What Drives Markov Regime Switching Behavior of Stock Markets? The Swiss Case

Martin Hess[1] [2]

ITAM, Mexico

April 22, 2002

[1]Instituto Tecnológico Autónomo de México, Av. Camino a Santa Teresa #930, Col. Héroes de Padierna, 10700 Mexico D.F., Mexico. email: mhess@itam.mx, http://ciep.itam.mx/~mhess.

[2]I thank Harris Dellas, Andreas Grünbichler, James Hamilton and Mark Watson for helpful comments. Part of this paper was written when I was affiliated with Study Center Gerzensee and University of Bern, Switzerland.

Abstract

To optimally account for dynamic and nonlinear changes in the stock market return distribution we evaluate competing Markov regime switching model setups for the Swiss stock market. We find that the stochastic movement is optimally tracked by time-varying first and second moments and including a memory effect. Besides the superior dynamic properties this setup exihibits appealing economic interpretations.

Keywords: Markov regime switching, asymmetric stock return distribution, optimal dynamic setup, memory effect.

JEL Codes: G12

1 Introduction

During the last decade, researchers have presented Markov switching regime models as an accurate way of modeling time-variation in stock market return distributions. These models may reveal patterns that go beyond traditional stylized facts. Besides its appealing stochastic properties this methodology excels through its intuitive simplicity and, more importantly, by their empirical ability to accurately track economic movements. Economic interpretations ensure that the success of these models are not simply a product of the choice of the sample but promise to provide persistently good results across different assets and time.

Their nature of not relying on linearities enables discrete state models to accurately capture typical stock market patterns such as jumps or crashes. Such movements have presented a long-standing challenge in our understanding of financial markets as these abrupt changes may involve sharp breaks in the relationships among variables. Other statistical features that lead other methods to fail and that may be captured by Markov switching are asymmetries, fat tails, volatility clustering or mean reversion. Economic explanations for these types of time-variation in a series point into two directions.

The first one attributes nonlinearities to particular behavior of market

participants (i.e. noise traders). There is a large literature that reports that speculative trading may cause fads, bubbles or even market crashes. Flood and Hodrick (1990) point out that any such behavior can be interpreted as evidence of misspecified fundamentals. Based on the seminal paper by Hamilton (1989), Funke, Hall and Sola (1994), van Norden and Vigfusson (1998) and van Norden and Schaller (1999) show that Markov Switching models with unobserved regimes are able to fit these types of stock market behavior well. Dewachter (2001) shows the explanatory power of regime switching models for foreign exchange traders. Furthermore, Jeanne and Rose (2002) link regimes and noise trading in a theoretical approach.

The second strand of economic explanations relates stock market movements to macroeconomic fundamental influences. Numerous studies confirm that the conditional moments of stock returns are business cycle dependent. Cecchetti, Lam and Mark (1990) show that switching in economic growth influences the distribution of stock returns via the dividends. Hamilton and Lin (1996) report that the driving force of conditionally switching moments are economic recessions. Generally, researchers have found a strongly countercyclical effect of real activity on stock return volatility (e.g. Schwert, 1989 or Campbell et al., 2001), whereas there exists no unanimity about the effect of real activity on the mean return (see e.g. Campbell, 1999). Perez-Quiros

and Timmermann (2001) describe the effect of output on higher order moments in a discrete state model and find significant influences. There also exists a broad theoretical literature relating business conditions and return distribution (e.g. Ebell, 2001).

In the literature, the setups of Markov switching models to accurately capture the time-variation of the stock return distribution diverge substantially. On the basis of Hamilton (1989) the simple mean switching setup has been modified in numerous ways but for stock markets the extensions have never been evaluated against each other. Turner, Startz and Nelson (1989) build a model allowing for the mean and variance to differ across the regimes. Hamilton and Susmel (1994) observe that a conditionally heteroskedastic Markov regime-switching model provides a better statistical fit than ARCH models without switching. Since then, the use of the regime-switching method for modelling dynamics and asymmetries in stock market returns has become very popular and the variety of different setups is growing fast (e.g. Chauvet and Potter (2000), Perez-Quiros and Timmermann (2001), Ang and Bekaert (2002) or Ang and Chen (2002)).

While research has repeatedly demonstrated the superiority of the discrete state approach over standard stochastic processes (see e.g. Dewachter, 2001) the issue of what parameters ought to be modeled as regime-switching

has not been extensively discussed. Literature lacks of studies that systematically evaluate competing models focussing on the driving moments of time-variation. The contribution of this study is to fill this gap by comparing different discrete state setups and by analyzing the effects of modeling different parameters as state-dependent for the Swiss stock market. We do not restrict to a goodness-of-fit analysis and to an assessment of the causes of the difference of model accurateness but extend the literature by evaluating the models out-of-sample as there may be enormous differences between how well a model fits in-sample and its forecasting power. A good fit may be therefore of little guidance if the prediction errors of a model are large relative to other setups. A systematic evaluation of competing models also allows us to get a deeper insight of the fundamental factors influencing the time-variation of the return distribution.

We find that within a group of stochastic return processes it is optimal to model time-varying mean and variance in an autoregressive process with one lag. Parallel to the findings of Dewachter (2001) for the foreign exchange market we show that also in the stock market the switching process is more pronounced for the variance than for the mean. Moreover, modeling a memory effect by including information about lagged states is crucial for a switching pattern which ought to be both, economically intuitive and

useful for practitioners. This effect leads to a substantially enhanced regime persistence and to a more reliable identification of the unobserved states.

The paper is organized as follows. Section 2 presents the general Markov switching model and the particular setups that we estimate. Section 3 contains a description of the data and of the employed estimation technique. In section 4 we report the estimation results and present a deeper statistical and interpretational analysis of the obtained model estimations. Section 5 concludes.

2 Methodology

2.1 General Setup

The Markov Switching models we propose are based on Hamilton (1989) whose general class of mixture distributions allows to approximate classes of flexible density functions. These conditional distributions may generate a wider range of skewness and kurtosis values than standard distributions and hence capture particular time-series behavior of a time series as discussed. The mixture density function of the returns r_t is calculated as the state probability weighted conditional density of r_t summed over all K discrete states s_t. Hence, at each time t, r_t will be drawn from a different distribution.

In its general form, the mixture density can be represented as

$$f\left(r_t; I_t\right) = \sum_{j=1}^{K} \Pr\left(s_t | I_t; \theta\right) f\left(r_t | I_t; s_t = j; \alpha\right) \tag{1}$$

where I_t represents the information set in period t, α a set of parameter values characterizing the conditional return density and θ is a vector containing α and vec(P) with P denoting the state transition probabilities of the regime switching process.

We present time variation assuming that the series follows a stochastic process. Thereby, mean return, variance and serial correlations of stock returns may depend on the regimes s_t. The general form of the time series model is

$$\psi(L) r_t = \psi(L) \mu\left(s_t\right) + \sigma\left(s_t\right) \varepsilon_t \tag{2}$$

with $\psi(L) = 1 - \left[\phi^1\left(s_{\widetilde{t}}\right)\right] L - \left[\phi^2\left(s_{\widetilde{t}}\right)\right] L^2 - \ldots - \left[\phi^p\left(s_{\widetilde{t}}\right)\right] L^p$, where L denotes a lag operator and $\widetilde{t}$ does not multiply with L. Various assumptions about the switching behavior of these three parameters combined with different stochastic processes yield a set of competing setups that we evaluate against each other in order to identify the most suitable model. In the following we will treat this general framework more specifically by introducing assumptions about the switching process of the regimes and about the stochastic process of the stock returns.

Let s_t be a random variable with a probability that it takes on value j depending only on its own past realizations. Then $\Pr\{s_t = j | s_{t-1} = i\} = p_{ij}$ denotes the constant transition probability that state j follows after a realization i. All combinations of p_{ij} are collected in the transition probability matrix P. The exact stochastic process of the discrete multiperiod regimes s_t following a first order Markov chain depends on the order of the process for the one-period regimes S_t and on the number of its actual states, N[1].

For the latter, we propose three different model specifications for conditional stock returns. We assume r_t to follow different stochastic processes with a normally distributed error term and two discrete one-period states,

[1] We define the state s_t as a series of consecutive realizations of the actual regime S_t. This allows to transform every K state, $m+1^{st}$ order Markov chain into an equivalent N state, first order model. If $i, k, ..., l$ are past realizations, transition probability p_{ij} equals $P\{S_t = j | S_{t-1} = i, S_{t-2} = k, ..., S_{t-m} = l\}$.

i.e. $S_t \in \{0,1\}$:

$$
\begin{aligned}
r_t &= (\mu_0 + (\mu_1 - \mu_0) S_t) + \left(\phi_0^1 + \left(\phi_1^1 - \phi_0^1\right) S_t\right) r_{t-1} \\
&\quad + (\sigma_0 + (\sigma_1 - \sigma_0) S_t) \varepsilon_t \qquad (3) \\
r_t - \mu_0 - (\mu_1 - \mu_0) S_t &= \left(\phi_0^1 + \left(\phi_1^1 - \phi_0^1\right) S_t\right) (r_{t-1} - \mu_0 - (\mu_1 - \mu_0) S_{t-1}) \\
&\quad + (\sigma_0 + (\sigma_1 - \sigma_0) S_t) \varepsilon_t \qquad (4) \\
r_t - \mu_0 - (\mu_1 - \mu_0) S_t &= \left(\phi_0^1 + \left(\phi_1^1 - \phi_0^1\right) S_t\right) (r_{t-1} - \mu_0 - (\mu_1 - \mu_0) S_{t-1}) \\
&\quad + \left(\phi_0^2 + \left(\phi_1^2 - \phi_0^2\right) S_t\right) (r_{t-2} - \mu_0 - (\mu_1 - \mu_0) S_{t-2}) \\
&\quad + (\sigma_0 + (\sigma_1 - \sigma_0) S_t) \varepsilon_t \qquad (5)
\end{aligned}
$$

where $\varepsilon_t \sim Niid(0,1)$. Note, that the explicit differentiation between models with and without historical regime persistence made by Kim, Nelson and Startz (1998) is implicitly introduced in our setup by the assumptions about the stochastic process. While there is no regime persistence in the first model, stock returns may depend on up to 2 prior regimes, S_{t-1} and S_{t-2}, in the two latter ones. This implies a switching process for S_t of order two and three, respectively. We will refer to the processes in equations (3) to (5) as models a, b and c in the same order.

The number of the multiperiod states s_t following a first order Markov chain is defined as $K = N^h$ where h is the order of the Markov chain for S_t. Given equations (3) to (5) and two actual regimes S_t, i.e. $N = 2$, the

number of discrete states for s_t are $K = 2, 4$ and 8, respectively.

The second model variation besides the conditional distribution of r_t concerns the parameters which we allow to be time dependent. In order to identify the economic origin of the regimes we alternatively let μ, μ and σ^2 or μ, σ^2 and ϕ_m be state-dependent[2]. We denote these three setups *1*, *2* and *3* depending on the number of time-variant parameters. The combination of the order of the Markov process and of the number of time-varying parameters amounts to a total of nine competing model setups which are denoted *a1*, *a2*, etc.

2.2 A Practical Example

To illustrate the approach, we take model b represented by equation (4) as an example. This stochastic process depends on the actual and one lagged regime (i.e. S_t and S_{t-1}) and thus implies the following stochastic process

[2] Hamilton and Susmel (1994) propose a switching ARCH (SWARCH) model allowing the parameters of such a model to stem from different regimes. Their findings of weekly stock market data is that short-run dynamics are governed by an ARCH process but that these effects almost completely die out after one month. This motivates our choice of making the unconditional variance state-dependent.

for the discrete regimes:

$$
\begin{aligned}
s_t &= 1 \text{ if } S_t = 0 \wedge S_{t-1} = 0 \\
s_t &= 2 \text{ if } S_t = 1 \wedge S_{t-1} = 0 \\
s_t &= 3 \text{ if } S_t = 0 \wedge S_{t-1} = 1 \\
s_t &= 4 \text{ if } S_t = 1 \wedge S_{t-1} = 1
\end{aligned} \tag{6}
$$

In this case, we define the transition probability matrix P of this four-state, first-order process as

$$
P = \begin{bmatrix} p_{11} & 0 & p_{13} & 0 \\ 1-p_{11} & 0 & 1-p_{13} & 0 \\ 0 & 1-p_{42} & 0 & 1-p_{44} \\ 0 & p_{42} & 0 & p_{44} \end{bmatrix} \tag{7}
$$

where p_{ij} denotes the transition probability $\Pr(s_t = j | s_{t-1} = i)$[3].

Model c includes two lagged regimes which implies eight discrete states and hence, a dimension of P of 8×8. Analogously to model b we define the states in model c as $s_t = 1, ..., 4$ by including $S_{t-2} = 0$ to states 1 to 4 in equation (6) and as $s_t = 5, ..., 8$ when including $S_{t-2} = 1$ instead. In the case of model a, s_t and S_t are identically the same[4]. As these discrete

[3] A zero probability is denoted by 0 and assigned to states that cannot follow each other.

[4] For homogeneity reasons we redefine the regimes of model a to take on similar values

states are unobserved, we assume each state to prevail at time t with the conditional probability $\Pr(s_t = j | I_t, \theta)$, where the information set I_t consists of the vector of past returns $\{r_{t-1}\}$.

Model *b2* thus represents demeaned stock returns following an AR(1) as in equation (4) with regime dependent moments μ_{S_t} and σ_{S_t} of the Gaussian distribution but with fixed serial correlation ϕ_{S_t}. This model hence describes four different conditional distributions for r_t:

$$\begin{aligned} r_t | I_{t-1}; s_t &= 1; \theta \sim N\left(\mu_0, \sigma_0^2\right), \\ r_t | I_{t-1}; s_t &= 2; \theta \sim N\left(\frac{\mu_1 - \phi\mu_0}{1-\phi}, \sigma_1^2\right), \\ r_t | I_{t-1}; s_t &= 3; \theta \sim N\left(\frac{\mu_0 - \phi\mu_1}{1-\phi}, \sigma_0^2\right), \\ r_t | I_{t-1}; s_t &= 4; \theta \sim N\left(\mu_1, \sigma_1^2\right). \end{aligned} \tag{8}$$

The evaluation of the model setup that yields the best fit is guided by the two dimensions of variations that distinguish the different models. First, we identify the optimal number of time-varying parameters with a likelihood ratio test separately for each of the three stochastic processes a, b and c[5].

as for models b and c, i.e. $s_t = 1$ if $S_t = 0$ and $s_t = 2$ if $S_t = 1$.

[5] Given the rich empirical evidence we take the switching behavior of stock returns as a fact and do not test whether it is present at all. Garcia (1998) suggests a test method for the presence of regime switching. This test uses non-standard asymptotic likelihood ratio distribution as the transition probabilities are not identified under the null hypothesis of no regime switching.

Using the characteristics of the asymptotic χ^2_J-distribution of random excess returns under the null hypothesis, we calculate the likelihood ratio

$$LR = 2\left[\mathcal{L}\left(\widehat{\theta}\right) - \mathcal{L}\left(\widetilde{\theta}\right)\right] \sim \chi^2_J \tag{9}$$

where $\mathcal{L}\left(\widetilde{\theta}\right)$ and $\mathcal{L}\left(\widetilde{\theta}\right)$ denote the log likelihood under the unrestricted maximum likelihood estimator and under the null. J denotes the number of restrictions. Once the optimal number of regime-dependent parameters is robustly identified for each stochastic return process we then run tests based on the Akaike information criterion to determine the optimal stochastic model.

After the identification of the best fitting model we run out-of-sample simulations to evaluate the forecasting accurateness of the different setups. Hence, we account for the fact that an optimal model fit does not necessarily imply a good out-of-sample performance. We forecast the mean with all nine models and compare it to a standard AR(1) forecast and to a random walk[6]. We also compare the results of a simple trading strategy of investing in the stock market if the forecasted return is larger than the risk free rate and being invested in the risk-free asset, otherwise.

[6] To remain consistent with our estimates we generate the return forecasts on the basis of the optimal in-sample regimes.

3 Data and Estimation

3.1 Data description

Our analysis uses the Datastream Swiss stock market index sampled on a monthly basis from January 1973 to December 1998 representing 312 data points. By the end of the observation period this broad value-weighted index includes 265 stocks with over 99% of the market capitalization. Returns are computed as the first difference of end-of-month log prices. Monthly mean and standard deviation of the series over the observed interval are 0.57% and 4.80%, respectively. The returns exhibit a significant first-order serial correlation. The return distribution is different from Gaussian at all significance levels with a skewness of -1.11 and a kurtosis of 8.40. This points to the typical stock return pattern of fat tails which are more pronounced to the left hand of the distribution.

We choose monthly returns due to the presence of more noise at higher frequencies which makes it more difficult to isolate cyclical variations and hence, obscures the analysis of the driving moments of switching behavior. Given the high number of parameter estimates the number of observations for data at lower frequency seems insufficient. The choice of monthly frequency of equity returns is also actual practice in comparable studies.

For the simulation of the investment strategies we select the 1-month Euro Swiss francs rate as the return on a riskfree investment. The mean return and standard deviation from this risk-free investment are 0.35% and 0.73%, respectively.

Figure 1 displays the absolute returns as a variance proxy of the market index. The volatility clustering over the 26 year record is very distinct and should be well fitted by an estimation of a Markov-switching process.

Figure 1: Absolute Returns of the Swiss Stock Market

The switching pattern in the conditional mean returns (not displayed) is expectedly less pronounced. This confirms Dewachter (2001) who observes that time-variation mean stock returns is frequently dominated by time variation in the variance. However, modelling mean returns to be regime-dependent has advantages. Most importantly, it allows an intuitive interpretation of bullish and bearish markets and hence, is of interest for portfolio management issues[7].

[7] Hess (2001) presents an illustrative example of how regimes may enter an asset allocation process as conditioning variables. He finds a large information content contained in regime-dependent models for the cross-section of stock returns.

3.2 Estimation procedure

We estimate the parameters θ by maximizing the log likelihood function assuming a Gaussian conditional return distribution $f(r_t|I_{t-1}, s_t = j; \alpha)$.

$$\mathcal{L}(s_t, \theta, \pi) = \log \pi + \sum_{t=1}^{T}\sum_{s=1}^{K} \log \left(\begin{array}{c} \Pr\left(s_t = j|r_{t-1}, \theta\right) \\ \times f(r_t|I_{t-1}; s_t = j; \alpha) \end{array} \right). \qquad (10)$$

where T is the number of observations in the sample.

We perform the estimation in an iterative process where alternately the parameter vector $\widehat{\theta}$ is estimated with the EM-algorithm and the conditional state probabilities $\Pr\left(s_t = j|I_{t-1},\theta\right)$ for $t = 1, 2, ..., T$ are generated in a state space environment. To optimize accuracy of these conditional probabilities we compute smoothed inferences based information until period T.

Our main objective, however, is to identify the conditional smoothed probability that the stock market is in state $S_t = i$ at period t, $\Pr\left(S_t = i|I_T,\theta\right)$. To compute them we simply sum up the two related marginal probabilities. In our previous example of a first-order stochastic process of a demeaned data series, we get:

$$\begin{aligned} \Pr\left(S_t = 0|I_T,\theta\right) &= \Pr\left(s_t = 1|I_T,\theta\right) + \Pr\left(s_t = 3|I_T,\theta\right) \qquad (11) \\ \Pr\left(S_t = 1|I_T,\theta\right) &= \Pr\left(s_t = 2|I_T,\theta\right) + \Pr\left(s_t = 4|I_T,\theta\right) \end{aligned}$$

Note, that very often S_t is of more interest than s_t as it describes the actual

regime of the stock market. Based on $\Pr(S_t = 0|I_T,\theta)$ and assuming a good model fit, financial analysts may identify periods of bullish and bearish periods and calm and turbulent periods in setups a and b, respectively.

4 Results

4.1 Model evaluation

Table 1 displays the estimates of all elements of θ for each of the models estimated for the period 1973:01 to 1998:12 at monthly frequency.

Table 1: Parameter Estimates for Regime Switching Models

A first look at the different values in table 1 shows various interesting facts. First, the conditional mean estimates in the models of group *1* are significantly different than in the other groups. This difference to models where additional parameters are regime-switching points to the fact that the switching behavior of μ might not be sufficient to optimally describe stock market returns. Within group *1* we observe two different behaviors. Model *a1* does not seem to accurately replicate the behavior of the stock market as the two means are quite close to each other and the transition probabilities indicate no persistence at all. Values close to 0.5 rather point to a random behavior than to two distinct states. Another sign of the inappropriateness

of model *a1* is shown in figure 2 displaying values for $\Pr(S_t = i|I_T;\theta)$ that are also close to 0.5 which points to a very bad identification of the states. As the two means are relatively close to each other, all the conditional probabilities in $\Pr(S_t = i|I_T;\theta)$ are close to 0.5 which implies that this setup is unable to accurately identify the states[8,9].

Figure 2: Conditional State Probabilities for $S_t = 0$

The estimates of models *b1* and *c1* show that the inclusion of lagged regimes substantially improves the fit of the states by estimating conditional state probabilities $\Pr(S_t = i|I_T;\theta)$ that are either close to zero or one. The different characteristics of the conditional means and of the transition probabilities compared to model *a1* in table 1 clearly emerges from the first column of panels of figure 2. The models with lagged regimes display that the numeric maximization of the log likelihood function identifies one sin-

[8] When the model identifies a high return market the corresponding conditional probability is on average 54.75% and hence reveals a very low ability to distinguish it from the low return state. The average conditional probability for the low state return is 52.94% and lies far off the 100% which would mean perfect observability.

[9] For illustration purposes, the solid line represents regimes that may be considered as being observed. We construct this series by assigning probability one to the regime which is more likely to prevail, i.e. we choose a cutoff point for defining the the observed regimes of $P(S_t = j|I_T;\theta)$=0.5.

gle state during most of the time and four isolated crash events[10]. As the state switches back immediately after a crash event these models produce regimes that clearly differ from the idea of periods of a bullish or a bearish markets. Even though the models *b1* and *c1* almost perfectly identify the two states S_t and hence, yield a good fit they are not very informative because of the few observations of the low return state. Thus, the difference to time-invariant models is small. Furthermore, these models do not represent a state in the spirit of a persistent regime nor are they forecastable from which the use of investment decisions is limited[11].

The rest of the models all produce quite similar and hence, robust results among each other with conditional means that lie between the results of the models with fixed variance and serial correlation. Common to all these models is the fact that estimations identify regimes of jointly low returns and high variances and vice versa. This result is largely consistent with the literature (e.g. Glosten, Jagannathan and Runkle, 1993, van Norden and Schaller, 1999). Thus, the stock market tends to exhibit smooth price

[10] Specifically, the oil price shock, the crashes in 1987 and 1990 and the Latin American crisis in 1998 are isolated events forming one regime with a monthly mean return of less than -15%.

[11] The difficulty to identify the states in *a1* and the unpredictability of the crashes provide also support for the efficiency of the stock markets, i.e. returns are unpredictable.

increases and turbulent periods of falling prices. While the difference in return variation is small the standard deviation of stock returns that are between 2.3 and 2.7 times higher in high-volatility than in low-volatility regimes. As all Markov switching models with regime persistence capture these asymmetries in a similar way we need to run specification tests for identifying the optimal setup.

The results of the likelihood ratio tests confirm that for all stochastic processes the homoskedasticity assumption (group *1*) is too restrictive, i.e. a regime-switching setup with the mean return as the only state-dependent variable is underspecified. We therefore conclude that a time-varying variance is a necessary condition to appropriately describe the fluctuations of the Swiss stock market.

Table 2: Specification Test Statistics

Comparing groups *2* and *3* we observe that the assumption of fixed autoregressive parameters is not restricting. The estimates of time-varying autocorrelation coefficients tend to be close to each other. Moreover there is no indication that either the sign or the value of the serial correlation coefficients would be linked to one particular state. Rather, they depend on the choice of the stochastic process for r_t. Therefore, the non-significance in the likelihood ratio test statistics for the models *a2* to *c2* does not surprise.

From this we conclude that an appropriate description uses time-varying mean and variance but constant autoregressive coefficients represented by the middle column of panels in figure 2.

Evaluating the three stochastic processes within group *2* by the Akaike information criterion (AIC) reveals that equation (5) does not represent an ideal stochastic process. Its AIC value of 3.15 is higher than the ones of the two AR(1) equations with both a value of 3.13. Although model *a2* has the advantage that less parameters need to be estimated we identify model *b2* as an ideal setup, mainly because of the regime characteristics.

Model *b2* shows considerably larger regime persistence, and hence there are more observations of regime $S_t = 1$. The ratio of observations of regime $S_t = 0$ to regime $S_t = 1$ observations is 74/237 in model *b2* as compared to 23/288 in model *a2* implying a reasonably long sample when investigating portfolio management issues in each state separately. Moreover, short regime durations also lack the interpretation of cyclical behavior and provide therefore less possibilities for economic interpretation. As illustrated in Hess (2001) this persistence allows for a much broader use of the switching regime methodology for portfolio management as the states are more easily forecastable and both states are sufficiently frequent.

A further advantage of model *b2* is displayed by figure 2 which shows

that the identification of the regimes in model *b2* is more certain than in *a2*[12]. Generally, the average conditional probability of the prevailing state $\Pr(S_t = i|I_T,\theta)$ is higher for all models with lagged states, illustrated by a much steadier graph in the second and third row panels of figure 2. The result that the states are more distinct appears to stem from the memory effect introduced by the conditional demeaning of the lagged returns in equations (4) and (5). To our knowledge, the effect of introducing regime switches in lagged returns has not yet been empirically analyzed in the field of financial economics.

Modeling stock returns as stochastic processes is mostly justified by the presence of noise traders which through slow information processing introduce serial correlation. When including lagged regimes we assume that the noise traders not just look at past returns but also set them into relation with the regimes at that time. Hence, they are able to make selective interpretations of past returns depending on the value of $(\mu_1 - \mu_0) S_{t-i}$ which indicates whether a bullish or a bearish regime has prevailed and thus informs about the expected returns at $t - i$.

In summary, we favor a model with time-varying mean and variance

[12] For model *b2* $\Pr(S_t = 0|I_T,\theta)$ and $\Pr(S_t = 1|I_T,\theta)$ are on average 93.14% and 89.63%, respectively, while for model *a2* the corresponding figures are 90.13 and 81.62%.

which includes a memory effect in the regimes. Despite a similar statistical evaluation, the insights of model *b2* can be expected to be more relevant than of model *a2* due to its economic interpretations and its higher accurateness for portfolio management purposes.

4.2 Characteristics of Optimal Model

The difference between the mean returns across the regimes is 1.19% which corresponds to an important annualized mean difference of over 14% setup *b2*. Due to the relatively large standard errors of the mean estimates the regime switches in mean returns are not significant. This observation is in line with the estimates by Ang and Bekaert (2001) for the US market but does not confirm the findings of van Norden and Schaller (1999) who detect significance. The volatility is 2.3 times higher during turbulent than during calm periods. This factor estimate is expectedly somewhat lower than in comparable studies as the estimated regimes do not exhibit patterns of short and isolated variance bursts.

The durations of the two assumedly observed regimes are 1.2 years and 3.4 years for $S_t = 0$ and $S_t = 1$, respectively[13]. The scarcity of switches in the center panel of figure 2 illustrates the estimated regime persistence and

[13] Hamilton and Susmel (1994) report a similar regime persistence for the US market while most other studies report one state of only short duration.

the similarities with figure 1 suggests that the regimes are mainly driven by volatility.

The optimal regime switching setup also reveals economic relevance of the chosen model. Narratively interpreted, the center panel of figure 1 shows that the first structural break coincides with the end of the oil price shock and stock markets pick up and enjoy ten years of smooth and steady growth. Increased volatility due to large gains is shown by the regime switch in 1985 and the peaks in 1987 and 1990 indicate the respective crashes. The latest regime has been prevailing since 1997 when markets got to their all time high and weakened abruptly due to the Asian and Latin American crisis. The regimes are not clearly identified when $\Pr(S_t = i|I_T;\theta)$ is neither close to 1 nor to zero as for example in the summer rally of 1987 preceding the crash or the smooth decline in 1994.

4.3 Forecasting Based Strategy Results

To test what the in-sample regimes imply for the performance of stock return prediction we simulate portfolio strategies based on forecasts conditioned on a regime switching model. Table 3 presents the summary statistics for the one-month-ahead predictions of all setups versus an AR(1) process and a random walk as a benchmark. It also reports the outcomes of a simple

trading strategy simulation with either a 100% investment in a risk-free asset or in the stock market index based on the return forecast.

Table 3: Conditional Forecasting Models

Generally, the difference in forecasting performance, measured as root mean squared error with respect to the true returns, between regime switching models and with a standard AR(1)-model is economically insignificant, while they all seem to slightly outperform a random walk. The difference, however, is too small to be informative for market participants and hence not significant enough to reject the efficient market hypothesis[14]. Model *b2*, identified as the most appropriate in-sample market model, is interestingly the second worst of all regime-switching models in terms of out-of-sample forecasting ability. This again is evidence of the fact that a good in-sample outcome does not imply much about out-of-sample results.

Investors who follow a trading strategy of switching between risk-free assets and stocks, face big differences among the models with an average monthly return ranging from 0.30% to 0.81% at marginal transaction costs of 1% per trade. Without transaction costs, the strategy results are more homogenous indicating that the implied strategies between the models are

[14] Basing the analysis on regime forecasts and not on optimized states would further deteriorate the results.

varying a lot with respect to the trading activity. Interestingly, all the models but one beat the strategy of simply holding the stock market. Model *b2* again performs poorly against the models just allowing their means to be state-dependent and just being able to identify the crash-periods (models *b1*, *c1*). While these two models do hardly any better in forecasting future return exactly, they seem to be quite good in market timing. Apart from this exception, the models perform in general poorly as the outperformance over a random walk is very modest. These results show that introducing regime-switches does not significantly enhance the power of stochastic models in explaining the time-series of stock returns.

5 Summary

Recent literature has repeatedly reported the statistical and economic appropriateness of regime switching models to analyze and describe the stochastic behavior of stock returns. However, the literature lacks analyses that jointly investigate the driving moments of time-variation and systematic evaluation of competing model setups.

This study attempts to fill this gap by analyzing monthly Swiss stock market returns over a 26-year sample by comparing alternative regime switching models. As a good model fit does not ensure accurate out-of-the sample

performance we extend the study by empirically analyzing conditional return forecastability. The outcomes of in-sample simulations based on the optimized regimes and of out-of-sample predictions are then compared to standard forecasts.

Stock returns modeled as a stochastic process with memory effect and constant autoregressive parameter yield the best fit among nine different setups. Besides the good fit, this model exhibits the highest state persistence. The regimes contain a high economic content which makes them a valuable and easily tractable information variable for conditional portfolio strategies. However, the forecasting accurateness of this setup is relatively poor. This suggests that the choice of the regime switching model setup needs to be carefully chosen in response to the objective of the study. Further research with a larger variety of models may possibly help identifying a model with the desirable properties of both, successful in-sample and out-of-sample performance.

References

[1] Ang, A. and G. Bekaert, 2002, International Asset Allocation with Regime Shifts, *Review of Financial Studies,* forthcoming.

[2] Ang, A. and J. Chen, 2002, Asymmetric Correlations of Equity Portfolios, *Journal of Financial Economics*, 63(3), 443-494.

[3] Campbell, J. Y., 1999, Asset Prices, Consumption, and the Business Cycle, in: Taylor J. and M. Woodford, eds., *Handbook of Macroeconomics*, Amsterdam, North-Holland.

[4] Campbell, J. Y., M. Lettau, B. G. Malkiel and Y. Xu, 2001, Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk, *Journal of Finance*, 56(1), 1-43.

[5] Chauvet, M. and S. Potter, 2000, Coincident and Leading Indicators of the Stock Market, *Journal of Empirical Finance*, 7(1), 87-111.

[6] Cecchetti, S. G., P.-S. Lam and N. C. Mark, 1990, Mean Reversion in Equilibrium Asset Prices, *American Economic Review*, 80(3), 398-418.

[7] Dewachter, H., 2001, Can Markov Switching Models Replicate Chartist Profits in the Foreign Exchange Market?, *Journal of International Money and Finance*, 20(1), 25-41.

[8] Ebell, M. C., 2001, Why Are Asset Returns More Volatile during Recessions? A Theoretical Explanation, Study Center Gerzensee Working Paper 01.

[9] Flood, R. P. and R. J. Hodrick. 1990, On Testing for Speculative Bubbles, *Journal of Economic Perspectives*, 4(2), 85-101.

[10] Funke, M., S. Hall and M. Sola, 1994, Rational Bubbles During Poland's Hyperinflation: Implications and Empirical Evidence, *European Economic Review*, 38(6), 1257-1276.

[11] Garcia, R., 1998, Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models, *International Economic Review*, 39(3), 763-88.

[12] Glosten, L. R., R. Jagannathan, and D. E. Runkle, 1993, On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48(5),1779-1801.

[13] Hamilton, J. D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, 57, 357-384.

[14] Hamilton, J. D. and R. Susmel, 1994, Autoregressive Conditional Heteroskedasticity and Changes in Regime, *Journal of Econometrics*, 64(2), 307-333.

[15] Hamilton, J. D. and G. Lin, 1996, Stock Market Volatility and the Business Cycle, *Journal of Applied Econometrics*, 11(5), 573-593.

[16] Hess, M. K., 2001, Timing and Diversification: A State-Dependent CAPM Approach, mimeo.

[17] Jeanne, O. and A. K. Rose, 2002, Noise Trading and Exchange Rate Regimes, *Quarterly Journal of Economics*, forthcoming.

[18] Kim, C. J., C. R. Nelson and R. Startz, 1998, Testing for Mean Reversion in Heteroskedastic Data Based on Gibbs-Sampling-Augmented Randomization, *Journal of Empirical Finance*, 5(2), 131-54.

[19] van Norden, S. and H. Schaller, 1999, Speculative Behavior, Regime Switching and Stock Market Fundamentals, in: *Nonlinear Time Series Analysis of Economic and Financial Data*, Philip Rothman, ed., Kluwer, 321-356.

[20] van Norden, S. and R. Vigfusson, 1998, Avoiding the Pitfalls: Can Regime-Switching Tests Reliably Detect Bubbles?, *Studies in Non-*

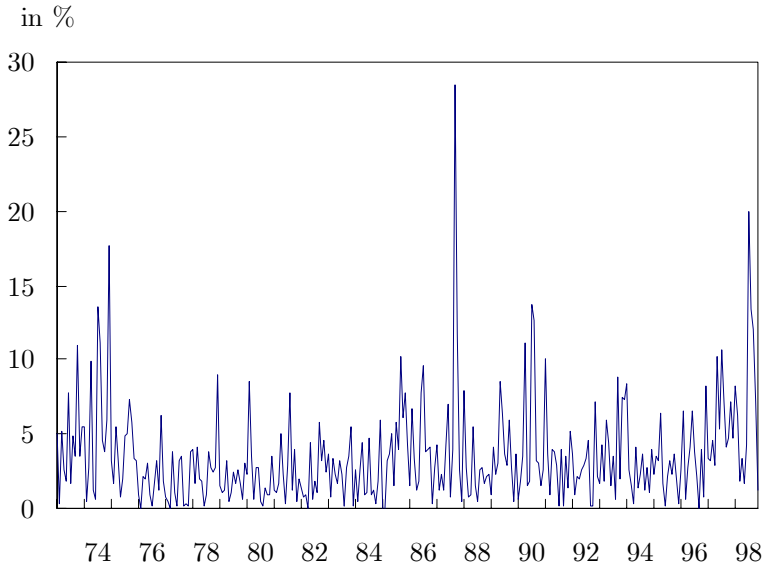
Linear Dynamics and Econometrics, 3(1), 1-22.

[21] 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.

[22] Schwert, G. W., 1989, Why Does Stock Market Volatility Change over Time?, *Journal of Finance,* 44(5), 1115-1153.

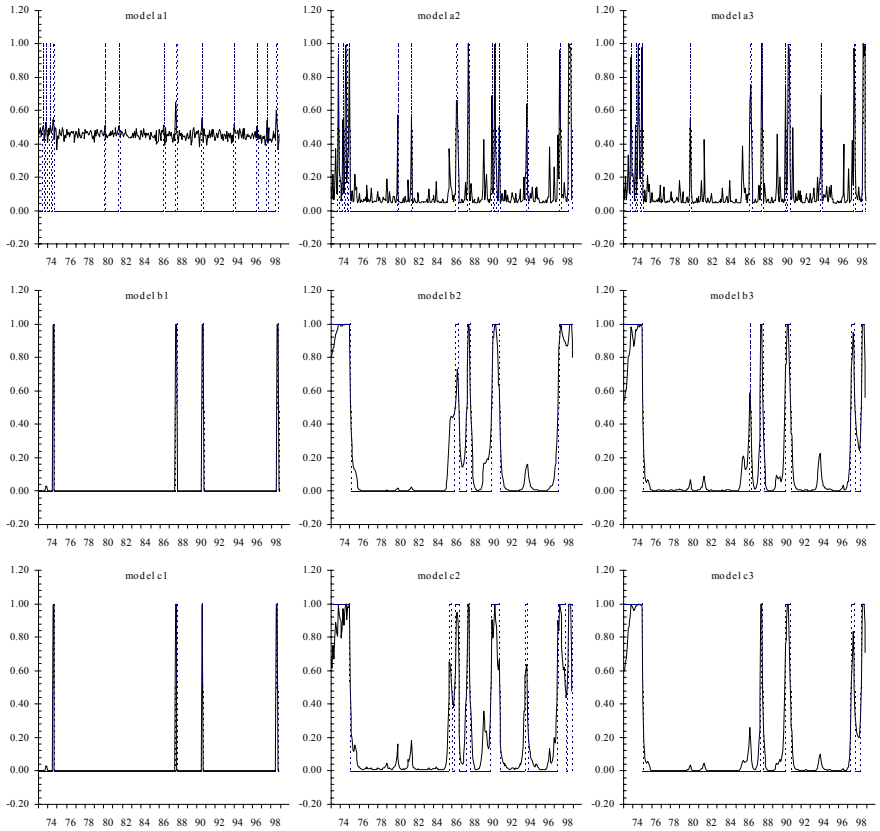
[23] Turner, C. M., R. Startz and C. R. Nelson, 1989, A Markov Model of Heteroscedasticity, Risk and Learning in the Stock Market, *Journal of Financial Economics*, 25(1), 3-22.

Figure 1: Absolute Returns of the Swiss Stock Market



Note: The graph displays absolute values of monthly log returns of the Datastream Swiss stock market index for the sample 1973:01 to 1998:12.

Figure 2: Conditional State Probabilities for $S_t = 0$



Notes: The displayed graphs display conditional regime probabilities for Swiss stock market returns for the sample 1973:01 to 1998:12. The regime-switching models *a1* to *c3* are described in section 2.1. The solid line shows the smoothed conditional state probabilities $\Pr(S_t = 0|I_T;\theta)$. The dashed line shows the assumedly observed state probabilities of the stock market $\Pr^*(S_t = j|I_T;\theta)$. It takes on the value one if $\Pr(S_t = 0|I_T;\theta)$ is larger

than the cutoff probability 0.5, and zero otherwise.

Table 1: Parameter Estimates for Regime Switching Models

	Model								
	a1	*b1*	*c1*	*a2*	*b2*	*c2*	*a3*	*b3*	*c3*
μ_0	0.24	-15.51	-15.46	-1.64	-0.30	-0.80	-1.73	-1.79	-3.11
	(0.18)	(1.14)	(0.12)	(0.74)	(2.86)	(0.16)	(0.59)	(2.08)	(0.17)
μ_1	0.71	0.99	0.99	0.88	0.89	0.96	0.83	1.00	1.06
	(0.24)	(0.24)	(0.24)	(7.54)	(5.90)	(0.10)	(0.35)	(0.42)	(0.11)
σ_0	4.78	4.03	4.03	8.98	7.38	7.97	9.00	8.79	9.20
	(0.54)	(0.23)	(0.23)	(1.16)	(0.59)	(0.64)	(1.32)	(0.50)	(0.74)
σ_1				3.33	3.22	3.17	3.34	3.41	3.49
				(0.18)	(0.26)	(0.25)	(0.21)	(0.19)	(0.28)
$\phi_{1,0}$	0.13**	-0.04	-0.03	0.13	0.11*	0.10**	0.04	0.05	0.02
	(0.06)	(0.05)	(0.02)	(0.18)	(0.06)	(0.02)	(0.32)	(0.06)	(0.22)
$\phi_{1,1}$							0.16	0.05	0.05
							(0.37)	(0.05)	(0.37)
$\phi_{2,0}$			-0.02			-0.08**			-0.09
			(0.02)			(0.02)			(0.22)
$\phi_{2,1}$									0.05
									(0.37)
p_{1j}	0.46	0.00	0.00	0.26	0.99	0.67	0.26	0.84	0.98
	(0.03)	(0.50)	(0.00)	(0.23)	(0.20)	(0.38)	(0.23)	(0.28)	(0.29)
p_{2j}	0.54	0.00	0.00	0.86	0.01	0.45	0.87	0.03	0.00
	(0.03)	(0.11)	(0.00)	(0.21)	(0.26)	(0.41)	(0.21)	(0.19)	(0.00)
p_{3j}		1.00	1.00		1.00	0.91		0.95	0.42
		(0.50)	(0.65)		(0.20)	(0.42)		(0.28)	(0.96)
p_{4j}		0.01	0.00		0.04	0.04		0.03	0.46
		(0.11)	(0.11)		(0.26)	(0.41)		(0.19)	(0.45)
p_{5j}			0.00			0.75			0.99
			(0.00)			(0.38)			(0.29)
p_{6j}			0.00			0.60			0.00
			(0.00)			(0.41)			(0.00)
p_{7j}			1.00			0.88			0.52
			(0.65)			(0.42)			(0.96)
p_{8j}			0.01			0.03			0.60
			(0.11)			(0.41)			(0.45)

Notes: The displayed figures are maximum likelihood estimates for Swiss stock market returns for the sample 1973:01 to 1998:12.The regime-switching models *a1* to *c3* are described in section 2.1. μ: mean, σ: standard deviation, ϕ_i: i-th autoregressive coefficient. For state-dependent coefficients 0 and 1 subscripts denote the regime. Coefficients with values only in the subscript 0 row are time-invariant. p_{ij}: transition probability of

discrete states $\Pr(s_t = j | s_{t-1} = i)$. For each model the displayed transition probabilities represent a subset of vec(P), where P denotes the transition probability matrix. Only the first unrestricted estimates of each column of P, denoted by $s_t = j$, are displayed whereas both, the values restricted to zero and the values $1 - p_{ij}$ are omitted. Standard errors in parenthesis. ** and * denote significance at the 95% and a 90% level, respectively.

Table 2: Specification Test Statistics

	Model								
	a1	*b1*	*c1*	*a2*	*b2*	*c2*	*a3*	*b3*	*c3*
$\mathcal{L}$	-925.92	-895.84	-895.48	-896.18	-892.38	-891.56	-896.17	-891.73	-889.94
LR	59.50**	6.94**	7.85**	0.03	1.29	3.25	-	-	-
AIC	3.13	3.07	3.16	3.13	3.13	3.15	3.13	3.14	3.55

Notes: The displayed figures are estimated for Swiss stock market returns for the sample 1973:01 to 1998:12. The regime-switching models *a1* to *c3* are described in section 2.1. $\mathcal{L}$ = log likelihood (see equation (10)); AIC = Akaike information criterion, LR = likelihood ratio. The likelihood ratios LR reported for the restricted models is calculated with the unrestricted model of the same stochastic process (i.e. a, b, or c) that contains one additional state-dependent parameter. E.g. the likelihood ratio of *a1* is calculated from the log likelihoods of *a1* and *a2*. ** and * denote significance at the 95% and a 90% level, respectively.

Table 3: Conditional Forecasting Models

Model	1-step-ahead Forecast		Trading Strategy no transaction costs				Trading Strategy with transaction costs			
	RMS	MAE	μ	σ	S	EWS	μ	σ	S	EWS
RW	6.31	4.57	0.70	3.43	0.20	70.05	0.47	3.43	0.14	-16.12
$AR(1)$	4.75	3.41	0.74	3.54	0.21	89.66	0.55	3.55	0.15	6.50
a1	4.75	3.36	0.74	3.54	0.21	89.66	0.55	3.55	0.15	6.50
b1	4.61	3.30	0.89	4.37	0.20	177.50	0.81	4.38	0.19	117.44
c1	4.61	3.30	0.83	4.38	0.19	132.56	0.78	4.40	0.18	97.18
a2	4.75	3.36	0.75	3.55	0.21	97.57	0.57	3.56	0.16	12.05
b2	4.85	3.43	0.62	3.77	0.16	29.74	0.51	3.79	0.13	-9.49
c2	4.87	3.47	0.49	3.69	0.13	-12.48	0.30	3.71	0.08	-51.54
a3	4.76	3.36	0.74	3.48	0.21	94.35	0.55	3.48	0.16	8.10
b3	4.82	3.42	0.67	3.73	0.18	50.28	0.59	3.74	0.16	18.93
c3	4.79	3.39	0.62	3.90	0.16	26.51	0.56	3.92	0.14	6.13
optimal	-	-	2.21	2.57	0.86	19461.39	1.98	2.60	0.76	9636.70

Notes: All displayed figures are estimated for Swiss stock market returns for the sample 1973:01 to 1998:12. RW = random walk, $AR(1)$ = unrestricted, the regime-switching models *a1* to *c3* are described in section 2.1. The measures for 1-step-ahead forecast accurateness RMS and MAE denote root mean squared error relative to the true return (in %) and mean absolute error (in %), respectively. The trading strategy consists in investing 100% of the assets either in stocks or in risk-free investments, depending on the higher forecasted return. Transaction costs are 2% for a round trip. μ = mean; σ = standard deviation. S = Sharpe ratio; EWS = end wealth surplus relative to a buy-and-hold strategy (in %). Same result for models $AR(1)$ and *a1* as in no period, a different investment has been taken. 'optimal' refers to a strategy under perfect foresight. We assume a random walk with a drift term equal to the historic performance measured by a

moving average over 24 periods. The innovations are N(0,σ^2) where the estimate of σ^2 is the historic volatility over 24 months.