Extracting bull and bear markets from stock returns*

John M. Maheu† Thomas H. McCurdy‡ Yong Song§

Preliminary
May 2009

Abstract

Bull and bear markets are important concepts used in both industry and academia. We propose a new Markov-switching model for the identification of bull and bear regimes for stock returns. Traditional methods to partition the market index into bull and bear regimes are called dating algorithms. Such an approach sorts returns ex post based on a deterministic rule. Statistical inference on returns or investment decisions require more information from the return distribution. Our model fully describes the return distribution while treating bull and bear regimes as unobservable. The model consists of 4 states, two govern the bull regime and two govern the bear regime, which allows for rich and heterogeneous intra-regime dynamics. As a result the model can capture bear market rallies and bull market corrections. A Bayesian estimation approach accounts for parameter and regime uncertainty and provides probability statements regarding future regimes and returns. Applied to 123 years of data our model provides superior identification of trends in stock prices. We evaluate both the econometric specification and the economic value of the model.

*The authors are grateful for comments from seminar participants at McMaster University and to the Social Sciences and Humanities Research Council of Canada for financial support.

†Department of Economics, University of Toronto and RCEA, jmaheu@chass.utoronto.ca

‡Rotman School of Management, University of Toronto and CIRANO, tmccurdy@rotman.utoronto.ca

§Department of Economics, University of Toronto, ysong1@gmail.com

1 Introduction

About the only certainty in the stock market is that, over the long haul, overperformance turns into underperformance and vice versa (dshort.com, February 22, 2009)

There is a widespread belief both by investors, policy makers and academics that low frequency trends do exist in the stock market. Traditionally these positive and negative low frequency trends have been labelled as bull and bear markets respectively. If these trends do exist, then it is important to extract them from the data to analyse their properties and consider their use as inputs into investment decisions and risk assessment.

Traditional methods of identifying bull and bear markets are based on an ex post assessment of the peaks and troughs of the price index. Formal dating algorithms based on a set of rules for classification are found in Gonzalez, Powell, Shi, and Wilson (2005), Lunde and Timmermann (2004) and Pagan and Sossounov (2003). Most of this work is closely related to the dating methods used to identify turning points in the business cycle (Bry and Boschan (1971)). A significant drawback of this approach is that a turning point can only be identified several observations after it occurs. The latent nature of bull and bear markets is ignored and these methods cannot be used for statistical inference on returns or for investment decisions which require more information from the return distribution.

For adequate risk management and investment decisions, we need a probability model for returns and one for which the distribution of returns changes over time. For time series that tend to be cyclical, for example, due to business cycles or bull and bear stock markets, a popular model has been a two-state regime switching model in which the states are latent and the mixing parameters are estimated from the available data. One popular family is Markov-switching (MS) models for which transitions between states are governed by a Markov chain.

Our paper investigates whether we can probabilistically identify low frequency trends (sometimes referred to as primary trends) in a stock market index and also capture the salient features of each phase of the market. We propose a Markov-switching structure which jointly characterizes the unobservable bull and bear market regimes for stock returns, allows intra-regime dynamics, and provides a full description of the return distribution. This approach allows uncertainty about the market regime to be incorporated into out-of-sample forecasts.

Hamilton (1989) applied a two-state MS model to quarterly U.S. GNP growth rates in order to identify business cycles and estimate 1st-order Markov transition probabilities associated with the expansion and recession phases of those cycles. Durland and McCurdy (1994) extended this model to allow the transition probabilities to be a function of duration in the state and applied this duration-dependent MS model to business cycles.

There have been many applications of regime-switching models to stock returns. For example, Hamilton and Lin (1996) relate business cycles and stock market regimes. Cecchetti, Lam, and Mark (2000) and Gordon and St-Amour (2000) derive implications of regime-switching consumption for equity returns. Maheu and McCurdy (2000) allow duration-dependent transition probabilities, as well duration-dependent intra-state dynamics for returns and volatilities. Guidolin and Timmermann (2002), Guidolin and Timmermann (2005), Guidolin and Timmermann (2008), Perez-Quiros and Timmermann (2001) and Turner, Startz, and Nelson (1989), among others, explore the implications of nonlinearities due to regimes switches for asset allocation and/or predictability of returns.

We allow 4 latent states, two govern the bull regime and two govern the bear regime. This structure allows for rich and time-varying intra-regime dynamics. In particular, the model can accommodate short-term reversals (secondary trends) within each regime of the market. For example, in the bull regime it is possible to have a series of persistent negative returns (a correction), despite the fact that the expected long-run return (primary trend) is positive in that regime. Bear markets often exhibit persistent rallies which are subsequently reversed as investors take the opportunity to sell with the result that the expected long-run return is still negative.

Separating short-term reversals in the market from the primary trend is an important empirical regularity that a model must capture for it to be able to reproduce the salient features of the market. This approach is consistent with the definition of bull and bear markets used by Sperandeo (1990) and Chauvet and Potter (2000). It is also consistent with the design of the Lunde and Timmermann (2004) filter to capture long-run structure in stock prices.

Each bull and bear regime has two states. We identify the model by imposing the long-run mean of returns to be negative in the bear market and positive in the bull market. We also impose that the overall mean for returns be positive, while allowing for very different dynamics in each regime. We consider several versions of the model in which the variance dynamics are decoupled from the mean dynamics. We find that a model in which the first and second moment are coupled provides the best fit to the data.

A Bayesian estimation approach accounts for parameter and regime uncertainty and provides probability statements regarding future regimes and returns. Applied to 123 years of data our model provides superior identification of trends in stock prices.

One important difference with our specification is that the richer dynamics in each regime allow us to extract bull and bear markets in higher frequency data. As we show, a problem with a two-state Markov-switching model applied to higher frequency data is that it results in too many switches between the high and low return states. In other words, it is incapable of extracting the low frequency trends in the market. In high frequency data it is important to allow for short-term reversals in the regime of the market.

Our model provides a realistic identification of bull and bear markets and closely matches the output from traditional dating algorithms. The model also provides a good fit to the statistics of the cycle. The use of 4 states is important to the success of our approach. Relative to a two-state model we find that market regimes are more persistent and there is less erratic switching. According to Bayes factors, our 4-state model of bull and bear markets is strongly favored over several alternatives including a two-state model, and different variance dynamics.

Of primary importance is the fact that our model can tell us the probability of a bull or bear regime in real time, unlike the dating algorithms. It can also produce out-of-sample forecasts, something we explore in this paper. We consider several outputs from the model to perform market timing strategies to assess the economic value of the trends we extract from the data. In out-of-sample exercises the model provides valuable probability statements concerning the predictive density of returns. These probability statements are used to signal long, short and cash positions that allow an investor to improve on a pure cash position or a buy and hold strategy.

This paper is organized as follows. The next section describes the data, Section 3 discusses existing ex post market dating algorithms. Section 4 summarizes the benchmark model and develops our proposed specification, and estimation and model comparison are discussed in Section 5. Section 6 presents results including parameter estimates, probabilistic identification of bull and bear regimes, and an analysis of the economic value of our proposed model through market timing strategies and Value-at-Risk forecasts. Section 7 concludes.

2 Data

We begin with 123 years of daily returns (33926 observations). The 1885-1925 daily returns are from Schwert (1990). The 1926-2007 daily returns are the Center for Research on Security Prices (CRSP) value weighted including distributions (VWRETD) index returns for the NYSE+AMEX+NASDAQ stock exchanges. We convert daily returns to continuously compounded returns by taking the natural logarithm of the gross return. We construct weekly continuously compounded returns from the daily continuously compounded returns by cumulating daily returns from Wednesday close to Wednesday close of the following week. If a Wednesday is missing, we use Tuesday close. If the Tuesday is also missing, we use Thursday. Weekly returns are scaled by 100 so they are percentage returns. Unless otherwise indicated, henceforth returns implies continuously compounded percentage returns. Summary statistics are shown in table 1.

3 Bull and Bear Dating Algorithms

Ex post sorting methods for classification of stock returns into bull and bear phases are called dating algorithms. Such algorithms attempt to use a sequence of rules to isolate patterns in the data. A popular algorithm is that used by Bry and Boschan (1971) to identify turning points of business cycles. Pagan and Sossounov (2003) adapted this algorithm to study the characteristics of bull/bear regimes in monthly stock prices. First a criterion for identifying potential peaks and troughs is applied; then censoring rules are used to impose minimum duration constraints on both phases and complete cycles. Finally, an exception to the rule for the minimum length of a phase is allowed to accommodate 'sharp movements' in stock prices.

The Pagan and Sossounov (2003) adaptation of the Bry-Boschan (BB) algorithm can be summarized as follows:

1. Identify the peaks and troughs by using a window of 8 months.
2. Enforce alternation of phases by deleting the lower of adjacent peaks and the higher of adjacent troughs.
3. Eliminate phases less than 4 months unless changes exceed 20%.
4. Eliminate cycles less than 16 months.

Window width and phase duration constraints will depend on the particular series and will obviously be different for smoothed business cycle data than for stock prices. Pagan and Sossounov (2003) provide a detailed discussion of their choices for these constraints.

There are alternative dating algorithms or filters for identifying turning points. For example, the Lunde and Timmermann (2004) (LT) algorithm identifies bull and bear markets using an cumulative return threshold of 20% to locate peaks and troughs moving forward.[1] They define a binary market indicator variable I_t which takes the value 1 if the stock market is in a bull state at time t and 0 if it is in a bear state. The stock price at the end of period t is labelled P_t.

Our application of their filter can be summarized as: Use a 6-month window to locate the initial local maximum or minimum. Suppose we have a local maximum at time t_0, in which case we set $P_{t_0}^{\max} = P_{t_0}$.

1. Define stopping-time variables associated with a bull market as

$$\tau_{\max}(P_{t_0}^{\max}, t_0 \mid I_{t_0} = 1) = \inf\{t_0 + \tau : P_{t_0+\tau} \geq P_{t_0}^{\max}\}$$

$$\tau_{\min}(P_{t_0}^{\max}, t_0 \mid I_{t_0} = 1) = \inf\{t_0 + \tau : P_{t_0+\tau} \leq 0.8 P_{t_0}^{\max}\}$$

[1] Lunde and Timmermann (2004) explore alternative thresholds and also asymmetric thresholds for switching from bull versus from bear markets. For this description we use a threshold of 20%.

2. One of the following happen.

   - If $\tau_{\max} < \tau_{\min}$, bull market continues, update the new peak value $P^{\max}_{t_0+\tau_{\max}} = P_{t_0+\tau_{\max}}$ discard previous peak at time t_0 and set $I_{t_0+1} = \cdots I_{t_0+\tau_{\max}} = 1$. Goto 1 above.

   - If $\tau_{\max} > \tau_{\min}$, we find a trough at time $t_0 + \tau_{\min}$ and we have been in a bear market from $t_0 + 1$ to $t_0 + \tau_{\min}$, $I_{t_0+1} = \cdots = I_{t_0+\tau_{\min}} = 0$. Record the value $P^{\min}_{t_0+\tau_{\min}} = P_{t_0+\tau_{\min}}$ and mark time t_0 as one peak. Goto 1 below for bear market.

On the other hand suppose t_0 is a local minimum.

1. Bear market stopping times are

$$\tau_{\min}(P^{\min}_{t_0}, t_0 \mid I_{t_0} = 0) = \inf\{t_0 + \tau : P_{t_0+\tau} \leq P^{\min}_{t_0}\}$$

$$\tau_{\max}(P^{\min}_{t_0}, t_0 \mid I_{t_0} = 0) = \inf\{t_0 + \tau : P_{t_0+\tau} \geq 1.2 P^{\min}_{t_0}\}$$

2. One of the following happens.

   - If $\tau_{\min} < \tau_{\max}$, bear market continues, update the trough point forward, $P^{\min}_{t_0+\tau_{\min}} = P_{t_0+\tau_{\min}}$ discard previous trough value at time t_0 and set $I_{t_0+1} = \cdots = I_{t_0+\tau_{\min}} = 0$. Goto 1.

   - If $\tau_{\min} > \tau_{\max}$ we have a peak at $t_0 + \tau_{\max}$ and have been in a bull market from $t_0 + 1$ to $t_0 + \tau_{\max}$, $I_{t_0+1} = \cdots = I_{t_0+\tau_{\min}} = 1$. Record the value $P^{\max}_{t_0+\tau_{\max}} = P_{t_0+\tau_{\max}}$ and mark time t_0 as a trough and goto 1 above for the bull market.

This process is repeated until the last data point.

The classification into bull and bear regimes using these two filters is found in Table 2. There are several features to note. First, the sorting of the data is broadly similar but with important differences. For example, during the 1930s the BB approach finds many more switches between the market phases than the LT routine does. More recently, both identify 1987-10 as a trough but the subsequent bull phase ends in 1998-07 for LT but 2000-03 for BB. Generally, the BB filter identifies more bull and bear markets (31 and 31) than the LT filter (24 and 25). The average bear duration is 53.7 (BB) and 47.2 (LT) weeks while the average bull durations are very different, 152.4 (BB) and 217.0 (LT). In other words, the different parameters and assumptions in the filtering methods can result in a very different classification of market phases.

Although such dating algorithms can filter the data to locate different regimes, they cannot be used for forecasting or inference. A two-step approach which involves first sorting the data into market regimes and then following with an econometric model conditional on regimes is possible. Such an approach ignores uncertainty from regime estimation and does not allow it to be incorporated into the second step of estimation and forecasting. In addition, since the sorting rule focuses on the first moment, it does not characterize the full distribution of returns. The latter is required if we wish to derive features of the regimes that are useful for measuring and forecasting risk.

Further, the dating algorithms sort returns into a particular regime with probability zero or one. However, the data provides more information; investors may be interested in estimated probabilities associated with the particular regimes. Such information can be used to answer questions such as 'How likely is it that the market could turn into a bear next month?' or 'Are we in a bear market now or just a correction'? Probabilistic modeling of latent states can help answer such questions.

Nevertheless, the dating algorithms are still very useful. For example, we use the BB and LT algorithms to sort data simulated from our candidate parametric models in order to determine whether the latter can match commonly perceived features of bull and bear markets.

4 Models

In this section, we briefly review a benchmark two-state model, our proposed 4-state model, and some alternative specifications of the latter used to evaluate robustness of our best model.

4.1 Two-State Markov-Switching Model

The concept of bull and bear markets suggests cycles or trends that get reversed. Since those regimes are not observable, as discussed in Section 1, two-state latent-variable MS models have been applied to stock market data. A two-state 1st-order Markov model can be written

$$r_t|s_t \sim N(\mu_{s_t}, \sigma^2_{s_t}) \tag{4.1}$$

$$p_{ij} = P(s_t = j|s_{t-1} = i) \tag{4.2}$$

$i = 0, 1$, $j = 0, 1$. We impose $\mu_0 < 0$ and $\mu_1 > 0$ so that $s_t = 0$ is the bear market and $s_t = 1$ is the bull market.

Modeling of the latent regimes, regime probabilities, and state transition probabilities, allows explicit model estimation and inference. In addition, in contrast to dating algorithms or filters, forecasts are possible. Investors can base their investment decisions on the posterior states or the whole forecast density.

4.2 New 4-state Model

Consider the following general $K+1$ state first-order Markov-switching model for returns

$$\begin{aligned} r_t|s_t &\sim N(\mu_{s_t}, \sigma^2_{s_t}) &(4.3)\\ p_{ij} &= P(s_t = j|s_{t-1} = i) &(4.4)\end{aligned}$$

$i = 0, ..., K,\ j = 0, ..., K$. We will focus on a 4-state model, $K = 3$. Without any additional restrictions we cannot identify the model and relate it to market phases. Therefore, we consider the following restrictions. First, the states $s_t = 0, 1$ are assumed to govern the bear market; we label these states as the *bear regime.* The states $s_t = 2, 3$ are assumed to govern the bull market; these states are labeled the *bull regime.* Each regime has 2 states which allows for positive and negative periods of price growth within each regime. In particular

$$\begin{aligned} \mu_0 &< 0 \text{ (bear negative growth)}, &(4.5)\\ \mu_1 &> 0 \text{ (bear positive growth)},\\ \mu_2 &< 0 \text{ (bull negative growth)},\\ \mu_3 &> 0 \text{ (bull positive growth)}.\end{aligned}$$

This structure can capture short-term reversals in market trends. Each state can have a different variance and can accommodate autoregressive heteroskedasticity. Therefore, conditional heteroskedasticity within each regime can be captured.

Consistent with the 2 states in each regime the transition matrix is

$$P = \begin{pmatrix} p_{00} & p_{01} & 0 & p_{03} \\ p_{10} & p_{11} & 0 & 0 \\ 0 & 0 & p_{22} & p_{23} \\ p_{30} & 0 & p_{32} & p_{33} \end{pmatrix} \qquad (4.6)$$

so that each regime will tend to persist and move between positive and negative returns but can escape to the other regime with probabilties p_{03} and p_{30}.[2] The unconditional probabilities associated with P can be solved (Hamilton (1994))

$$\pi = (A'A)^{-1}A'e \qquad (4.7)$$

where $A' = [P' - I,\ \iota]$ and $e' = [0, 0, 0, 0, 1]$ and $\iota = [1, 1, 1, 1]'$.

[2] Note that only a negative return bear state can exit to a positive return bull state, and only a positive return bull state can exit to a negative return bear state. This improves identification of regimes. It implies, for example, that a bear market rally can never immediately precede a transition into a bull market. In other words, a bear market rally that turns into a bull market is labelled a bull market.

Using the matrix of unconditional state probabilities given by (4.7), we impose the following conditions on long-run returns in each regime,

$$E[r_t|\text{bear regime}, s_t = 0, 1] = \frac{\pi_0}{\pi_0 + \pi_1}\mu_0 + \frac{\pi_1}{\pi_0 + \pi_1}\mu_1 < 0 \tag{4.8}$$

$$E[r_t|\text{bull regime}, s_t = 2, 3] = \frac{\pi_2}{\pi_2 + \pi_3}\mu_2 + \frac{\pi_3}{\pi_2 + \pi_3}\mu_3 > 0, \tag{4.9}$$

along with a restriction that bull markets last longer than bear markets, $\pi_0+\pi_1 < \pi_2+\pi_3$. We impose no constraint on the variances.

The equations (4.5) and (4.6), along with equations (4.8) and (4.9), serve to identify bull and bear regimes.[3] The bull (bear) regime has a long-run positive (negative) return. Each market regime can display short-term reversals that differ from their long-run mean. For example, a bear regime can display a bear market rally (temporary period of positive returns), even though its long-run expected return is negative. Similarly for the bull market.

4.3 Other Models

Besides the 4-state model we consider several other specifications and provide model comparisons among them. The dependencies in the variance of returns are the most dominate feature of the data. This structure may adversely dominate dynamics of the conditional mean. The following specifications are included to investigate this issue.

4.3.1 Restricted 4-State Model

This is identical to the 4-state model in Section 4.2 except that inside a regime the return innovations are homoskedastic. That is, $\sigma_0^2 = \sigma_1^2$ and $\sigma_2^2 = \sigma_3^2$. In this case, the variance within each regime is restricted to be constant although the overall variance of returns can change over time due to switches between regimes.

4.3.2 Markov-Switching Mean and i.i.d. Variance Model

In this model the mean and variance dynamics are decoupled. This is a robustness check to determine to what extent the variance dynamics might be driving the regime transitions. This specification is identical to the Markov-switching model in Section 4.2 except that only the conditional mean follows the Markov chain while the variance

[3]If the transition probabilities p_{30} or p_{03} are sufficienty small, it is possible for s_t to become trapped in a bear regime ($s_t = 0, 1$) or a bull regime ($s_t = 2, 3$). This could result in most, or all, of our data being sorted into a single regime. To remove this possibility of an absorbing state we impose the bounds $p_{03} > 0.01$, and $p_{30} > 0.01$ which ensures a minimum of transitions between bull and bear regimes. The Gibbs sampling approach to posterior simulation imposes each of these constraints in estimation.

follows an independent i.i.d mixture. That is,

$$r_t|s_t = \mu_{s_t} + z_t \tag{4.10}$$

$$z_t \sim \sum_{i=1}^{L} \eta_i \mathbf{N}(0, \sigma_i^2) \tag{4.11}$$

$$p_{ij} = P(s_t = j|s_{t-1} = i) \tag{4.12}$$

$i, j = 0, ..., K$, $\sum_{i=1}^{L} \eta_i = 1$ and $\eta_i \geq 0$. For identification, $\sigma_1^2 < \sigma_2^2 < \cdots < \sigma_K^2$ is imposed along with the constraints used for the conditional mean in the previous section.

5 Estimation and Model Comparison

5.1 Estimation

In this section we discuss Bayesian estimation for the most general model introduced in Section 4.2 assuming there are $K + 1$ total states, $k = 0, ..., K$. The other models are estimated in a similar way with minor modifications.

There are 3 groups of parameters $M = \{\mu_0, ..., \mu_K\}$, $\Sigma = \{\sigma_0^2, ..., \sigma_K^2\}$, and the elements of the transition matrix P. Let $\theta = \{M, \Sigma, P\}$ and given data $I_T = \{r_1, ..., r_T\}$ we augment the parameter space to include the states $S = \{s_1, ..., s_T\}$ so that we sample from the full posterior $P(\theta, S|I_T)$. Assuming conditionally conjugate priors $\mu_i \sim N(m_i, n_i^2)$, $\sigma_i^{-2} \sim G(v_i/2, s_i/2)$ and each row of P following a Dirichlet distribution, allows for a Gibbs sampling approach following Chib (1996). Gibbs sampling iterates on sampling from the following conditional densities given startup parameter values for M, Σ and P:

- $S|M, \Sigma, P$
- $M|\Sigma, P, S$
- $\Sigma|M, P, S$
- $P|M, \Sigma, S$

Sequentially sampling from each of these conditional densities results in one iteration of the Gibbs sampler. Dropping an initial set of draws to remove any dependence from startup values, the remaining draws $\{S^{(j)}, M^{(j)}, \Sigma^{(j)}, P^{(j)}\}_{j=1}^{N}$ are collected to estimate features of the posterior density. Simulation consistent estimates can be obtained as sample averages of the draws. For example, the posterior mean of the state dependent mean and standard deviation of returns are estimated as

$$\frac{1}{N}\sum_{j=1}^{N} \mu_k^{(j)}, \quad \frac{1}{N}\sum_{j=1}^{N} \sigma_k^{(j)}, \tag{5.1}$$

for $k = 0, ..., K$ and are simulation consistent estimates of $E[\mu_k | I_T]$ and $E[\sigma_k | I_T]$ respectively.

The first sampling step of $S|M, \Sigma, P$ involves a joint draw of all the states. Chib (1996) shows that this can be done by a so-called forward and backward smoother through the identity

$$p(S|\theta, I_T) = p(s_T|\theta, I_T) \prod_{t=1}^{T-1} p(s_t|s_{t+1}, \theta, I_t). \tag{5.2}$$

The forward pass is to compute the Hamilton (1989) filter for $t = 1, ..., T$

$$\begin{aligned} p(s_t = k|\theta, I_{t-1}) &= \sum_{l=0}^{K} p(s_{t-1} = l|\theta, I_{t-1}) p_{lk}, \quad k = 0, ..., K, & (5.3) \\ p(s_t = k|\theta, I_t) &= \frac{p(s_t = k|\theta, I_{t-1}) f(r_t|I_{t-1}, s_t = k)}{\sum_{l=0}^{K} p(s_t = l|\theta, I_{t-1}) f(r_t|I_{t-1}, s_t = l)}, \quad k = 0, ..., K. & (5.4) \end{aligned}$$

Note that $f(r_t|I_{t-1}, s_t = k)$ is the normal pdf $N(\mu_k, \sigma_k^2)$. Finally, Chib (1996) has shown that a joint draw of the states can be taken sequentially from

$$p(s_t|s_{t+1}, \theta, I_t) \propto p(s_t|\theta, I_t) p(s_{t+1}|s_t, P), \tag{5.5}$$

where the first term on the right-hand side is from (5.4) and the second term is from the transition matrix. This is the backward step and runs from $t = T-1, T-2, ..., 1$. The draw of s_T is taken according to $p(s_T = k|\theta, I_T), k = 0, ..., K$.

The second and third sampling steps are straightforward and use results from the linear regression model. Conditional on S we select the data in regime k and let the number of observations of $s_t = k$ be denoted as T_k. Then $\mu_k|\Sigma, P, S \sim N(a_k, A_k)$,

$$a_k = A_k \left(\sigma_k^{-2} \sum_{t \in \{t|s_t = k\}} r_t + n_k^{-2} m_k \right), \quad A_k = (\sigma_k^{-2} T_k + n_k^{-2})^{-1}. \tag{5.6}$$

A draw of the variance is taken from

$$\sigma_k^{-2}|M, P, S \sim G\left((T_k + v_k)/2, \left(\sum_{t \in \{t|s_t = k\}} (r_t - \mu_k)^2 + s_k \right) /2 \right) \tag{5.7}$$

Given the conjugate Dirichlet prior on each row of P, the final step is to sample $P|M, \Sigma, S$ from the Dirichlet distribution (Geweke (2005)).

An important byproduct of Gibbs sampling is an estimate of the smoothed state probabilties $P(s_t|I_T)$ which can be estimated as

$$\widehat{p(s_t = i|I_T)} = \frac{1}{N}\sum_{j=1}^{N}\mathbf{1}_{s_t=i}(S^{(j)}) \tag{5.8}$$

for $i = 0, ..., K$.

At each step, if a parameter draw violates any of the prior restrictions discussed in Section 4.2 then it is discarded. For the 4-state model we set the independent priors as

$$\mu_0 \sim N(-2,2)\mathbf{1}_{\mu_0<0}, \mu_1 \sim N(1,2)\mathbf{1}_{\mu_1>0}, \tag{5.9}$$

$$\mu_2 \sim N(-1,2)\mathbf{1}_{\mu_2<0}, \mu_3 \sim N(2,2)\mathbf{1}_{\mu_3>0}, \tag{5.10}$$

$$(p_{00}, p_{01}, p_{03}) \sim \mathbf{Dirichlet}(0.6, 0.37, 0.03), \tag{5.11}$$

$$(p_{10}, p_{11}), (p_{23}, p_{22}) \sim \mathbf{Dirichlet}(0.5, 0.5), \tag{5.12}$$

$$(p_{33}, p_{32}, p_{30}) \sim \mathbf{Dirichlet}(0.03, 0.27, 0.7), \sigma_i^{-2} \sim G(1/20, 1/2). \tag{5.13}$$

These priors are informative but cover a wide range of empirically relevant parameter values.

5.2 Model Comparison

If the marginal likelihood can be computed for a model it is possible to compare models based on Bayes factors. Non-nested models can be compared as well as specifications with a different number of states. Note that the Bayes factor penalizes over-parameterized models that do not deliver improved predictions.[4] For the general Markov-switching model with $K+1$ states, the marginal likelihood for model $\mathcal{M}_i$ is defined as

$$p(r|\mathcal{M}_i) = \int p(r|\mathcal{M}_i, \theta)p(\theta|\mathcal{M}_i)d\theta \tag{5.14}$$

which integrates out parameter uncertainty. $p(\theta|\mathcal{M}_i)$ is the prior and

$$p(r|\mathcal{M}_i, \theta) = \prod_{t=1}^{T} f(r_t|I_{t-1}, \theta) \tag{5.15}$$

is the likelihood which has S integrated out according to

$$f(r_t|I_{t-1}, \theta) = \sum_{k=0}^{K} f(r_t|I_{t-1}, \theta, s_t = k)p(s_t = k|\theta, I_{t-1}). \tag{5.16}$$

[4] This is referred to as an Ockham's razor effect. See Kass and Raftery (1995) for a discussion on the benefits of Bayes factors.

The term $p(s_t = k|\theta, I_{t-1})$ is available from the Hamilton filter. Chib (1995) shows how to estimate the marginal likelihood for MS models. His estimate is based on re-arranging Bayes' theorem as

$$\widehat{p(r|\mathcal{M}_i)} = \frac{p(r|\mathcal{M}_i, \theta^*)p(\theta^*|\mathcal{M}_i)}{p(\theta^*|r, \mathcal{M}_i)} \tag{5.17}$$

where θ^* is a point of high mass in the posterior pdf. The terms in the numerator are directly available above while the denominator can be estimated using additional Gibbs sampling runs.[5]

A log-Bayes factor between model $\mathcal{M}_i$ and $\mathcal{M}_j$ is defined as

$$\log(BF_{ij}) = \log(p(r|\mathcal{M}_i)) - \log(p(r|\mathcal{M}_j)). \tag{5.18}$$

Kass and Raftery (1995) suggest interpreting the evidence for $\mathcal{M}_i$ versus $\mathcal{M}_j$ as: not worth more than a bare mention for $0 \leq \log(BF_{ij}) < 1$; positive for $1 \leq \log(BF_{ij}) < 3$; strong for $3 \leq \log(BF_{ij}) < 5$; and very strong for $\log(BF_{ij}) \geq 5$.

6 Results

6.1 Parameter Estimates and Implied Distributions

Model estimates for the 2-state Markov-switching model are found in Table 3. This specification displays a negative conditional mean along with a high conditional variance and a high conditional mean with a low conditional variance. Both regimes are highly persistent. These results are consistent with the sorting of bull and bear regimes in Maheu and McCurdy (2000) and Guidolin and Timmermann (2005).

Estimates for our new 4-state model are found in Table 4. Recall that states $s_t = 0, 1$ capture the bear regime while states $s_t = 2, 3$ capture the bull regime. Each regime contains a state with a positive and a negative conditional mean. Consistent with the 2-state model, volatility is highest in the bear regime. In particular, the highest volatility occurs in the bear regime in state 0. This state also delivers the lowest expected return. The highest expected return and lowest volatility is in state 3 which is part of the bull regime.

All states are persistent with estimates of p_{ii} ranging from 0.68 to 0.97. Compared to the bear regime, there are more frequent switches inside the bull regime, since the probabilities p_{22} and p_{33} are smaller than p_{00} and p_{11}.

The unconditional distribution of the states is reported in Table 5. The probability of the bear regime is $\pi_0+\pi_1 = 0.441$ while the probability of a bull regime is $\pi_2+\pi_3 = 0.559$.[6]

[5] The integrating constant in the prior pdf is estimated by simulation.

[6] In the following discussion, posterior quantities are computed using the average of the Gibbs sam-

The last column of Table 6 reports posterior quantities associated with the bull and bear market regimes. The weekly conditional means for the bear and bull regimes are -0.04% and 0.27% respectively. The intra-regime returns for this model will exhibit conditional heteroskedasticity. Estimates of the unconditional standard deviations are 3.08 (bear) and 1.43 (bull).

Table 6 also allows a comparison of some regime statistics for the 2-state and 4-state models. For example, the expected duration of regimes is longer in the 4-state model. That is, by allowing heterogeneity within a regime in our 4-state model, we switch between bull and bear markets less frequently.

In the 2-state model, the expected return and variance are fixed within a regime. In this case, the only source of regime variance is return innovations. In contrast, the average variance for each regime in the 4-state model can be attributed to changes in the conditional mean as well as to the average conditional variance of the return innovations. For instance, the average variance of returns in the bear regime can be decomposed as $\mathrm{Var}(r_t|s_t = 0, 1) = \mathrm{Var}(E[r_t|s_t]|s_t = 0, 1) + E[\mathrm{Var}(r_t|s_t)|s_t = 0, 1]$, with a similar result for the bull regime. For the bear phase, the mean dynamics account for a small share 2% of the total variance, while for the bull it is larger at 11%.[7]

The MS-2 model has identical higher order moments in each market. The MS-2 assumes normality in both markets while the MS-4 shows that the data is at odds with this. Skewness in both markets is significantly negative in the MS-4 and has different kurtosis levels in each regime. For instance, the bear market displays a high kurtosis of 5.39 while the bull market has a relatively lower value of 3.51.[8] The bear market has thicker tails and captures more extreme events.

The expected future return, conditional on starting in each of the four states, is shown in Figure 1. This Figure plots, for each state $i = 0., , , .3$, expected future weekly return, $E[r_{t+h}|s_t = i]$, for a range of weeks h, that is, for forecast horizons $t + h$. The term structure of expected returns in each state can differ significantly, but for a long enough forecast horizon they converge to the unconditional mean of returns.

Figures 2 and 3 provide more details on the distributional features of each state. The first figure is the predictive densities in each of the 4 states. All states have either different location and/or different tail shapes. Integrating the 2 states in each regime, generates the implied predictive densities for the bull and bear market illustrated in Figure 3. Also included is the implied unconditional distribution of returns. As mentioned above, the bull regime has a positive mean and more mass around 0 relative to the bear distribution which has much thicker tails and a negative mean. The unconditional distribution is a mixture of these 2 regime densities.

pling draws. In general this will differ from the results derived from the posterior mean of the model parameters.

[7]This is computed as 0.22/(0.22+9.32) and 0.22/(0.22+1.77).

[8]The implied unconditional skewness is -0.61 and kurtosis 7.92.

An important feature of the 4-state model relative to the 2-state version is that the 4-state model allows the realized conditional mean associated with a particular regime to change over time. This is because a regime can have a different sequence of states realized during different historical periods. As an example, consider Figure 4. Data from the 4-state model is simulated; the top panel displays the cumulative return, the second panel the realized state, and the bottom panel the average conditional return in a bull or bear regime.[9] The figure shows that the realized conditional mean assoicated with a particular regime will be different over time. In addition, the returns will display heteroskedasticity inside a regime. These are features that the simpler 2-state parameterization cannot capture. In that model all bull (bear) markets have an identical conditional mean and conditional variance.

6.2 Model Comparisons

One can conduct formal model comparisons based on the marginal likelihoods reported in Table 7. The constant mean and variance model performs the worst. The next model has a constant mean but allows the variances to follow a 4-state i.i.d. mixture. Following this are models with a 2-state versus a 4-state Markov-switching conditional mean – both combined with a 4-state i.i.d. variance as in Section 4.3.2. In both cases, the additional dynamics that are introduced to the conditional mean of returns provides a significant improvement. However, these specifications are strongly dominated by their counterparts which allow a common 2 (and 4) state Markov chain to direct both conditional moments. These specifications capture persistence in the conditional variance.

Note that the log-Bayes factor between the 2-state MS and the 4-state MS in the conditional mean restricted to have only a 2-state conditional variance (Section 4.3.1) is large at $41.3 = -13387.2 - (-13428.5)$. This improved fit comes when additional conditional mean dynamics (going from 2 to 4 states) are added to the basic 2-state MS model. The best model is the 4-state Markov-switching model. The log-Bayes factor in support of the 4-state versus the 2-state model is $107 = -13321.5 - (-13428.5)$. This is very strong evidence that the 4-state specification provides a better fit to weekly returns.

The Markov-switching models specify a latent variable that directs low frequency trends in the data. As such, the regime characteristics from the population model are not directly comparable to the dating algorithms of Section 3. Instead we consider the dating algorithm as a lens to view both the CRSP data and data simulated from our models. Using parameter draws from the Gibbs sampler, we simulate return data from a model and then apply both the BB and the LT dating algorithm to those simulated returns. This is done many times[10] and the average and 0.70 density intervals of these

[9] For instance, the average conditional return for the sequence of bull states in $\{1, 0, 3, 2, 3, 3, 0\}$ is $\frac{1}{4}\mu_2 + \frac{3}{4}\mu_3$.

[10] 10,000 simulations each of 6389 observations.

statistics are reported in Tables 8 and 9. The tables also include the statistics from the CRSP data and a $*$ indicates that a model's 0.70 density interval does not contain the CRSP statistic.

Based on these results, the 2-state model is unable to account for 6 of the data statistics while the 4-state model cannot account for 4 for the BB dating approach. It is not surprising that the 2-state model fails to capture the intra-regime dynamics in the second panel of Table 8 as this is what the 4-state model is designed to do. Based on the LT dating algorithm the MS-4 has 1 more statistic within its density interval than the MS-2. We conclude that the 4-state model generally does as well and often better than the 2 state model, nevertheless, there is some room for improvement. The average bear negative return duration and the average bull positive return duration are difficult for the models to match, producing values too low relative to the data.

The dating of the market regimes using the LT approach are found in the top panel of Figure 5. The shaded portions under the cumulative return denote bull markets while the white portions of the figure are the bear markets. Below this panel is the smoothed probability of a bull market, $P(s_t = 2|I_T) + P(s_t = 3|I_T)$ for the 4-state model. The final plot in Figure 5 is the smoothed probability of a bull market, $P(s_t = 1|I_T)$ from the 2-state model. The 4-state model produces less erratic shifts between market regimes, closely matches the trends in prices, and generally corresponds to the dating algorithm. The two-state model is less able to extract the low frequency trends in the market. In high frequency data it is important to allow intra-regime dynamics, such as short-term reversals.

Note that the success of our model should not be based on how well it matches the results from dating algorithms. Rather this comparison is done to show that the latent-state MS models can identify bull and bear markets with similar features to those identified by conventional dating algorithms. Beyond that, the Markov-switching models presented in this paper provide a superior approach to modeling stock market trends as they deliver a full specification of returns along with latent market dynamics. Such an approach permits out-of-sample forecasting which we turn to next.

6.3 Market Timing

To investigate the value of the model in its ability to identify and predict trends in stock returns, we consider some simple market timing strategies based on the predictive density.

The predictive density for future returns based on current information at time t is computed as

$$p(r_t|I_{t-1}) = \int f(r_t|\theta, I_{t-1})p(\theta|I_{t-1})d\theta \qquad (6.1)$$

which involved integrating out both state and parameter uncertainty using the posterior distribution $p(\theta|I_{t-1})$. From the Gibbs sampling draws $\{S^{(j)}, M^{(j)}, \Sigma^{(j)}, P^{(j)}\}_{j=1}^{N}$ based on data I_{t-1} we approximate the predictive density as

$$\widehat{p(r_t|I_{t-1})} = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=0}^{K} f(r_t|\theta^{(i)}, I_{t-1}, s_t = k)p(s_t = k|s_{t-1}^{(i)}, \theta^{(i)}) \tag{6.2}$$

where $f(r_t|\theta^{(i)}, I_{t-1}, s_t = k)$ follows $N(\mu_k^{(i)}, \sigma_k^{2(i)})$ and $p(s_t = k|s_{t-1}^{(i)}, \theta^{(i)})$ is the transition probability.

Consider an investor with wealth W_{t-1} who has the option to invest in a risk-free asset yielding r_f or to take a long or short position in the market. The investor takes a long position, using the proportion α_L of wealth, if the model indicates that

$$P(r_t > r_f|I_{t-1}) > CV_L \tag{6.3}$$

where CV_L is a specified probability threshold. Next period's wealth will be

$$W_t = \alpha_L W_{t-1}\frac{P_t}{P_{t-1}} + (1-\alpha_L)(1+r_f)W_{t-1} \tag{6.4}$$

Analogously, the investor takes a short position, using a proportion α_S of wealth if

$$P(r_t < 0|I_{t-1}) > CV_S \tag{6.5}$$

where CV_L is a specified probability threshold. In this case the next period wealth is

$$W_t = \alpha_S W_{t-1}\left((1+r_f) - \frac{P_t}{P_{t-1}}\right) + (1+r_f)W_{t-1}. \tag{6.6}$$

Otherwise, the investor puts all his or her wealth in the risk-free asset, in which case wealth next period will be $W_t = (1+r_f)W_{t-1}$.

We investigate the challenging out-of-sample period of 2008. At each point in the sample we re-estimate the model and compute the one-week-ahead predictive density and the associated probabilities required for the above investment strategy. The weekly risk-free rate was obtained from the website of Kenneth French. Given an initial wealth of \$100, Table 10 reports wealth at the end of 2008 for various values of $\alpha_L, CV_L, \alpha_S, CV_S$. The bottom panel of the table reports final wealth from leaving all wealth in the risk-free asset every period; as well as from buying and holding the market index for the full investment horizon.

The buy and hold strategy performs particularly poorly during this period, resulting in a loss of 40% of initial wealth; all other strategies result in better outcomes. The most profitable approach allows for short positions to be taken. For example, taking short positions with 80% of current wealth and long positions of 20% provides the best performance (Sharpe ratio of 0.104) amongst those strategies that we implemented with a return of 18%. This suggests that the model can provide effective signals concerning the direction of the market.

In addition, we forecast out-of-sample returns using the predictive mean versus the sample average. The model (predictive mean) achieves a mean-squared error of 17.32 versus 17.99 for the sample mean. This provides further evidence that the model is capturing trends in the data.

6.4 Value-at-Risk

An industry standard measure of potential portfolio loss is the Value-at-Risk (VaR). $\mathrm{VaR}_{(\alpha),t}$ is defined as the 100α percent quantile of the portfolio value or return distribution given information at time $t-1$. We compute $\mathrm{VaR}_{(\alpha),t}$ from the predictive density MS-4 model as

$$P(r_t < \mathrm{VaR}_{(\alpha),t} | I_{t-1}) = \alpha. \tag{6.7}$$

Given a correctly specified model the probability of a return of $\mathrm{VaR}_{(\alpha),t}$ or less is α.

To compute the Value-at-Risk from the MS-4 model we do the following. First, N draws from the predictive density are taken as follows: draw θ and s_{t-1} from the Gibbs sampler, a future state $\tilde{s}_t$ is simulated based on P and $\tilde{r}_t|\tilde{s}_t \sim N(\mu_{\tilde{s}_t}, \sigma^2_{\tilde{s}_t})$. From the resulting draws, the $\tilde{r}_t$ with rank $[N\alpha]$ is an estimate of $\mathrm{VaR}_{(\alpha),t}$.

Figure 6 displays the conditional VaR from 2005 – 2008 predicted by the MS-4 model, as well as that implied by the normal benchmark for $\alpha = 0.05$. At each point the model is estimated based on information up to $t-1$. Similarly, the bechmark, $N(0, s^2)$, sets s^2 to the sample variance using I_{t-1}. It is clear that the normal benchmark overestimates the VaR for much of the sample and then tends to understate it in 2007-2008. The MS-4 has a very different $\mathrm{VaR}_{(.05),t}$ over time becasue it takes into account the current regime. The potential losses increase considerably after 2007 as the model identifies a move from a bull to a bear market.

7 Conclusion

This paper proposes a new 4-state Markov-switching model to identify bull and bear markets in weekly stock market data. The model fully describes the return distribution while treating bull and bear regimes as unobservable. Of the 4 latent states, two govern the bull regime and two govern the bear regime. This allows for rich and heterogeneous intra-regime dynamics.

The model provides a realistic identification of bull and bear markets and closely matches the output from traditional ex post dating algorithms. The model also provides a good fit to the statistics of the cycle. Relative to a two-state model we find that market regimes are more persistent and there is less erratic switching. Model comparisons show that the 4-state specification of bull and bear markets is strongly favored over several alternatives including a two-state model, as well as various alternative specifications for variance dynamics.

In out-of-sample exercises the model provides probability statements concerning the predictive density of returns. These probability statements are used to signal long, short and cash positions that allow an investor to improve on a pure cash position or a buy and hold strategy.

References

Bry, G., and C. Boschan (1971): *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs.* NBER, New Yor.

Cecchetti, S., P. Lam, and N. Mark (2000): "Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to be True?," *American Economic Review*, 90(4), 787–805.

Chauvet, M., and S. Potter (2000): "Coincident and leading indicators of the stock market," *Journal of Empirical Finance*, 7, 87–111.

Chib, S. (1995): "Marginal likelihood from the Gibbs output," *Journal of the American Statistical Association*, 90(432), 1313–1321.

Chib, S. (1996): "Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models," *Journal of Econometrics*, 75, 79–97.

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.

Geweke, J. (2005): *Contemporary Bayesian Econometrics and Statistics.* Wiley.

Gonzalez, L., J. G. Powell, J. Shi, and A. Wilson (2005): "Two centuries of bull and bear market cycles," *International Review of Economics and Finance*, 14, 469–486.

Gordon, S., and P. St-Amour (2000): "A Preference Regime Model of Bull and Bear Markets," *American Economic Review*, 90(4), 1019–1033.

Guidolin, M., and A. Timmermann (2002): "International asset allocation with regime shifts," *Review of Financial Studies*, 15, 1137–1187.

——— (2005): "Economic Implications of Bull and Bear Regimes in UK Stock and Bond Returns," *The Economic Journal*, 115(111-143).

——— (2008): "International Asset Allocation under Regime Switching, Skew, and Kurtosis Preferences," *Review of Financial Studies*, 21(2), 889–935.

Hamilton, J. D. (1989): "A New Approach to the Economic Analysis of Nonstationary 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.

Kass, R. E., and A. E. Raftery (1995): "Bayes Factors," *Journal of the American Statistical Association*, 90(420), 773–795.

Lunde, A., and A. G. Timmermann (2004): "Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets," *Journal of Business & Economic Statistics*, 22(3), 253–273.

Maheu, J. M., and T. H. McCurdy (2000): "Identifying Bull and Bear Markets in Stock Returns," *Journal of Business & Economic Statistics*, 18(1), 100–112.

Pagan, A. R., and K. A. Sossounov (2003): "A simple framework for analysing bull and bear markets," *Journal of Applied Econometrics*, 18(1), 23–46.

Perez-Quiros, G., and A. Timmermann (2001): "Business Cycle Asymmetries in Stock Returns: Evidence from Higher Order Moments and Conditional Densities," *Journal of Econometrics*, 103(1-2), 259–306.

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

Sperandeo, V. (1990): *Principles of Professional Speculation*. New York: Wiley.

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.

Table 1: Weekly Return Statistics (1885-2007)[a]

N	Mean	standard deviation	Skewness	Kurtosis	J-B[b]
6389	0.173	2.290	-0.50	7.3	5178*

[a] Continuously compounded returns
[b] Jarque-Bera normality test: p-value = 0.00000

Table 2: BB and LT Dating Algorithm Turning Points

Troughs		Peaks		Troughs		Peaks	
BB[a]	LT[b]	BB	LT	BB	LT	BB	LT
	1885-04			1938-03	1938-03	1939-10	1939-10
		1887-05		1940-05	1940-05	1941-09	1941-09
1887-10		1890-06	1890-06	1942-04	1942-04	1946-05	1946-05
1890-12	1890-12	1892-03	1892-03	1947-05	1947-05	1948-06	
1893-07	1893-07	1895-09	1895-09	1949-06		1953-03	
1896-08	1896-08	1899-09		1953-09		1957-07	
1900-06		1902-09	1902-09	1957-11		1960-01	
1903-10	1903-10	1906-10	1906-10	1960-09		1961-12	1961-12
1907-11	1907-11	1909-09		1962-06	1962-06	1966-02	
1910-07		1912-10		1966-10		1968-12	1968-12
1914-12		1916-11	1916-11	1970-05	1970-05	1971-04	
1917-12	1917-12	1919-11	1919-11	1971-11		1973-01	1973-01
1921-06	1921-06	1929-09	1929-09	1974-10	1974-10	1980-11	1980-11
	1929-11		1930-04	1982-08	1982-08	1983-06	
	1930-12		1931-02	1984-07		1987-08	1987-08
1932-06	1932-06		1932-09	1987-10	1987-10		1998-07
	1933-03		1933-07		1998-10	2000-03	2000-03
	1933-10	1934-02		2002-10	2002-10	2007-10	2007-10
1935-03		1937-03	1937-03				

[a] BB: Bry and Boschan algorithm using Pagan and Sossounov parameters
[b] LT: Lunde and Timmermann algorithm

Table 3: MS Two-State Model Estimates

	Mean	Median	Std	0.95 DI
μ_0	-0.42	-0.42	0.13	(-0.68, -0.18)
μ_1	0.31	0.31	0.02	(0.26, 0.35)
σ_0	4.12	4.12	0.12	(3.88, 4.38)
σ_1	1.57	1.57	0.02	(1.52, 1.62)
p_{00}	0.94	0.94	0.01	(0.92, 0.96)
p_{11}	0.99	0.99	0.002	(0.98, 0.99)

This table reports the posterior mean, median, standard deviation and 0.95 density intervals for model parameters.

Table 4: MS 4-State Model Estimates

	Mean	Median	Std	0.95 DI
μ_0	-0.78	-0.78	0.144	(-1.109, -0.515)
μ_1	0.26	0.26	0.049	(0.162, 0.351)
μ_2	-0.38	-0.33	0.225	(-1.046, -0.103)
μ_3	0.70	0.70	0.069	(0.566, 0.836)
σ_0	4.74	4.73	0.176	(4.417, 5.109)
σ_1	2.02	2.01	0.066	(1.901, 2.161)
σ_2	1.63	1.64	0.092	(1.432, 1.809)
σ_3	1.05	1.05	0.062	(0.930, 1.173)
p_{00}	0.89	0.89	0.021	(0.842, 0.925)
p_{01}	0.08	0.08	0.018	(0.049, 0.119)
p_{03}	0.03	0.03	0.007	(0.021, 0.049)
p_{10}	0.03	0.03	0.006	(0.022, 0.046)
p_{11}	0.97	0.97	0.006	(0.954, 0.978)
p_{22}	0.68	0.70	0.091	(0.448, 0.816)
p_{23}	0.32	0.30	0.091	(0.184, 0.553)
p_{30}	0.01	0.01	0.002	(0.010, 0.018)
p_{32}	0.22	0.21	0.053	(0.124, 0.330)
p_{33}	0.77	0.77	0.053	(0.656, 0.864)

The posterior mean, median, standard deviation and 0.95 density intervals for model parameters.

Table 5: Unconditional State Probabilites

	mean	0.95 DI
π_0	0.126	(0.093, 0.162)
π_1	0.315	(0.225, 0.376)
π_2	0.230	(0.119, 0.338)
π_3	0.329	(0.242, 0.436)

The posterior mean and 0.95 density intervals associated with the posterior distribution for π from Equation (4.7).

Table 6: Implied Regime Statistics for MS Models

	MS-2	MS-4
bear mean	-0.42	-0.04
	(-0.68, -0.18)	(-0.126, -0.001)
bear duration	16.5	113.3
	(11.9, 22.7)	(69.1, 162.6)
bear standard deviation	4.12	3.08
	(3.89, 4.38)	(2.88, 3.34)
bear variance from $\text{Var}(E[r_t\|s_t]\|s_t = 0, 1)$	0.00	0.22
		(0.11, 0.38)
bear variance from $E[\text{Var}(r_t\|s_t)\|s_t = 0, 1]$	17.0	9.32
	(15.1, 19.2)	(8.06, 10.9)
bear skewness	0	-0.40
		(-0.54, -0.27)
bear kurtosus	3	5.39
		(4.45,5.98)
bull mean	0.31	0.27
	(0.26, 0.35)	(0.21, 0.33)
bull duration	71.0	142.9
	(50.6, 100.0)	(100.4, 190.7)
bull standard deviation	1.57	1.43
	(1.52, 1.62)	(1.37, 1.50)
bull variance from $\text{Var}(E[r_t\|s_t]\|s_t = 2, 3)$	0.00	0.22
		(0.11, 0.36)
bull variance from $E[\text{Var}(r_t\|s_t)\|s_t = 2, 3]$	2.47	1.77
	(2.32, 2.63)	(1.46, 2.06)
bull skewness	0	-0.42
		(-0.53,-0.32)
bull kurtosus	3	3.51
		(3.26,3.81)

The posterior mean and 0.95 density interval for regime statistics.

Table 7: Log Marginal Likelihoods: Alternative Models

Model	$\log f(Y \mid \text{Model})$
Constant mean with constant variance	-14425.1
Constant mean with 4-state i.i.d variance	-13808.6
MS 2-state mean with 4-state i.i.d. variance (4.10 with $K + 1 = 2$)	-13698.6
MS 4-state mean with 4-state i.i.d. variance (4.10 with $K + 1 = 4$)	-13580.4
MS 2-state mean with coupled MS 2-state variance (4.1)	-13428.5
MS 4-state mean with coupled MS 4-state variance (4.3)	-13321.5
MS 4-state mean with coupled MS 2-state variance ($\sigma_0^2 = \sigma_1^2$, $\sigma_2^2 = \sigma_3^2$)	-13387.2

Table 8: Posterior bull/bear statistics by BB algorithm

	CRSP	MS-2	MS-4
Avg. number of bears	31	31.6	33.3
		(28, 35)	(30, 37)
Avg. bear duration	53.7	49.5	53.5
		(42.3, 56.7)	(46.2, 60.8)
Avg. bear amplitude[b]	-34.4	-33.2	-34.8
		(-38.7, -27.7)	(-40.0, -29.7)
Avg. bear return[c]	-0.64	-0.68	-0.66
		(-0.80, -0.56)	(-0.76, -0.55)
Avg. bear standard deviation[d]	2.44	2.71*	2.68*
		(2.51, 2.93)	(2.46, 2.91)
Avg. number of bulls	31	31.7	33.4
		(28, 36)	(30, 37)
Avg. bull duration	152.4	154.9	140.0
		(129.7, 180.8)	(118.5, 161.8)
Avg. bull amplitude	70.0	68.1	61.0*
		(57.4, 79.0)	(52.4, 69.8)
Avg. bull return	0.46	0.44	0.44
		(0.41, 0.47)	(0.41, 0.47)
Avg. bull standard deviation	1.95	2.00	1.99
		(1.89, 2.12)	(1.86, 2.12)
Avg. bear -ve return	-2.27	-2.37	-2.37
		(-2.59, -2.16)	(-2.56, -2.18)
Avg. bear +ve return	1.65	1.72	1.63
		(1.57, 1.87)	(1.49, 1.76)
Avg. bear -ve return duration	2.39	2.27*	2.25*
		(2.17, 2.36)	(2.17, 2.34)
Avg. bear +ve return duration	1.84	1.74*	1.82
		(1.66, 1.82)	(1.74, 1.90)
Avg. bull -ve return	-1.41	-1.35*	-1.38
		(-1.41, -1.29)	(-1.46, -1.31)
Avg. bull +ve return	1.54	1.64*	1.54
		(1.57, 1.70)	(1.45, 1.63)
Avg. bull -ve return duration	1.65	1.66	1.65
		(1.63, 1.69)	(1.62, 1.68)
Avg. bull +ve return duration	2.77	2.43*	2.65*
		(2.37, 2.49)	(2.56, 2.73)

[a] 70% density interval

[b] Aggregate return over one regime

[c] Average return in one regime

[d] Return standard deviation in one regime

Table 9: Posterior bull/bear statistics by LT Algorithm

	CRSP	MS-2	MS-4
Avg. number of bears	25	26.8	28.1
		(21, 32)	(23, 33)
Avg. bear duration	47.2	41.7	47.2
		(34.0, 49.3)	(38.7, 55.6)
Avg. bear amplitude[b]	-40.5	-39.1	-41.0
		(-43.6, -34.6)	(-45.6, -36.6)
Avg. bear return[c]	-0.86	-0.96	-0.89
		(-1.11, -0.81)	(-1.03, -0.75)
Avg. bear std[d]	3.34	3.10*	3.13
		(2.91, 3.29)	(2.91, 3.36)
Avg. number of bulls	24	26.4	27.8
		(21, 32)	(23, 33)
Avg. bull duration	217.0	209.2	190.4
		(158.5, 261.4)	(145.9, 234.9)
Avg. bull amplitude	88.2	83.5	74.5*
		(66.4, 100.8)	(60.8, 88.1)
Avg. bull return	0.41	0.40	0.40
		(0.37, 0.44)	(0.36, 0.44)
Avg. bull std	2.41	2.14*	2.17*
		(1.99, 2.30)	(2.01, 2.34)
Avg. bear -ve return	-2.70	-2.72	-2.66
		(-2.94, -2.50)	(-2.89, -2.44)
Avg. bear +ve return	1.86	1.83	1.71*
		(1.67, 1.98)	(1.58, 1.85)
Avg. bear -ve return duration	2.43	2.37	2.33
		(2.26, 2.48)	(2.23, 2.43)
Avg. bear +ve return duration	1.79	1.66*	1.74
		(1.58, 1.73)	(1.66, 1.82)
Avg. bull -ve return	-1.37	-1.36	-1.39
		(-1.41, -1.30)	(-1.46, -1.32)
Avg. bull +ve return	1.52	1.63*	1.53
		(1.57, 1.69)	(1.44, 1.62)
Avg. bull -ve return duration	1.72	1.69	1.69
		(1.66, 1.73)	(1.65, 1.72)
Avg. bull +ve return duration	2.69	2.39*	2.59*
		(2.33, 2.44)	(2.51, 2.67)

[a] 70% density interval

[b] Aggregate return over one regime

[c] Average return in one regime

[d] Return standard deviation in one regime

Table 10: Market Timing Out-of-Sample (2008)

	α_L	CV_L	α_S	CV_S	W_T	Sharpe ratio
	0.2	0.5	0.2	0.5	102.25	0.010
	0.5	0.5	0.5	0.5	102.18	0.003
	0.5	0.5	0.0	1.0	89.84	-0.239
	1.0	0.5	0.0	1.0	78.79	-0.243
	0.5	0.5	0.6	0.5	104.65	0.022
	0.5	0.5	0.8	0.5	109.59	0.048
	0.2	0.5	0.8	0.5	118.25	0.104
Buy&Hold	1.0		0.0		59.99	-0.244
Cash	0.0		0.0		101.86	

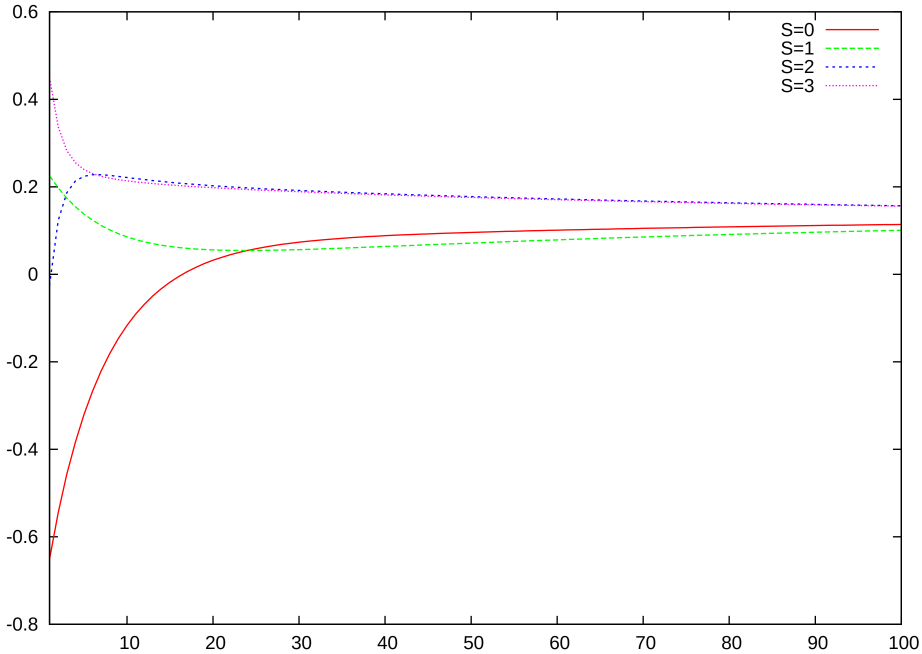

Figure 1: MS 4 States, Conditional Mean Dynamics: $E[r_{t+h}|s_t = S]$ versus h.

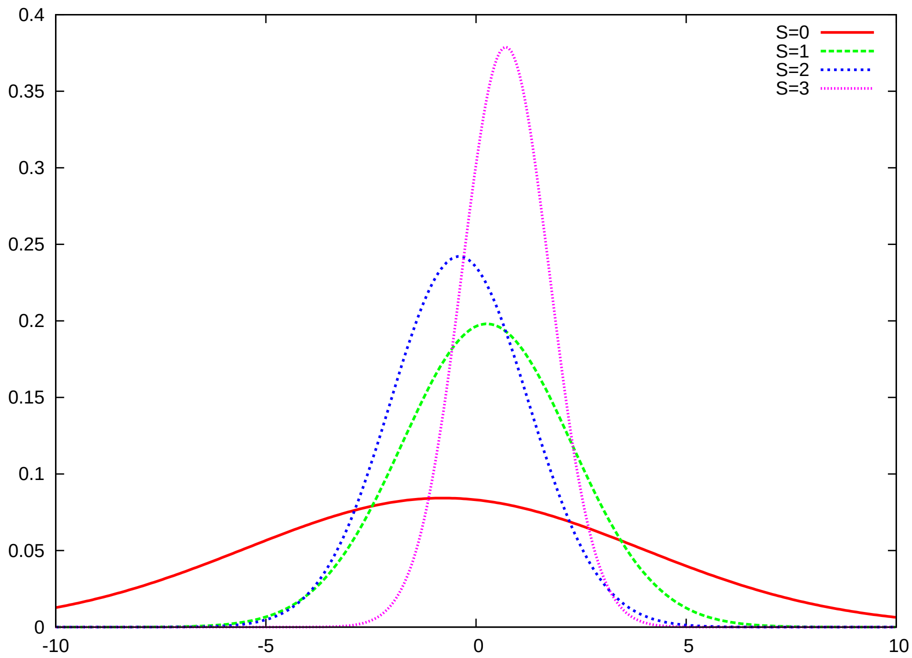

Figure 2: MS 4 States, State Density

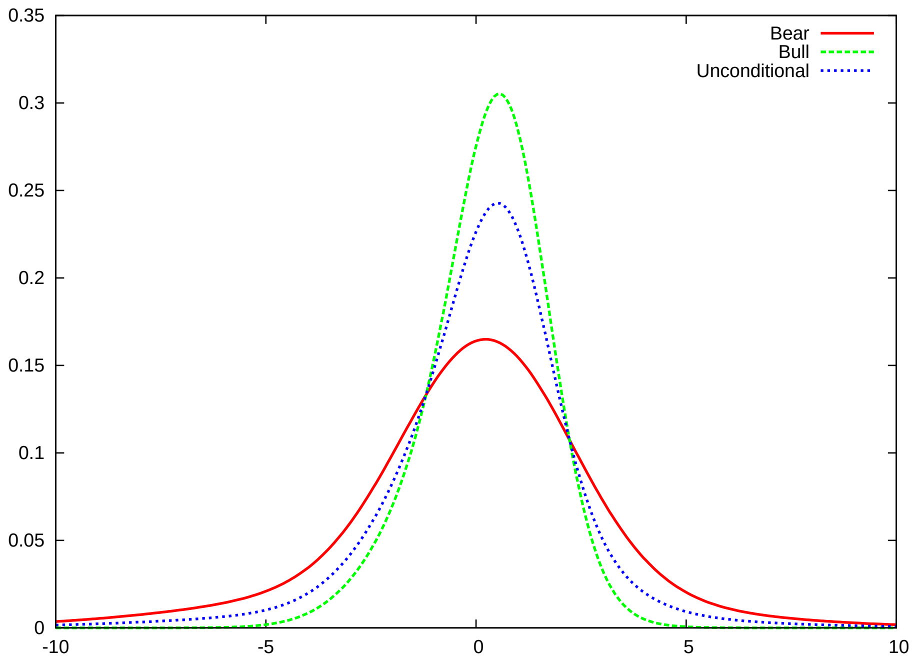

Figure 3: MS 4 States, Regime Density

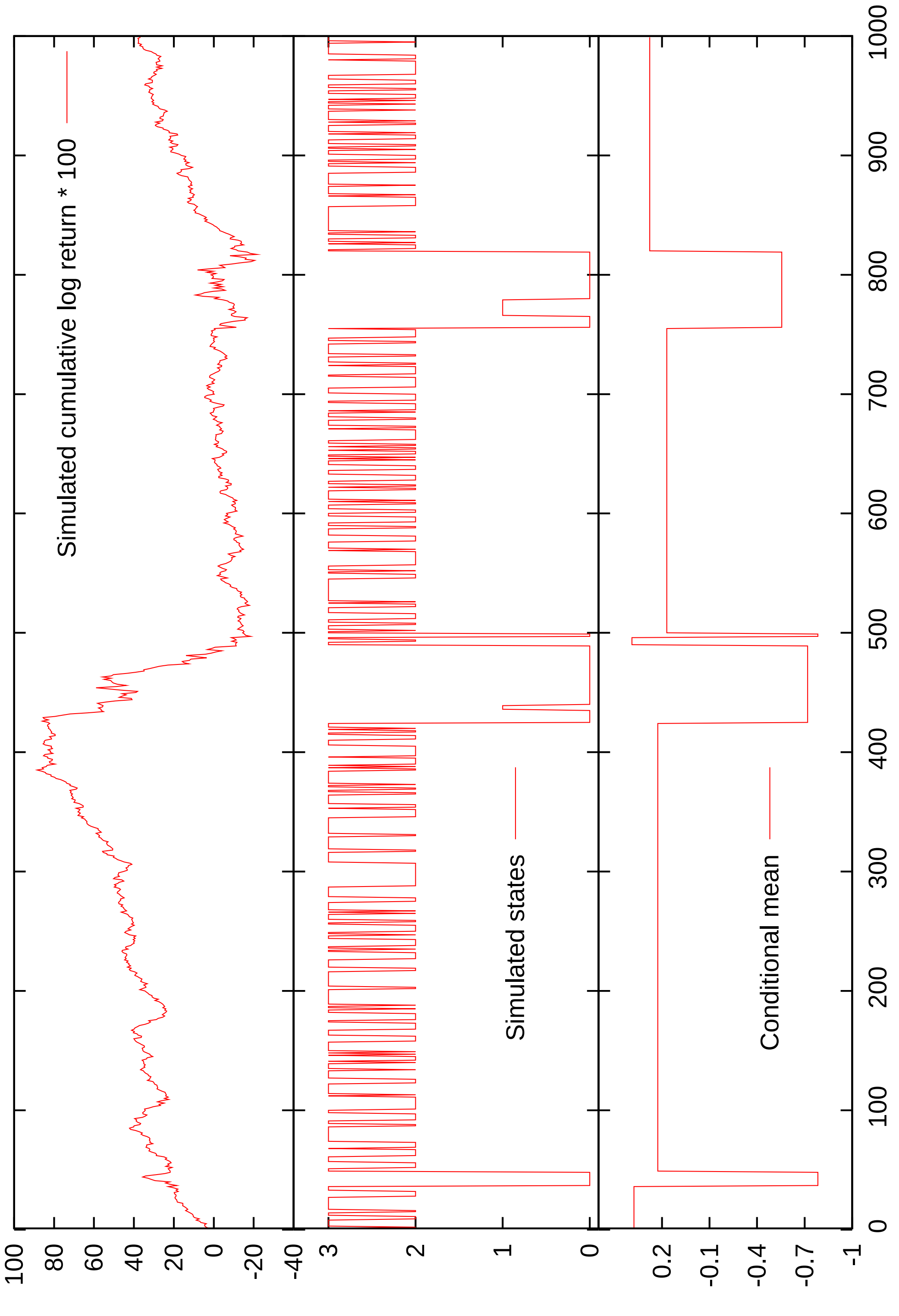

Figure 4: Simulated Returns, States and Average Regime Conditional Mean

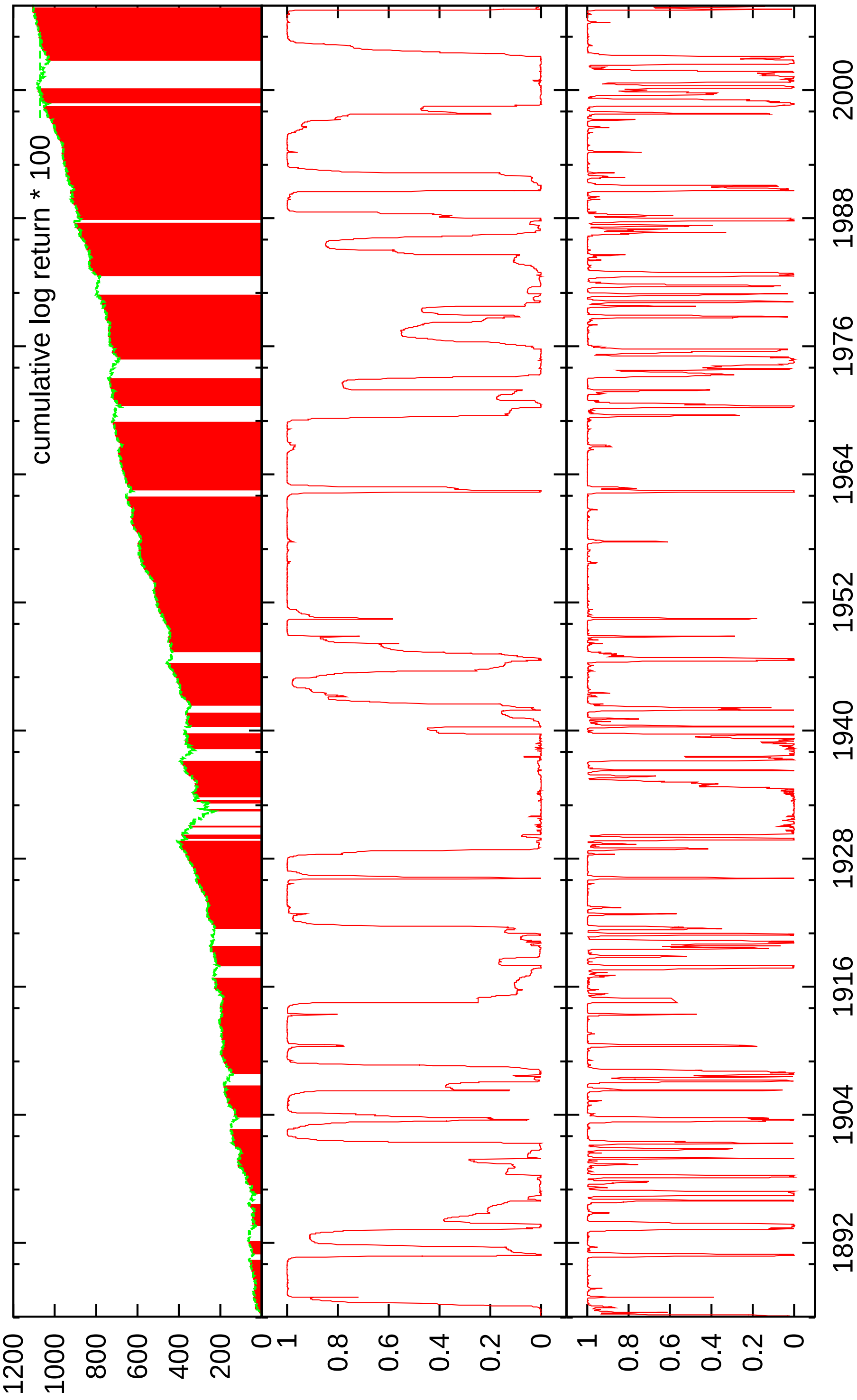

Figure 5: Cumulative Return (LT dating) & Smoothed Bull States from MS-4 and MS-2

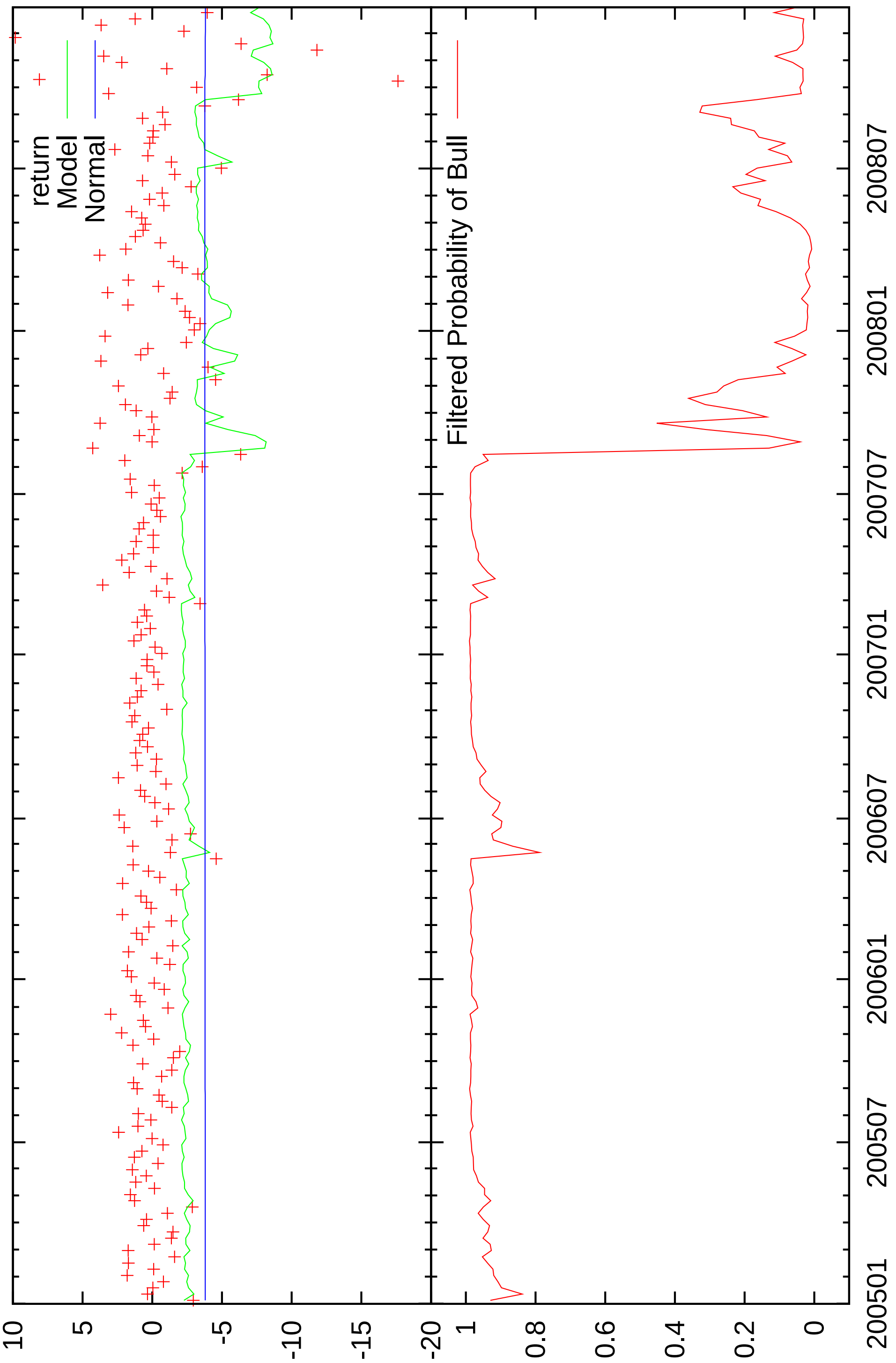

Figure 6: Value-at-Risk from MS-4 and Benchmark Normal distribution