)D(S_t) + \sum_{i=1}^{l} \phi_i (R_{t-i} - \mu(S_{t-i}) - \psi(S_{t-i})D(S_{t-i})) \\ &+ (\sigma(S_t) + \zeta(S_t)D(S_t))^2 \nu_t \\ \nu_t &\sim NID(0,1),\ S_t = 1,2 \end{aligned}$$

Transition Probabilities for all Models

$$P(S_t = i | S_{t-1} = i, D(S_{t-1}) = d) = \begin{cases} \frac{\exp(\gamma_1(i)+\gamma_2(i)d)}{1+\exp(\gamma_1(i)+\gamma_2(i)d)} & d \leq \tau \\ \frac{\exp(\gamma_1(i)+\gamma_2(i)\tau)}{1+\exp(\gamma_1(i)+\gamma_2(i)\tau)} & d > \tau \end{cases}$$

$i = 1,2$

Table 5: Duration-Dependent Markov Model Estimates

Parameter	DDMS-1	DDMS-2	DDMS-ARCH	DDMS-DD
$\mu(1)$	-.145 (.011)	-.003 (.009)	-.004 (.009)	-.014 (.007)
$\psi(1)$				.630e-3 (.001)
$\mu(2)$	.011 (.001)	.009 (.001)	.009 (.001)	.029 (.003)
$\psi(2)$				-.001 (.260e-3)
ϕ_1	.104 (.029)	.061 (.023)	.057 (.024)	.012 (.028)
ϕ_2	.022 (.027)	.002 (.022)	.006 (.023)	-.022 (.025)
ϕ_3	.022 (.027)	-.010 (.023)	.005 (.025)	-.002 (.025)
ϕ_4	.050 (.025)	.012 (.023)	.016 (.023)	.021 (.025)
ϕ_5	.072 (.026)	.065 (.022)	.057 (.022)	.080 (.025)
$\sigma(1)$	.044 (.744e-3)	.102 (.009)	$\omega(1)$: .008 (.002)	.175 (.011)
$\zeta(1)$				.014 (.002)
$\sigma(2)$		.036 (.002)	$\omega(2)$: .001 (.921e-4)	.201 (.007)
$\zeta(2)$				-.002 (.520e-3)
$\gamma_1(1)$	.634 (1.270)	-.765 (.494)	-1.034 (.572)	.617 (.337)
$\gamma_2(1)$	-1.024 (.960)	.341 (.087)	.342 (.096)	.085 (.045)
$\gamma_1(2)$	1.092 (.381)	.954 (.527)	.706 (.562)	1.027 (.433)
$\gamma_2(2)$	.219 (.032)	.248 (.071)	.242 (.062)	.122 (.040)
τ	20	12	14	16
lgl	3185.260	3320.325	3331.495	3357.416

Standard errors are in parenthesis. lgl is the loglikelihood value.

Table 6: Misspecification Tests

Test	DDMS-1	DDMS-2	DDMS-ARCH	DDMS-DD
$Q(6)$	19.63 [.003]	3.015 [.807]	3.603 [.730]	3.151 [.790]
$Q^2(6)$	346.11 [.000]	31.143 [.2e-4]	4.581 [.599]	3.255 [.776]

p-values are in square brackets. $Q(6)$ is the Ljung and Box (1978) portmanteau test for autocorrelation in the residuals with 6 lags and $Q^2(6)$ is the same test on the squared residuals for the respective models.

Table 7: DDMS-DD Estimates For Different Samples Periods, $\tau = 16$.

Parameter	1834-1925	1926-1995	1947-1995
$\mu(1)$	-.021 (.017)	-.016 (.007)	-.016 (.011)
$\psi(1)$	.002 (.006)	.25e-3 (.001)	.001 (.002)
$\mu(2)$	.024 (.005)	.038 (.003)	.037 (.003)
$\psi(2)$	-.001 (.39e-3)	-.002 (.35e-3)	-.002 (.32e-3)
ϕ_1	.044 (.035)	-.098 (.047)	-.136 (.054)
ϕ_2	-.026 (.032)	-.049 (.042)	-.137 (.048)
ϕ_3	.014 (.033)	-.026 (.037	-.041 (.046)
ϕ_4	.018 (.033)	.011 (.037)	.023 (.043)
ϕ_5	.067 (.034)	.091 (.036)	.095 (.042)
$\sigma(1)$	.174 (.021)	.166 (.012)	.160 (.016)
$\zeta(1)$	.018 (.006)	.015 (.002)	.014 (.004)
$\sigma(2)$	.224 (.014)	.180 (.008)	.178 (.012)
$\zeta(2)$	-.004 (.880e-3)	-.82e-3 (.001)	-.90e-3 (.001)
$\gamma_1(1)$	.753 (.965)	.702 (.394)	1.916 (1.096)
$\gamma_2(1)$	-.010 (.137)	.101 (.051)	-.146 (.154)
$\gamma_1(2)$	1.527 (.668)	1.216 (.505)	2.075 (.558)
$\gamma_2(2)$	.097 (.053)	.096 (.051)	.030 (.050)
lgl	1947.740	1414.608	1102.618

Standard errors are in parenthesis. lgl is the loglikelihood value.

Figure 1. Returns

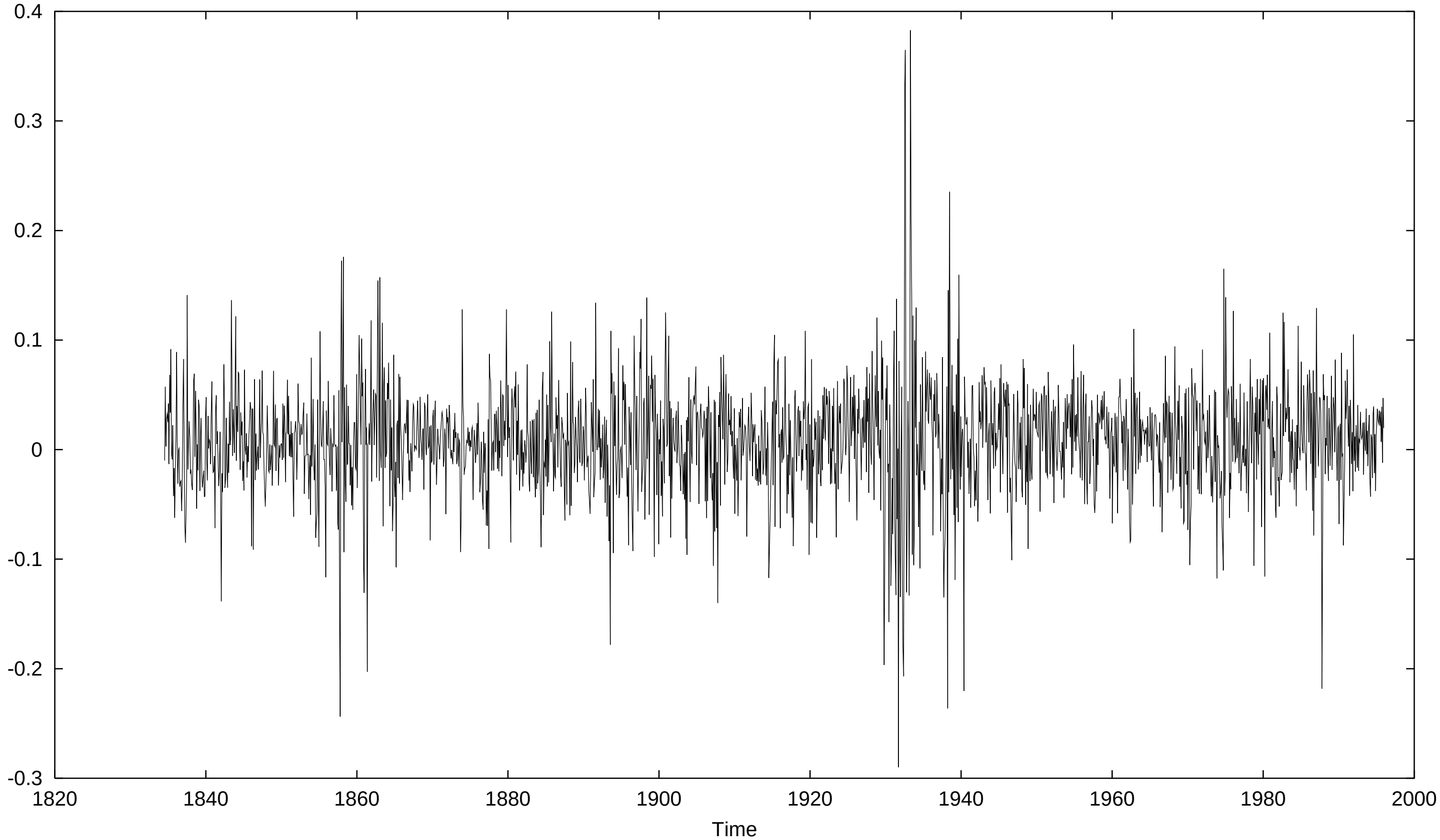


Figure 2. DDMS-2, Probability of State 2

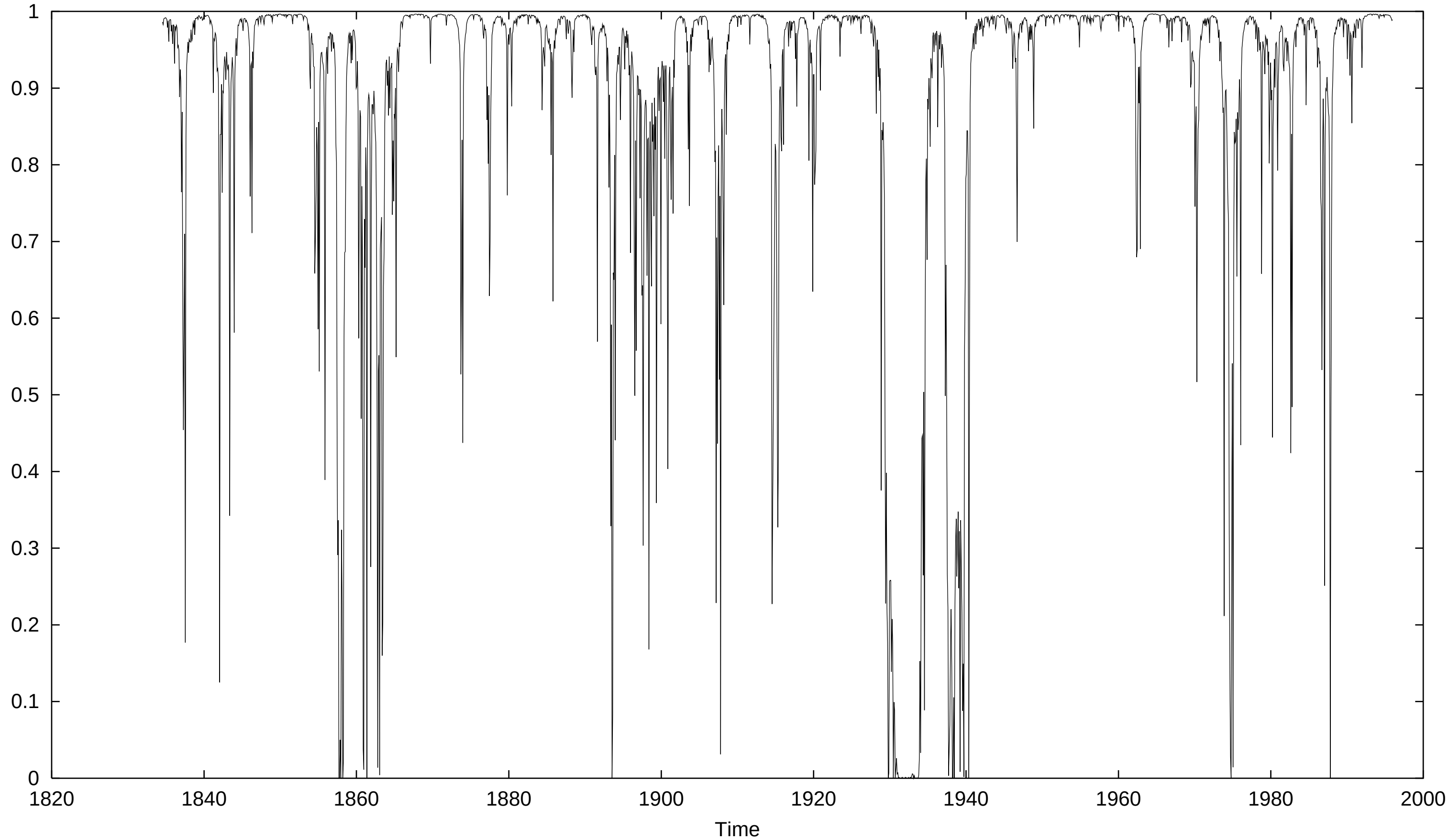


Figure 3. DDMS-ARCH, Probability of State 2

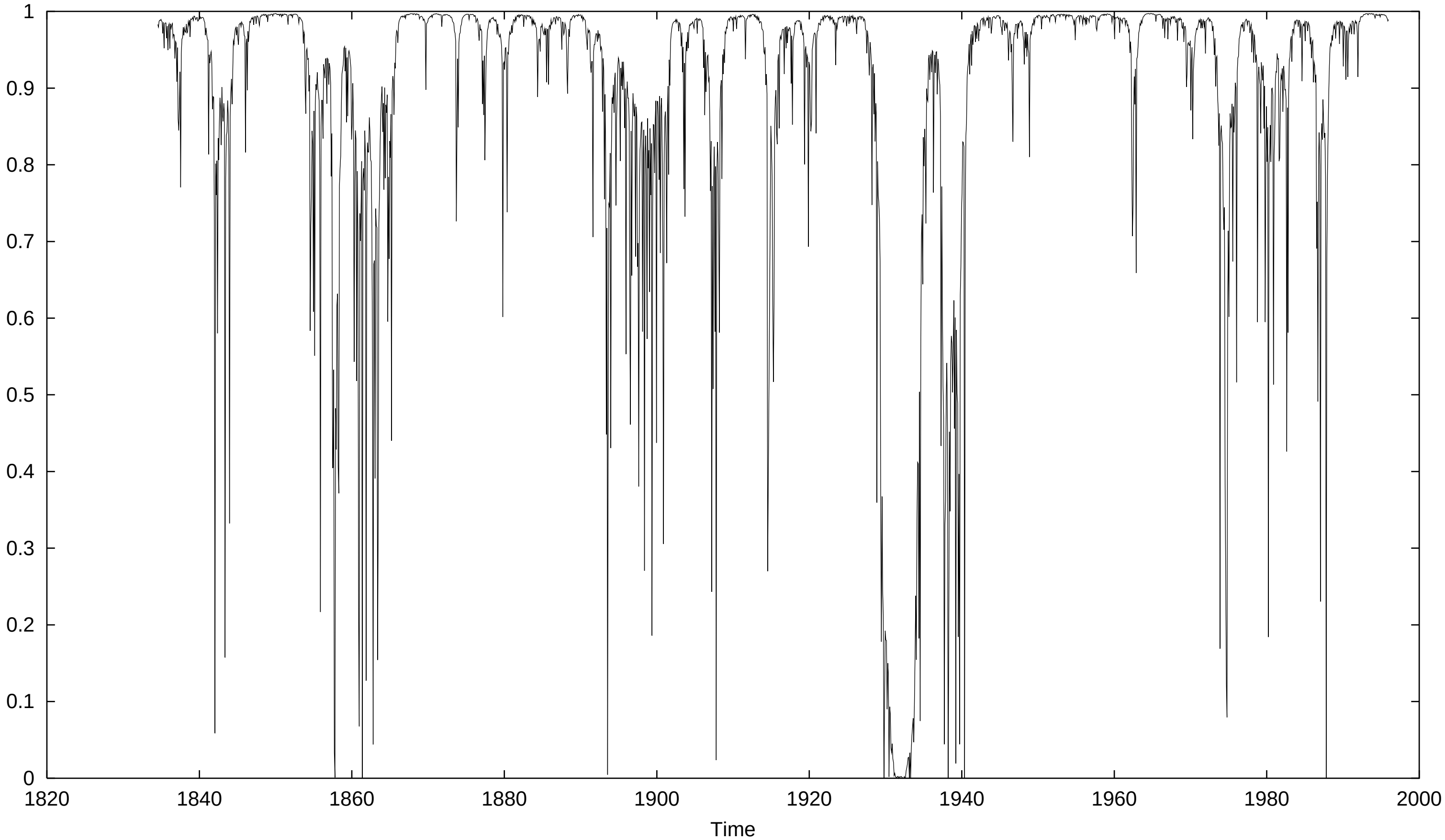


Figure 4. DDMS-ARCH, Expected Conditional Standard Deviation

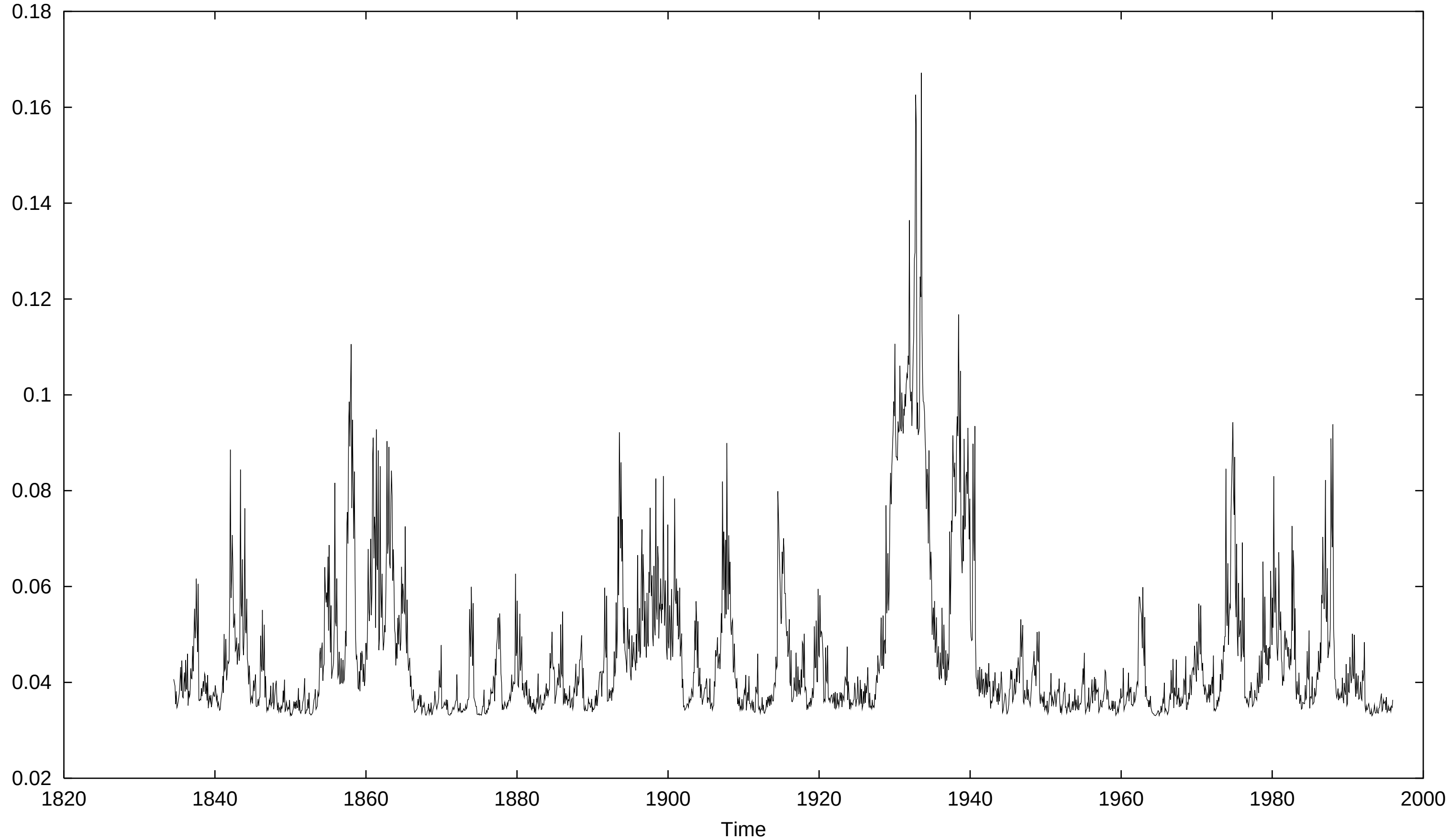


Figure 5. DDMS-DD, Expected Conditional Standard Deviation

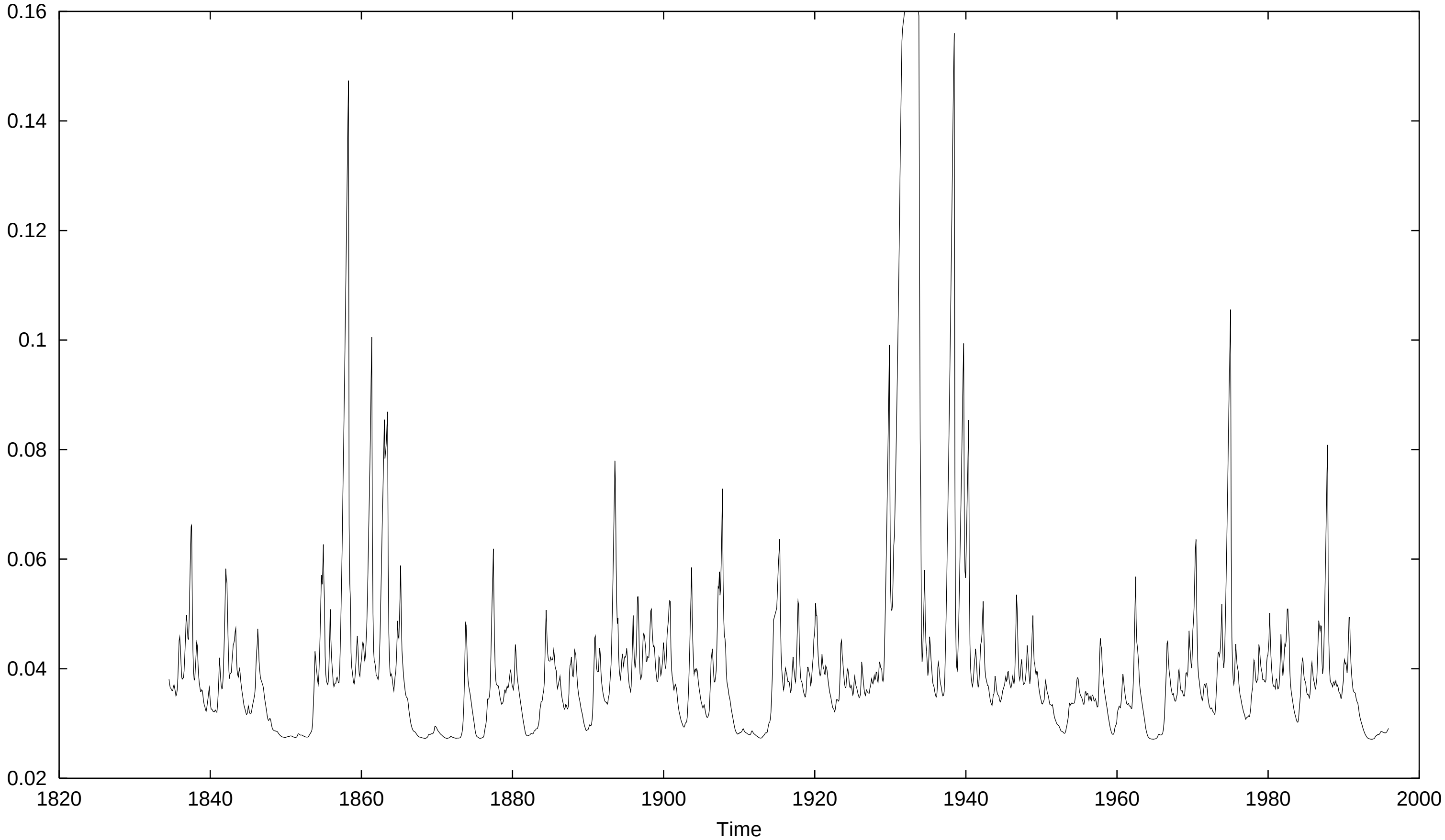


Figure 6. DDMS-ARCH: Transition Probabilities

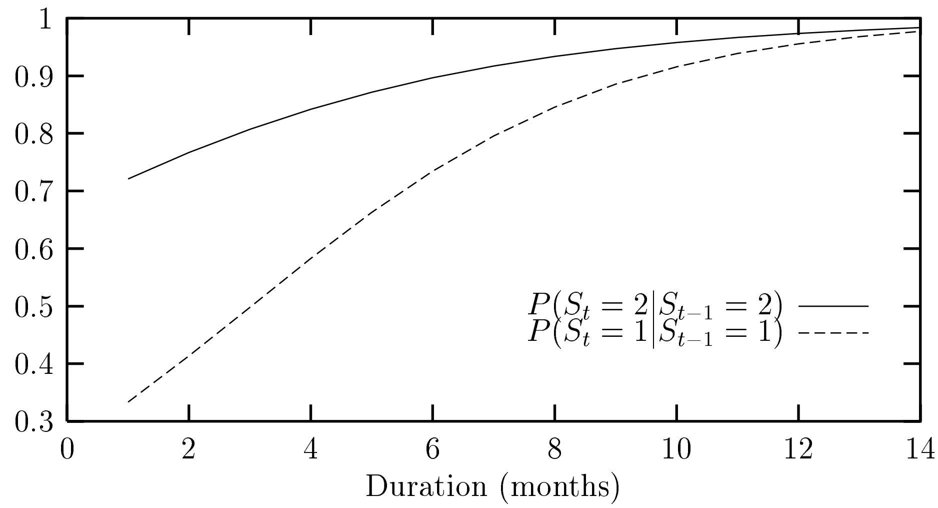


Figure 7. DDMS-DD, Transition Probabilities

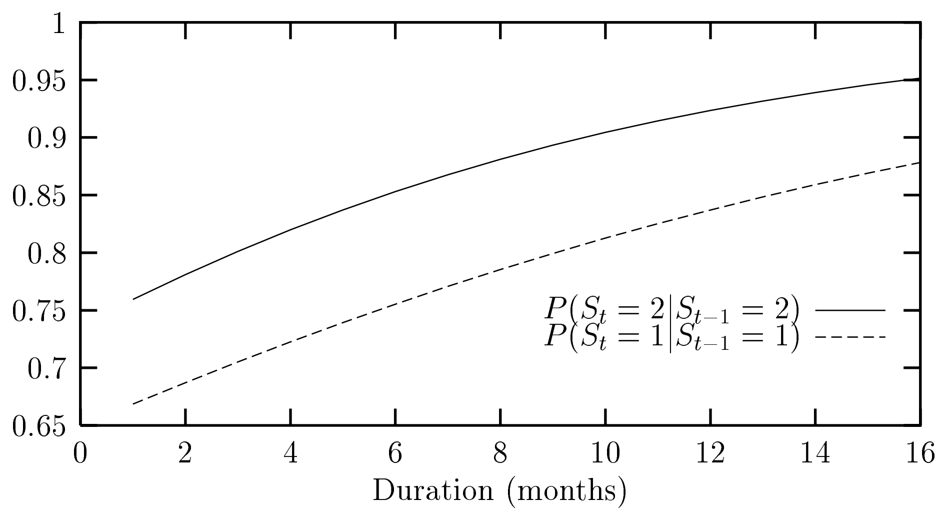


Figure 8. Conditional Return in State 2

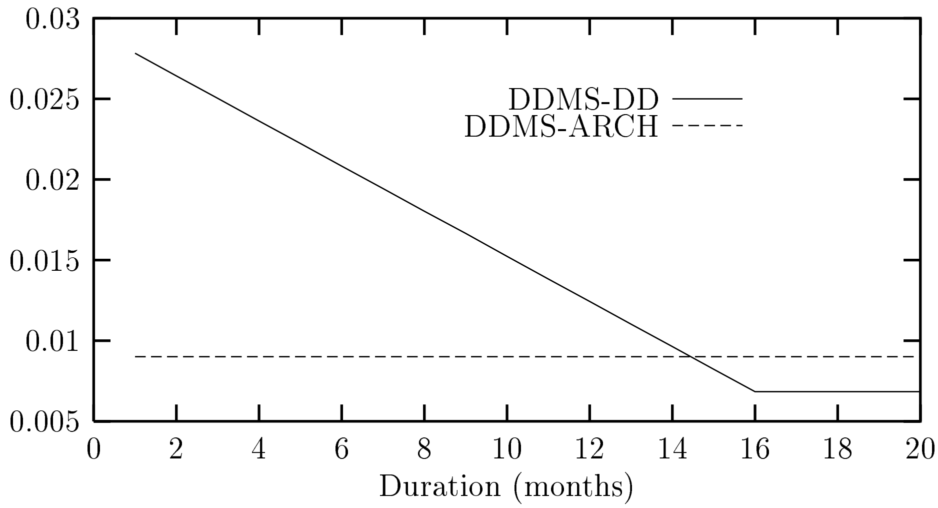
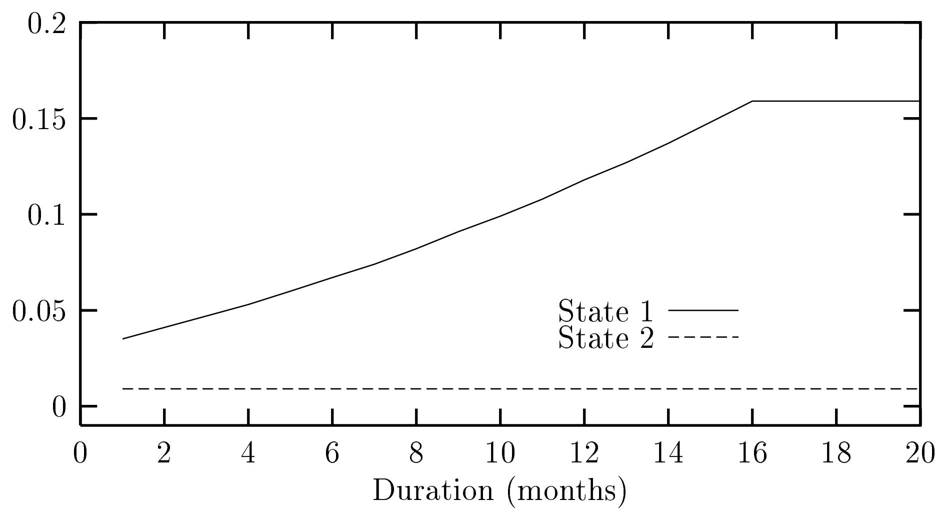


Figure 9. Conditional Standard Deviation from DDMS-DD Model



References

Bickel, P. J., Y. Ritov, and T. Rydén (1998): "Asymptotic normality of the maximum-likelihood estimator for general hidden Markov models," *forthcoming, Annals of Statistics.*

Blanchard, O., and M. Watson (1982): *"Bubbles, Rational Expectations, and Financial Markets" in Crisis in the Economic and Financial System.* Lexington Books, Lexington, MA.

Brock, W. A., W. D. Dechert, and J. A. Sheinkman (1987): "A Test for Independence Based on the Correlation Dimension," Working Paper 8702, University of Wisconsin, Dept. of Economics.

Cai, J. (1994): "A Markov Model of Switching-Regime ARCH," *Journal of Business & Economic Statistics*, 12(3), 309–316.

Cochran, S. J., and R. H. Defina (1995): "Duration Dependence in the US Stock Market Cycle: A Parametric Approach," *Applied Financial Economics*, 5, 309–318.

Davies, R. B. (1987): "Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74, 33–43.

DeLong, J., A. Shleifer, L. Summers, and R. Waldman (1990): "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98(4), 703–738.

Diebold, F., G. Rudebusch, and D. Sichel (1993): *"Further Evidence on Business Cycle Duration Dependence" in Business Cycles, Indicators, and Forecasting.* University of Chicago Press for NBER, 255-284.

Dueker, M. J. (1997): "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility," *Journal of Business & Economic Statistics*, 15(1), 26–34.

Durland, J. M., and T. H. McCurdy (1994): "Duration-Dependent Transitions in a Markov Model of U.S. GNP Growth," *Journal of Business & Economic Statistics*, 12(3), 279–288.

Garcia, R. (1995): "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," University of Montreal.

Gordon, S., and P. St-Amour (1999): "A Preference Regime Model of Bull and Bear Markets," Départment d'Économie, Université Laval.

Gray, S. F. (1996): "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics*, 42, 27–62.

Haessel, W. (1978): "Measuring Goodness of Fit in Linear and Nonlinear Models," *Southern Economic Journal*, 44(3), 648–652.

Hamilton, J. D. (1989): "A New Approach to the Economic Analysis of Non-stationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Hamilton, J. D., and G. Lin (1996): "Stock Market Volatility and the Business Cycle," *Journal of Applied Econometrics*, 11, 573–593.

Hamilton, J. D., and R. Susmel (1994): "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, 307–333.

Hansen, B. (1992): "The Likelihood Ratio Test Under Non-standard Conditions: Testing the Markov Trend Model of GNP," *Journal of Applied Econometrics*, 7, S61–S82.

——— (1996): "Erratum: The Likelihood Ratio Test Under Non-standard Conditions: Testing the Markov Trend Model of GNP," *Journal of Applied Econometrics*, 11, 195–198.

Kim, C.-J. (1994): "Dynamic linear models with Markov-switching," *Journal of Econometrics*, 60, 1–22.

Krishnamurthy, V., and T. Rydén (1998): "Consistent Estimation of Linear and Non-linear Autoregressive Models with Markov Regime," *Journal of Time Series Analysis*, 19(3), 291–307.

Lam, P. (1997): "A Markov Switching Model of GNP Growth with Duration Dependence," Department of Economics, Ohio State University.

Leroux, B. G. (1992): "Maximum-likelihood Estimation for Hidden Markov Models," *Stochastic Processes and Their Applications*, 40, 127–143.

Ljung, L., and G. Box (1978): "On a Measure of Lack of Fit in Time-Series Models," *Biometrika*, 67, 297–303.

Maheu, J. M. (1998): "Duration Dependence and Stochastic Bubbles in Asset Returns," Department of Economics, University of Alberta.

McQueen, G., and S. Thorley (1994): "Bubbles, Stock Returns, and Duration Dependence," *Journal of Financial and Quantitative Analysis*, 29(3), 379–401.

Merton, R. C. (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

Pagan, A. (1996): "The Econometrics of Financial Markets," *Journal of Empirical Finance*, 3, 15–102.

Rydén, T., T. Teräsvirta, and S. Åsbrink (1998): "Stylized Facts of Daily Returns Series and the Hidden Markov Model," *Journal of Applied Econometrics*, 13, 217–244.

Schaller, H., and S. van Norden (1997): "Fads or Bubbles," Bank of Canada, Working Paper 97-2.

Schwert, G. W. (1990): "Indexes of U.S. Stock Prices from 1802 to 1987," *Journal of Business*, 63(3), 399–426.

Tsay, R. S. (1986): "Nonlinearity Tests for Time Series," *Biometrika*, 73(2), 461–66.

Turner, C., R. Startz, and C. Nelson (1989): "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market," *Journal of Financial Economics*, 25, 3–22.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroskedasticity," *Econometrica*, 48(4).