Regime Switching Models in Stock Market Returns and the Economic Cycle

BY

YU CHEN
B.S., Zhejiang University, 2000

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Business Administration
in the Graduate College of the
University of Illinois at Chicago, 2012

Chicago, Illinois

Defense Committee:

Stanley L. Sclove, Chair and Advisor
Rong Chen, Rutgers University
Gilbert W. Bassett Jr., Finance
Jin Man Lee, Economics
Fangfang Wang

This dissertation is dedicated to my parents, for their love and support.

ACKNOWLEDGMENT

I would like to express my gratitude first to my advisor, Professor Sclove, for his unwavering support and mentoring during these years. My deep gratitude to Professor Rong Chen and Professor Gilbert Bassett, for their support and the great opportunities to work on various projects that opened my eyes. I want to thank Dr. Jim Man Lee and Dr. Fangfang Wang, for their valuable comments and suggestions on my dissertation. I also want to thank Professor Lon-Mu Liu and Professor Houston Stokes, for their teaching and advising in the early stage of my dissertation. I am grateful to all my colleagues in the Business Statistics program, for the intellectual, pleasant and encouraging environment, and for the great times we have had. Thank you all!

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation is the summary of my research in the field of regime switching effects on the US equity market and US economy. We used regime switching models on market index and factors to capture financial market condition changes. We then tested the relationship changes between different assets, and constructed a portfolio rebalance method that uses the market regime information. Also, we applied regime switching models to US GDP and GDP components, to find ways to interpret and predict the US business cycle, especially recessions.

In Chapter 1, an introduction is provided for each part of the work, with summary of my contribution, and forms of models and algorithms being used. More detailed backgrounds, literature views, procedures and results will be presented in the later chapters.

Chapters 2 is focused on regime switching models in the financial markets. In the first part we apply the regime switching model on capital asset pricing model (CAPM). The regimes are defined by applying Hidden Markov Model (HMM) (Hamilton, 1988) on the market index. And each of the regimes represents the market condition with high volatility and low volatility separately. With the regimes defined, we analyzed the relationship changes between a group of stocks on the return, volatility and the return and volatility after taking out the market effects, and observed the relationship changes with the market regime changes.

In the second part of Chapter 2, we applied HMM on Factor Models, where regimes are defined by the Fama French three factor model (Fama and French, 1992) and momentum factor. The estimated parameters under each regime also represent to high and low volatility periods. We tested the correlation changes between these factors, and found there are periods these factor are more correlated than usual, which can be of consideration when applying these factors to do prediction.

In the third part of Chapter 2, with the facts that the HMM regimes on market index and factors can represent different market conditions, and the relationship between assets change with such regimes. We constructed a simple method to see if such regime changes can be used on portfolio rebalance decision. We find that, with the regime into consideration, we can improve the Markowitz portfolio optimization results in a portfolio rebalance.

Chapter 3 is focused on the regime switching models on US economy. First, we applied HMM on US GDP growth rate. We match the regime probability and see the low growth rate regimes matches with NBER announced recession periods very well. Secondly, we use the same HMM technique on the US GDP components. With the regime probability series estimated by HMM, we compared the components regimes with the whole GDP regimes. The purpose is to find the possible leading components that can be used to predict the whole US GDP, or specifically the recession. We calculated various conditional probabilities between the GDP regime probabilities and lagged GDP components' probabilities, and the correlation between them. We found that some of the components do tend to lead the total GDP, or the whole economy when entering a recession.

CHAPTER 1

INTRODUCTION

This dissertation is a collection of work in the field of regime switching effects on the US equity market and US economy. We used regime switching models on market index and factors to capture financial market condition changes. We then tested the relationship changes between different assets, and constructed a portfolio rebalance method that uses the market regime information. Also, we applied regime switching models to US GDP and GDP components, to find ways to interpret and predict US business cycle, especially the recessions.

In Chapter 1, I provide an introduction to each part of the work with summary of my contribution, and forms of models and algorithms being used. More detailed backgrounds, literature views, procedures and results will be presented in the later chapters.

Chapter 2 is focused on regime switching models in the financial markets. In the first part we apply the regime switching model to the Capital Asset Pricing Model (CAPM). The regimes are defined by applying the Hidden Markov Model (HMM) (Hamilton, 1988) on the market index. Each of the regimes represents the market condition with high volatility and low volatility separately. With the regimes defined, we analyzed the relationship changes between a group of stocks on the return, volatility and the return and volatility after taking out the market effects, and observed the relationship changes with the market regime changes.

In the second part of Chapter 2, we applied HMM on Factor Models, where regimes are defined by the four factor model, which is Fama French three factors (Fama and French, 1992) and momentum factor. The estimated parameters under each regime also represent to high and low volatility periods. We tested the correlation changes between these factors, and found there are periods these factor are more correlated than usual, which can be of consideration when applying these factors to do prediction.

In the third part of Chapter 2, with the facts that the HMM regimes on market index and factors can represent different market conditions, and the relationship between assets change with such regimes, we constructed a simple method to see if such regime changes can be used on portfolio rebalance decision. We find that, with the regime into consideration, we can improve the Markowitz portfolio optimization results in a portfolio rebalance.

Chapter 3 is focused on regime switching models on US economy. First, we applied HMM on US GDP growth rate. We match the regime probability and see the low growth rate regimes matches with NBER announced recession periods very well. Secondly, we use the same HMM technique on the US GDP components. With the regime probability series estimated by HMM, we compared the components regimes with the whole GDP regimes. The purpose is to find the possible leading components that can be used to predict the whole US GDP, or specifically the recession. We calculated various conditional probabilities between the GDP regime probabilities and lagged GDP components probabilities, and the correlation between them. We found that some of the components do tend to lead the total GDP, or the whole economy when entering a recession.

1.1 Hidden Markov Model in Economic and Financial Time Series Analysis

The Hidden Markov Model is a parametric stochastic probability model that is used on time series data to segment the series into different states (regimes). It was widely used in signal processing in electronic devices, speech recognition, pattern recognition, etc. The Hidden Markov Model was first introduced to the econometric and financial world by (Hamilton, 1989a), and it became more and more popular in financial time series analysis. The Hidden Markov Model were extended in different ways, and many methodologies have been applied to provide decision making on various financial and economic problems. For example, (Turner et al., 1989), (Bhar and Hamori., 2004), (Nielsen and Olesen, 2001), (Chang, 2009), (Chang and Feigenbaum, 2008) employed Markov-switching models for the modeling of stock returns, and the relationship between stock, interest rate and bond. (Chen, 2001), (Erlandsson, 2002), (Gruss and Mertens, 2009), (Susmel and Kalimipalli, 2001), (Henry, 2009) applied the HMM on interest rate modeling.

(Dueker, 1997), (Pagan and Schwert., 1990), (Susmel, 2000), (Brunetti et al., 2003), (Chkili and Nguyen, 2011),(Shyu and Hsia, 2008), (Qiao et al., 2008), (Marcucci, 2005), (Hobbes et al., 2007), (Aray, 2008) applied regime switching models on stock volatility models, including regime switching ARCH/GARCH type models.

(Zhang et al., 2011) showed how the regime switching model can be applied to model stock market volatility changes responding to policy changes in certain countries when the financial crisis in 2007 occurred. (Engel and Hamilton., 1990), (Engel, 1994), (Bollen et al., 2000), (Klaassen, 2002), and (Beine et al., 2003) investigated regime switching in foreign exchange rates.

Our work focuses on regime changes of the market index and Fama French factors, how the correlation of stock returns or certain portfolio returns change under and how they can affect the decision of portfolio rebalancing.

1.2 Hidden Markov Model Specification

In a time-series state model, the underlying random vector is

$$\mathbf{X}_t = \begin{pmatrix} \mathbf{Y}_t \\ \mathbf{Z}_t \end{pmatrix}, t = 1, 2, \ldots, n.$$

The random variable $\mathbf{Y}_t$ is observable; it is in general a vector, though for much of the exposition below it will be a scalar. The variable Y is the object of our analysis. It may be a transform of another variable. For example, in Chapter 3 we have $Y_t = \ln P_t - \ln P_{t-1}$, where P_t is the GDP in quarter t.

1.2.1 The State Variable

We assume that the time series moves across K states. The vector $\mathbf{Z}_t$ consists of unobservable state indicators. The K states are indexed by $k = 1, 2, \ldots, K$. The element $Z_{kt} = 1$ if and only if the state is k at time t; the other Z_{jt} are then 0.

The number of states K may be determined depend on the problem the model is applied to. However, often it can be estimated by model-selection criterion such as BIC. BIC is a leading term in the expansion of $\ln \P(H_k|D)$, where H_k is the hypothesis that model k is true, and D represents all the data available. See also (Kashyap, 1982) for more detailed explanation.

If the values of the Z_{kt} were known, we could estimate the mean μ_k for the k-th state by

$$\hat{\mu}_k = \Sigma_{t=1}^{n} Z_{kt} Y_t / \Sigma_{t=1}^{n} Z_{kt}.$$

The denominator is the frequency of state k among the n time points, and the numerator is the sum of the observations in state k.

Equivalent to $\mathbf{Z}_t$, we define a scalar random variable Q_t which is equal to k if the state at t is k . $Q_t = k$ if and only if $Z_{kt} = 1$. The process of determine $\{Q_t\}$, is called the label process. And Q_t is a homogeneous process, a first-order Markov chain with stationary transition probabilities

$$\tau_{jk} = \P\{Q_t = k \,|\, Q_{t-1} = j\}.$$

The τ_{jk} are non-negative and sum to 1 for each j.

The conditional probability density function (p.d.f.) of Y_t given that $Q_{t-1} = j$ is the mixture p.d.f.

$$f(y_t|q_{t-1} = j) = \tau_{j1} f_1(y_t) + \tau_{j2} f_2(y_t) + \ldots + \tau_{jK} f_K(y_t),$$

where there are K component p.d.f.s $f_k(\cdot), k = 1, 2, \ldots, K$. In the parametric case, the component p.d.f.s $f_k(y) = g(y|\mathbf{\Theta}_k)$, where $\mathbf{\Theta}_k$ denotes the parameters of the k-th distribution. For example, $g(\cdot)$ might be the multivariate Normal p.d.f. Then the distributional parameters are $\mathbf{\Theta}_k = (\mu_k, \mathbf{\Sigma}_k)$, where μ_k and $\mathbf{\Sigma}_k$ are the mean vector and covariance matrix. The conditional probability that y_t is from the k-th component, given that $Q_{t-1} = j$, is

$$p(k|y_t, q_{t-1} = j) = \tau_{jk} f_k(y_t)/f(y_t|q_{t-1} = j),$$

where $f(y_t|q_{t-1} = j) = \Sigma_{k=1}^{K} \tau_{jk} f_k(y_t)$.

The total number of free parameters for K m-dimensional Normal components is Km for means, $Km(m+1)/2$ for covariance matrices, and $K(K-1)$ for transition probabilities, since each row of the transition probability matrix must add to 1. The total number of parameters is $K(K-1) + Km + Km(m+1)/2$. In the univariate case, $m = 1$ and the number of parameters is $K(K-1) + K + K = K(K-1) + 2K$.

1.2.2 The Likelihood

Given the states, the conditional joint p.d.f. is in the form

$$f(\mathbf{y}|\mathbf{q}) \; = \; \pi_{q_1} f_{q_1}(y_1) \tau_{q_1 q_2} f_{q_2}(y_2) \times \ldots \times \tau_{q_{n-1} q_n} f_{q_n}(y_n).$$

With $\mathbf{y}$ fixed, it can be considered as a function of the parameters. The unconditional likelihood corresponds to the p.d.f., which is the probability weighted sum of $f(\mathbf{y}|\mathbf{q})$ over all possible sets of labels, is in form of

$$f(\mathbf{y}) = \Sigma_{\mathbf{q}} f(\mathbf{y}|\mathbf{q})p(\mathbf{q}).$$

In dealing with historical data, if the main goal is to label the states at each time period, it is appropriate to deal with the conditional likelihood.

1.2.3 Model Estimation for Hidden Markov Model

The EM algorithm (Dempster et al., 1977) consists of an E (Expectation) step, estimating the states, and an M (Maximization) step which here means estimating the distributional parameters and transition probabilities.

1.2.3.1 Greedy Algorithm

The conditional probability $p(k|y_t)$ plays an important role in algorithms for estimating Hidden Markov Model. To maximize the conditional likelihood, one can first consider a greedy, one-step look-ahead algorithm. The greedy algorithm is not optimal, but it has a straightforward and fast estimation procedure. In the greedy algorithm, at any stage, it loops through the data Y_t from $t = 1$ to n. At each point t, having estimated the label q_{t-1} to be j, it will estimate the label q_t to be

$$\begin{aligned}\hat{q}_t &= \text{argmax}_{k=1,2,\ldots,K}\hat{p}(k|y_t, q_{t-1} = j) \\ &= \text{argmax}_{k=1,2,\ldots,K}\hat{\tau}_{jk}\hat{f}_k(y_t)/\hat{f}(y_t|q_{t-1} = j)\end{aligned}$$

Note that the denominator $\hat{f}(y_t|q_{t-1} = j)$ does not depend on k, so

$$\hat{q}_t = \text{argmax}_{k=1,2,\ldots,K}\hat{\tau}_{jk}\hat{f}_k(y_t).$$

In the EM algorithm the estimation of the k-th mean at any given iteration is

$$\hat{\mu}_k = \Sigma_{t=1}^n \hat{p}(k|y_t)y_t/\Sigma_{t=1}^n \hat{p}(k|y_t),$$

where the estimates $\hat{p}(k|y_t)$ are computed from the current values of parameters estimated.

We will use such estimation of $\hat{\mu}_k$ in the next section where the Baum-Welch algorithm is used to obtain the conditional probabilities. First note that one could use an algorithm with estimation of the form.

$$\hat{\mu}_k = \Sigma_{t=1}^n \hat{z}_{kt}y_t/\Sigma_{t=1}^n \hat{z}_{kt},$$

where $\hat{z}_{kt}$ are different from $\hat{p}(k|y_t)$ and are the "hard" estimates 0 or 1 according to the current label of y_t. The variances are estimated analogously. The estimates of the transition probabilities are just the observed relative frequencies of the transitions. The procedure begins with initial estimates of the state probabilities, the transition probabilities, the means, and the variances. Computing programs using the greedy algorithm are included in (Sclove, 1992).

1.2.3.2 Baum-Welch Algorithm

The Baum-Welch Algorithm, developed by Lloyd R. Welch in collaboration with Leonard E. Baum, is a powerful tool for examining and analyzing the results of continuing processes that proceed regularly in stepwise fashion – Markov processes. It has become an important tool in many fields, particularly in speech recognition.

The Baum-Welch algorithm estimates the probability distribution over the states of the hidden Markov chain at each time point.

In (Baum et al., 1970), there is a reference to a paper by Baum and Welch submitted to the Proceedings of the National Academic of Sciences (PNAS), but the archives of PNAS reveal no publication by Baum and Welch. Yet, everyone knows the algorithm as "The Baum-Welch algorithm".

The Baum-Welch algorithm is more complicated than the greedy algorithm or Viterbi algorithm (which will be mentioned later), but gives better estimation of the parameters Θ. The algorithm adjusts the estimates of Θ through iterations of a Forward-Backward procedure and finds the one that maximizes $P(Y|\Theta)$ which is the HMM likelihood, a function of the parameters Θ for the fixed data vector $Y = (y_1, y_2, \ldots, y_n)$.

Forward Procedure

First we define

$$\alpha_t(j) = P(y_1, y_2, y_3, ...y_t, q_t = j|\Theta).$$

So $\alpha_t(j)$ is the probability of the observation sequence up to time t and being in the state j at time t, given the current parameters of the model Θ. The $\alpha_t(j)$ can be computed inductively:

1. Initialization:

$$\alpha_1(j) = \pi_j f_j(y1), 1 \le j \le N,$$

2. for $t = 1, 2, \ldots, T-1, 1 \le j \le N$

$$\alpha_{t+1}(k) = \left[\sum_{j=1}^{N} \alpha_t(j)\tau_{j,k}\right] f_k(y_{t+1}),$$

3. So we have

$$P(Y|\Theta) = \sum_{k=1}^{N} \alpha_T(k),$$

Backward Procedure

We define

$$\beta_t(j) = P(y_{t+1}, y_{t+2}, y_{t+3}, ...y_T | q_t = j, \Theta).$$

So $\beta_t(j)$ is the probability of the observation sequence from t+1 to T, given the state j at time t and the parameters Θ of the model. (Notice the condition is different from the Forward Procedure). $\beta_t(j)$ also can be computed inductively:

Computation of $\beta_t(j)$:

1. Initialization:

$$\beta_T(j) = 1, 1 \leq j \leq N,$$

2. for $t = 1, 2, \ldots, T-1, 1 \leq j \leq N$,

$$\beta_t(j) = \sum_{k=1}^{N} \tau_{j,k} f_k(y_{t+1}) \beta_{t+1}(k),$$

3. and we have

$$P(Y|\Theta) = \sum_{j=1}^{N} \pi_j f_j(y_1) \beta_1(j),$$

Then we define $\gamma_t(j)$,

$$\gamma_t(j) = P(q_t = j | y, \Theta).$$

The quantity $\gamma_t(j)$ is the probability that state is j at time t given all the observations y and the current parameter θ of the model.

So, in the **E-Step**, we have $\gamma_t(j)$ by Bayes law:

$$\begin{aligned}\gamma_t(j) &= \frac{P(q_t = j, y|\Theta)}{P(y|\Theta)} \\ &= \frac{\alpha_t(j)\beta_t(j)}{P(y|\Theta)}.\end{aligned}$$

Also we define $\xi_t(j,k)$ as the probability of being in state j at time t and being in state k at time $t+1$, given the observation vector Y and the parameter Θ. So,

$$\begin{aligned}\xi_t(j,k) &= P(q_t = j, q_{t+1} = k)|Y, \Theta) \\ &= \frac{P(q_t = j, q_{t+1} = k, Y|\Theta)}{P(Y|\Theta)} \\ &= \frac{P(q_t = j, q_{t+1} = k, Y|\Theta)}{P(Y|\Theta)} \\ &= \frac{P(q_t = j, y_1, y_2, \ldots y_t|\Theta)P(q_{t+1} = k, y_{t+1}, y_{t+2}, \ldots y_T|\Theta)}{P(Y|\Theta)} \\ &= \frac{\alpha_t(j)P(q_{t+1} = k, y_{t+1}, y_{t+2}, \ldots y_T|\Theta)}{P(Y|\Theta)} \\ &= \frac{\alpha_t(j)P(q_{t+1} = k|q_t = j)P(y_{t+1}|q_{t+1} = k)P(y_{t+1}, y_{t+2}, \ldots y_T|q_{t+1} = k, \Theta)}{P(Y|\Theta)} \\ &= \frac{\alpha_t(j)\tau_{j,k}f_k(y_{t+1})\beta_{t+1}(k)}{P(Y|\Theta)}.\end{aligned}$$

And in the **M-Step**, the re-estimation formulas are:

$$\hat{\pi}_j = \gamma_1(j)$$
$$\hat{\tau}_{j,k} = \sum_{t=1}^{T-1} \xi_t(j,k) \Bigg/ \sum_{t=1}^{T-1} \gamma_t(j)$$
$$\hat{f}_j(w) = \sum_{\substack{t=1 \\ y_t = w}}^{T} \gamma_t(j) \Bigg/ \sum_{t=1}^{T} \gamma_t(j).$$

The parameters of the probability density function under each state $f_j(), j = 1, 2, \ldots N$ are re-estimated by weighted estimation.

Each E step and M step complete one iteration, and when the estimates of all the parameters Θ have converged, the iterations stop. And $\gamma_t(j)$ gives us the probability being in state j at time t.

1.2.3.3 Model Estimation with Other Methods

One of the very important algorithms for Hidden Markov Model is the Viterbi algorithm (Viterbi, 1967),(Forney, 1973). It is also a dynamic programming applied to the problem of finding the most likely state sequence in an HMM. The paper by Forney has a particularly good explanation. It is worth noting that in 1985, Dr. Andrew J. Viterbi co-founded QUALCOMM, Inc., and that the engineering school at the University of Southern California is the USC Viterbi School of Engineering. The Viterbi algorithm could be used for the maximization step of the EM algorithm, while the Baum-Welch algorithm gives maximum likelihood estimates for the mixture likelihood.

Another approach is provided by Hamilton (Hamilton, 1989a), which is an algorithm less complicated than Baum-Welch algorithm. The Viterbi algorithm and Baum-Welch algorithm are maximum likelihood approaches. There are other ways to estimate the HMM. We can use Markov Chain Monte Carlo (MCMC)(Gamerman, 1997), which is a Bayesian approach. Or we can use a Newton type procedure to find the parameters that maximize the likelihood function, which can be available in softwares like R, Matlab and the B34S system. The advantage of the latter two methods is they may handle more complicated models with more parameters involved, and one don't need to worry about the detailed procedures.

CHAPTER 2

REGIME SWITCHING UNDER CAPM AND FAMA FRENCH AND MOMENTUM FACTOR MODEL

It is commonly known that the financial market goes through different conditions, for example Bull and Bear markets, highly volatile periods and non-volatile periods. The prediction of market condition change, especially volatile market movement is very important to decision makings on asset allocation, risk control of trading and other investment behaviors. In our research, the market condition is modeled as conditional distribution of certain market indices and factors, or the relationship of such factors with assets. We use the Hidden Markov Model (HMM) to analyze market index and factors. Relationship changes of groups of stocks and portfolios are tested. Finally a portfolio rebalance strategy that uses the HMM market regime switching is compared with a strategy without considering regime change of the market. The results show that, when both strategy use the Markowitz portfolio optimization, the portfolio rebalance considering the HMM regimes is better than the one without considering the regimes.

2.1 Introduction

In our research, we use Hidden Markov Model to discover the regime changes on the time series of the Market Index, Fama French factors and momentum factor. We are interested in the relationship changes, variance-covariance changes between group of stocks or group of portfolios. It is natural to assume variance-covariance matrix will be different under different market conditions. We want to know if such changes happen with the HMM regime changes of the index and factors. The value of our research is, if there is evidence that the relationship of a group of asset indeed change with the HMM regime, it will be possible to consider the HMM regime when doing the portfolio rebalancing on the timing, and on the weight calculation. We apply the HMM first on the major financial market index, the S&P500 index, then on the Fama French factors and momentum factor. There are other indicators can be used to describe different financial behaviors. For example, the VIX for stock market volatility, certain industry index like Amex oil index for specific group of stocks. The regimes changes in these indices can be used as factors to describe specific market condition changes. In this thesis, we focus on the major indicators, just to show that it is possible to use the Hidden Markov Model results to guide portfolio rebalance. The results of Hidden Markov Model shows the indicators can go through 2 different states that behave differently mainly on the volatility part. This is in line with the common knowledge of the financial market. And the fact that the Fama French factors also go through different regimes with high and low volatility leads to some other interpretation as described later in the sections. In the empirical study, we use 16 oil industry stocks daily data and calculated the correlations matrix of several measurements, include return, residuals after regressing the returns on the market index, squared stock returns, squared residuals and so on. The reason we choose several measures is each of them shows a different relationship between stocks. The correlation of returns shows the relationship of the stock movement changes with the HMM regime, and the squared returns shows the relationship of the stock volatility changes with the regimes, residual after the regression shows the relationship outside the market effects. We observed the correlation difference between different S&P500 Index regimes. Other than market index like S&P500, the Fama French factors are also well known for asset pricing and portfolio management. Because such factor model uses multi factors for prediction, we are interested in the relationship between those factors. If one wants to use the 3 factors in a regression, it's important to know if and under what condition, the factors have a high correlation or low correlation.

Our results show clear correlation changes between these factors and 2 states regime changes are suggested by HMM for each factor.

With the regimes defined on market index and Fama French factors, and test shows correlation changes between stocks, we constructed a way of using Hidden Markov regime switching result to guide the portfolio rebalance. The purpose however is not to find a good rebalancing strategy that can beat other strategies immediately, but to show that the regime switching can have an effect on the portfolio optimization.

We used 49 Industry Portfolios and applied two ways of rebalancing methods, one uses past N days of daily return to perform the Markowitz portfolio rebalance, the other uses the data from the periods in the same regimes as current time to calculate the risk and expect return and do the portfolio optimization. We compare the weights and expected risk/return with the best optimized weights fitted with next M days data, which we considered the "true best optimized weights and performance". We can see better performance when using the Hidden Markov regime as guidance.

Capital Asset Pricing Model

The Market Model is a commonly used model that relates a stock's returns R_t to the market return M_t.

$$R_t = \alpha + \beta M_t + \varepsilon_t,$$

where

$$\varepsilon_t \sim N(0, \sigma^2).$$

In the Capital Asset Pricing Model (CAPM), we have

$$R_t - r_f = \beta(M_t - r_f) + \varepsilon_t,$$

where r_f is the risk free return rate (interest rate). CAPM indicates that the individual risk premium $R_t - r_f$ equals the market premium $(M_t - r_f)$ times β the sensitivity of the expected excess asset returns to the expected excess market returns.

The expected market rate of return is usually estimated by measuring the Geometric Average of the historical returns on a market portfolio (e.g. S&P 500). And the risk free rate of return used for determining the risk premium is usually the arithmetic average of historical risk free rates of return and not the current risk free rate of return.

Investors are interested in knowing the relationship between the behavior of an asset (stock, portfolio, or mutual fund) and the behavior of the stock market as a whole.

The value of the slope coefficient β in the linear regression is referred to as the mutual funds' "beta coefficient". Assuming the preceding model is true, investors can predict how the rates of return of an individual asset by

$$R_t = r_f + \beta(M_t - r_f) + \varepsilon_t,$$

So the CAPM can be regarded as representing a single-factor model of the asset price, where β is exposure to changes in value of the Market. It is a useful tool in determining if an asset being considered for a portfolio offers a reasonable expected return for risk. For example, if beta is greater than 1, the implication is that the return to the mutual fund will be greatly influenced by the behavior of the market and will move in the same direction as the change in the market return. If beta is between 0 and 1, the rates of return of the asset will be less sensitive to changes in market behavior but will also move in the same direction as the change in the market return.

A very commonly known behavior of the market that related to regime changes are Bull and Bear markets. By a common definition, a bear market is marked by a price decline of 20% or more in a key stock market index from a recent peak over a 12-month period.

On average, the stock market suffers a bear market every four or five years, defined as a drop of 20% in major indexes, such as the S&P 500. During an average bear market, the

S&P loses about 25% of its value over time. It usually takes 11 to 18 months for the market to hit bottom, which is about 2 to 5 quarters. (Lunde and Timmermann, 2004) studied time series dependence in the direction of stock prices by modeling the (instantaneous) probability that a bull or bear market terminates as a function of its age and a set of underlying state variables, such as interest rates. A random walk model is rejected both for bull and bear markets. Although it fits the data better, a generalized autoregressive conditional heteroscedasticity model is also found to be inconsistent with the very long bull markets observed in the data. The strongest effect of increasing interest rates is found to be a lower bear market hazard rate and hence a higher likelihood of continued declines in stock prices. (Ang and Timmermann, 2011) shows regime switching models can capture the stylized behavior of many financial series including fat tails, skewness, etc. (Hamilton, 1989a) suggested the periodic shift from a positive growth rate to a negative growth rate is part of the U.S. business cycle, and they can be used as criteria for defining and measuring economic recessions. (Maheu and McCurdy, 2000) use a Markov-switching model that incorporates duration dependence to capture nonlinear structure of the stock returns in both the conditional mean and the conditional variance.

Fama French Model

The Fama French three factor model is a model designed by Eugene Fama and Kenneth French to explain stock returns.

The traditional asset pricing model CAPM as mentioned above, uses the returns of the market to explain the returns of a portfolio or stock with parameter β. In the Fama

French model, three variables are used. Fama and French started with the observation that two classes of stocks have tended to do better than the market as a whole: small caps stocks and stocks with a high book-to-market ratio. So they added two factors to CAPM to reflect a stock or portfolio's exposure to these two factors.

$$R_t - r_f = \alpha + \beta_1(M_t - r_f) + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_t,$$

where R_t is the portfolio's return, r_f is the risk-free return, and M_t is the return of the whole stock market. SMB stands for "small (market capitalization) minus big", and HML stands for "high (book-to-market ratio) minus low". They represent the historic excess returns of small caps over big caps and of value stocks over growth stocks. With the SMB and HML calculated, the corresponding coefficients β_2 and β_3 are determined by linear regressions.

Portfolio Rebalancing

Portfolio rebalancing means adjust the weights of the securities in a portfolio so to make the expected risk and return inline with the original target. The purpose is usually to control risk. A portfolio's asset allocation may drift away from the original target over time, the portfolio can become too risky, or too conservative. And buying or selling securities in the portfolio is needed to bring the weights back to optimal level.

2.2 Models and Statistical Analysis

2.2.1 HMM for CAPM

In this section, we use the Markov Switching Model to segment S&P500 index to two regimes. We then tested the correlation change of 16 stocks from the oil industry under different regimes of the index, and show the correlation changes with the market regime. We also used DJIA index to do the segmentation. The market return series of S&P500 and DJIA looks similar and the regime periods tend to occur in the similar time periods, so we only give the parameter estimated and plot of the S&P500 index.

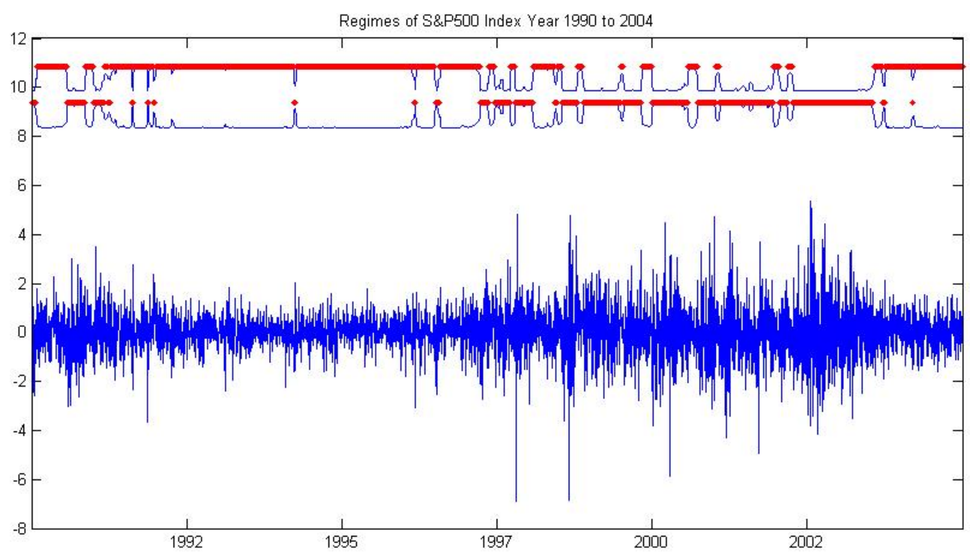


Figure 1. Regime changes and probabilities in S&P500 Index

Normal Mix Model			
Regime	Mean	Std.	
1	-0.0248	1.4314	
2	0.0691	0.6509	
	Trans. Prob.		
	0.9816	0.0184	
	0.0116	0.9880	
AR1 Model			
	α	β	Std.
1	-0.0194	-0.0120	1.4307
2	0.0653	0.0259	0.6505
	Trans. Prob.		
	0.9818	0.0182	
	0.0115	0.9881	

TABLE I

Estimated parameters for S&P500 index.

2.2.2 HMM For Fama French Factors

In this section we use the Hidden Markov Model to segment the Fama French three factors. We are interested in the regime switching of the factors, the parameters of the model under each regime and the change of the interaction between the factors. By studying the regime switching of these factors we hope to explore some investing strategy related to regime changes.

We fit the following mixed Normal HMM model to each of the three Fama French factors:

$$Y_t = \mu_{q_t} + \varepsilon_{q_t}, \qquad \varepsilon_{q_t} \sim N(0, \sigma^2_{q_t}), \qquad q_t = 1, 2.$$

Table II gives the estimated parameters. We show the time series plot together with the regime probabilities of each of the Fama French factors in Figure 2, Figure 3 and Figure 4.

TABLE II

Estimated parameters for regimes of Fama French 3 factors.

	Mkt-RF		SMB		HML	
Regime	μ	σ	μ	σ	μ	σ
1	-0.0387	1.4051	-0.0631	1.0078	0.0341	0.9686
2	0.0741	0.5928	0.0151	0.4649	0.0097	0.3542
	Trans.Prob.		Trans.Prob.		Trans.Prob.	
	0.9786	0.0214	0.9531	0.0469	0.9803	0.0197
	0.0134	0.9861	0.007	0.9927	0.0071	0.9925

2.3 Empirical Study

In modern portfolio theory, correlation has a very important role in reducing portfolio risk. During turbulent times the correlation tend to change and usually become higher. It will be highly desirable to rebalance the portfolio when a market change is noticed or predicted.

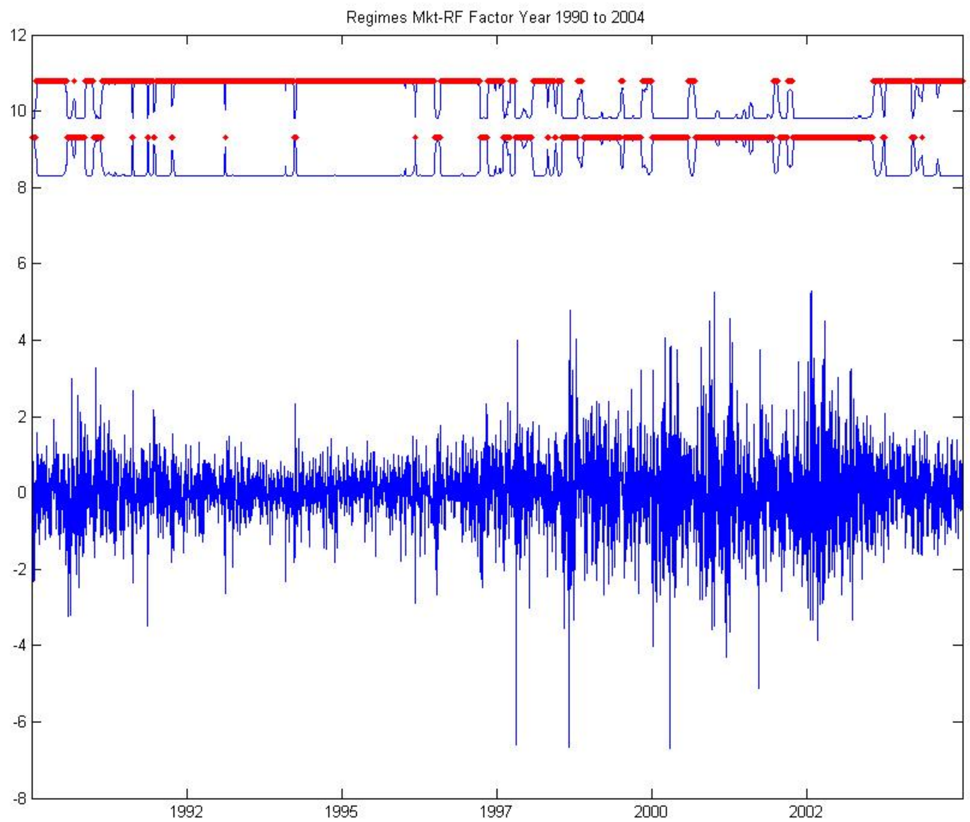


Figure 2. Regime changes of Market-RF factor

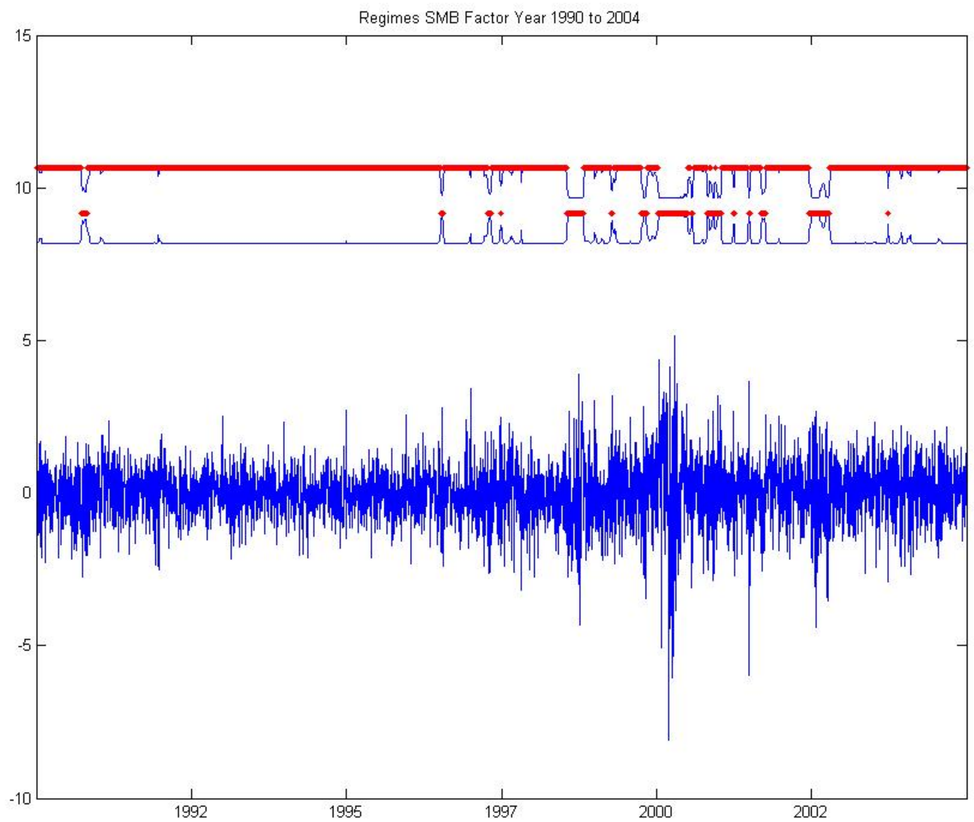


Figure 3. Regime changes of SMB factor

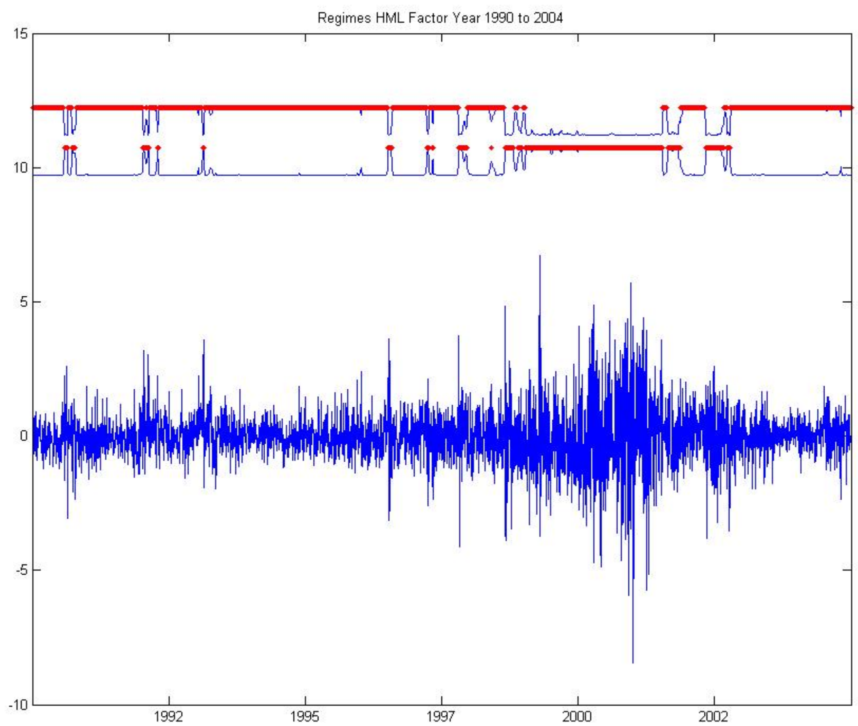


Figure 4. Regime changes of HML factor

It is interesting to see how the regime switching of the market return relates to the relationship changes between stocks such as correlation, and relationship between stocks and the market or factors.

Under the framework of the Hidden Markov Switching Model, we tested whether stock correlations are stable through different market index regimes.

We looked at how the correlations change over regimes. Do high correlations between stocks indeed go with highly volatile market regimes?

2.3.1 Stock Relationship Change under Regimes of a Market Index

The market index: We use the S&P500 daily returns in the period of January 1990 to December 2004 as market index.

Forms of models within each index regime: Regime change can have many different aspects including difference on volatility, different stock betas on the market, different return distributions, etc. Which aspects the regime results will reveal depend on the model we choose for the each of the regimes. For correlation testing, because the stock correlation tend to be higher on more volatile market, it is reasonable to choose models that can show the difference on market volatility. And to make the correlation comparison more clear, it is better to choose more states than two, and to see how the correlation differ under each level of market volatilities. So for the Hidden Markov Model on S&P500 Index, we defined four states instead of two. If the number of observations in one of the regime became very small and below a threshold, that regime will be dropped out in the process, and the number of states will be reduced. Also, we examine the number of data in each regime after the estimation and will discover if the data in any of the regime is too low and need to be dropped.

The S&P500 index return series Y was segmented to regimes belonging to one of the states, according to model

$$Y_t = \alpha_{q_t} + \beta_{q_t} Y_{t-1} + \varepsilon_{q_t}, \qquad \varepsilon_{q_i} \sim N(0, \sigma_{q_t}^2), \qquad q_t = 1, 2, 3, 4.$$

In each state, the market return is considered an AR(1) process with the residual following a Normal distribution with different standard deviations ε_{q_i}, where q_i being the state at time t.

Table III shows the parameters estimated for S&P500 index. We can see from state 1 to state 4, the standard deviation σ is decreasing. State 1 has the highest $\sigma = 2.18$, and state 4 has the lowest 0.45. The 4 regimes associate with market conditions with volatilities from high to low. Figure 5 shows the regime probabilities on the time series graph, comparing with the return series graph, it can be seen the probability of regime 1 tend to match with the periods the return series are more volatile.

The groups of stocks we tested include 16 oil stocks. Let $\rho_{ij|s}$ represent the correlation between stock i and j under regime s, the group of stocks within each regime were compared with $\rho_{ij|0}$, the correlation computed with the whole series without considering regime. For the correlation between each pair of stocks i, j, we tested if the correlation $\rho_{ij|s}$ is significantly different from $\rho_{ij|0}$.

Test of the correlations difference: For the group of stocks, we tested correlations of the following series:

1. The stock returns.
2. The residuals after regressing the stock return on the market index return.
3. The residuals after regressing the stock return on the market index return within each regime.
4. Squared stock returns.
5. Squared residual after regressing the stock return on the market index return.
6. Squared residuals after regressing the stock return on the market index return within each regime.

The reason that tests of 2 and 3 could be interesting is that they can help us answer the question:

- Are changes in correlations between stock returns related to the whole market return?
- Are the correlation changes caused by the difference of the beta within each regime and the beta for the whole period?

And in tests 4 to 6, to check the correlation of the volatility of the stocks, we use the squared return to approximate the daily volatility, and used squared residuals to approximate the volatility that are not explained by the market movements.

2.3.1 shows the results for the correlation tests for the 16 oil stocks under regimes of S&P500. The numbers inside the table show the number of tests that are significantly larger (or smaller) The total number of correlations calculated is $m(m-1)/2$ where m is the number of stocks. For m=16, the total number of tests is 120. The combination of 16 oil stocks and S&P500 index regime is a good illustration because.

- Some level of correlation exists between the oil industry stocks, so it is meaningful to the compare the correlation with the whole time period.
- The direct relationship comeing from the composite of index will be less likely to occur, such relationship could exist if we choose DJIA(Dow30) index, and use the stocks in the Dow30.

So this illustration can represent a group of stocks that are correlated naturally and the response to the general market condition changes.

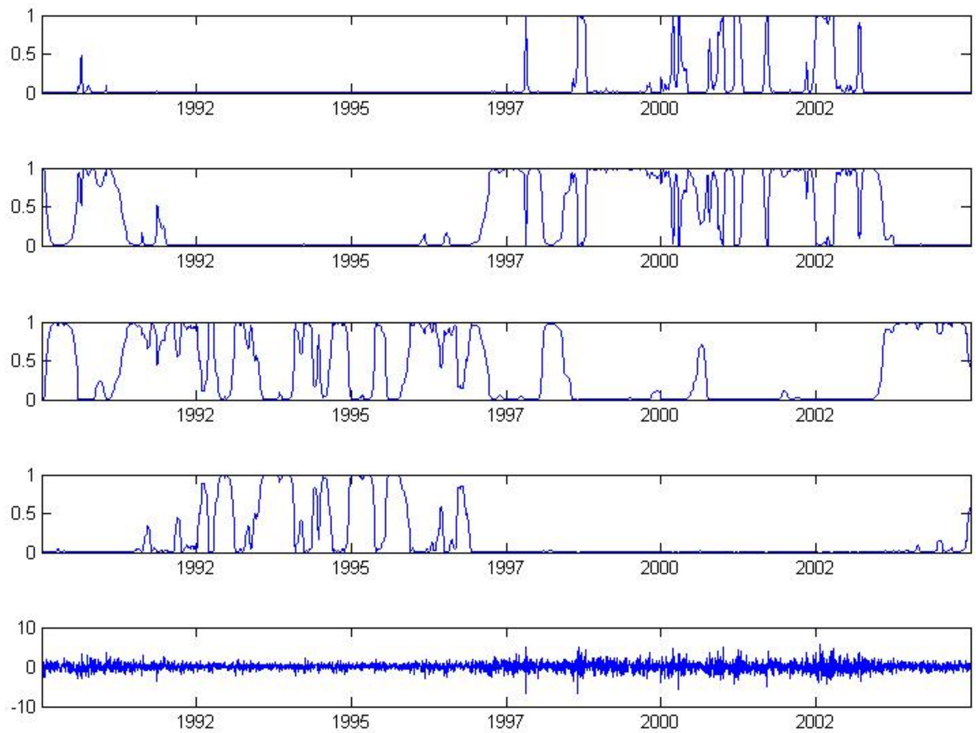


Figure 5. 4 states regime probabilities and return series of SP500 Index (1990-2004)

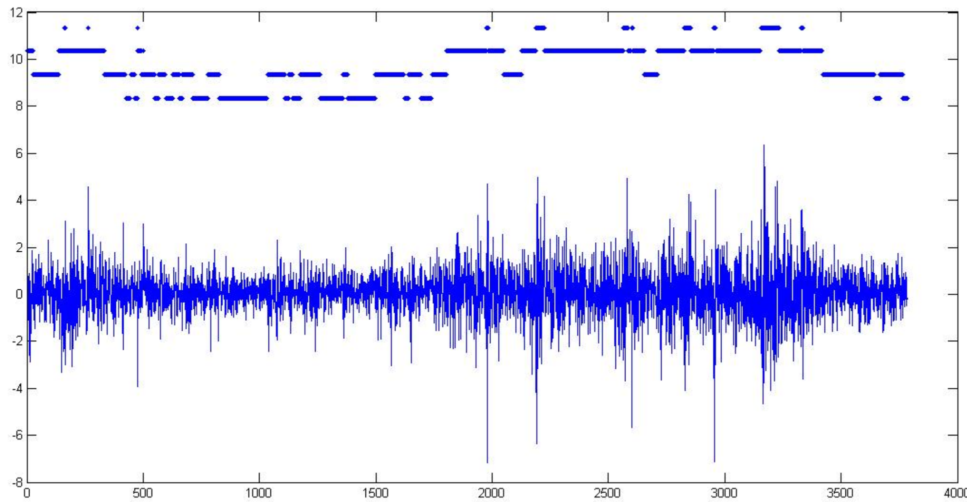


Figure 6. 4 states regime labeled and return series of SP500 Index

TABLE III

Parameters Estimated for S&P500 index with 4 states

4 states Hidden Markov AR(1) Model					
	State	1	2	3	4
	ϕ	-0.02751	0.001169	0.024002	0.095009
	σ	2.181334	1.126607	0.738927	0.450461
		1	2	3	4
Trans. Prob.	1	0.945525	0.054475	2.33E-23	1.70E-51
	2	0.010359	0.984626	0.005013	0.000000
	3	0.000000	0.004204	0.985563	0.009944
	4	0.000000	0.000000	0.023039	0.976023

Compare with correlation calculated with whole time period								
Regime	1		2		3		4	
Total Number 120	larger	smaller	larger	smaller	larger	smaller	larger	smaller
Return correlation	114	0	46	5	2	47	0	91
Flat beta residual	32	5	17	7	3	37	5	22
State different beta residual	32	4	14	7	3	36	5	22
Squared Return correlation	75	0	41	15	43	13	2	23
Squared State different beta residual	29	1	18	13	7	20	1	8
Squared State different beta residual	30	1	17	12	7	21	1	9

TABLE IV

Compare Correlation of 16 oil industry stocks under different S&P500 regimes

We can see that the correlation of the 16 stocks is higher in the high volatility regime (state 1) of the S&P500 index, 114 of the 120 correlations are significantly higher than correlation calculated without considering regimes, none became significantly lower. And in the low volatility regime (state 4), 91 of the 120 correlations are lower than non-regime correlation, none are significantly higher. States 2 and 3 showed similar patterns. These results suggested the correlation between certain stocks does change with the Hidden Markov Model estimated regimes.

We can also see the changes on the correlation of the residuals have a similar pattern, although not as strong as the pattern on the return, it suggested there are correlation changes on the returns that can't be explained by the stocks responding to the whole market movement.

The correlation of squared return and the squared residual also showed differences between different regimes. It suggested the correlations of the stock volatility are also different for each regimes, and there are volatility correlation that can't be explained by moving with the whole market.

Because the number of data selected from each regimes will be smaller than the whole period data. To test if the significance on correlation changes is due to the number of data we used to do calculation, we draw the same number of data in each regime but randomly from the whole time period and calculate the correlation, and test the difference. We did not see any of the same patterns and the numbers of significant differences are almost all close to zero.

The above tests suggested the Hidden Markov Regime Switching are able to reveal the market condition, and the stock relationship changes on return or volatility does exist between HMM regimes. And it is reasonable to apply HMM on portfolio rebalance and optimization strategies.

2.3.2 Correlation Changes among the Fama French Factors

Since Fama French Factors are also commonly used as stock return predictor, and they represent the market conditions, we are interested in testing the relationship changes between these factors themselves. Because these factors are usually applied in a multivariate regression model as explanatory variables on the stock return. The correlation between these factors are important on the predicting power of the Fama French three factor model.

We used the Fama French factors from year 1990 to 2004 and calculated the correlation of 60 days rolling window. Figure 7 shows the correlation plot of the Fama French three factors. We can see the correlation between these factors switch from time to time. And the direction of the changes are opposite between the 1st and 3rd plot before year 2003.

2.4 Portfolio Optimization with Index and Factor Regime Switching

In the earlier sections, we have shown that the HMM regimes do reveal different market conditions. And there are evidences showing stock correlations change with the market index regimes. There could be many different ways to consider regime switching in investing. In this section, we want to test if the HMM regime changes on the market index has effect on portfolio rebalancing.

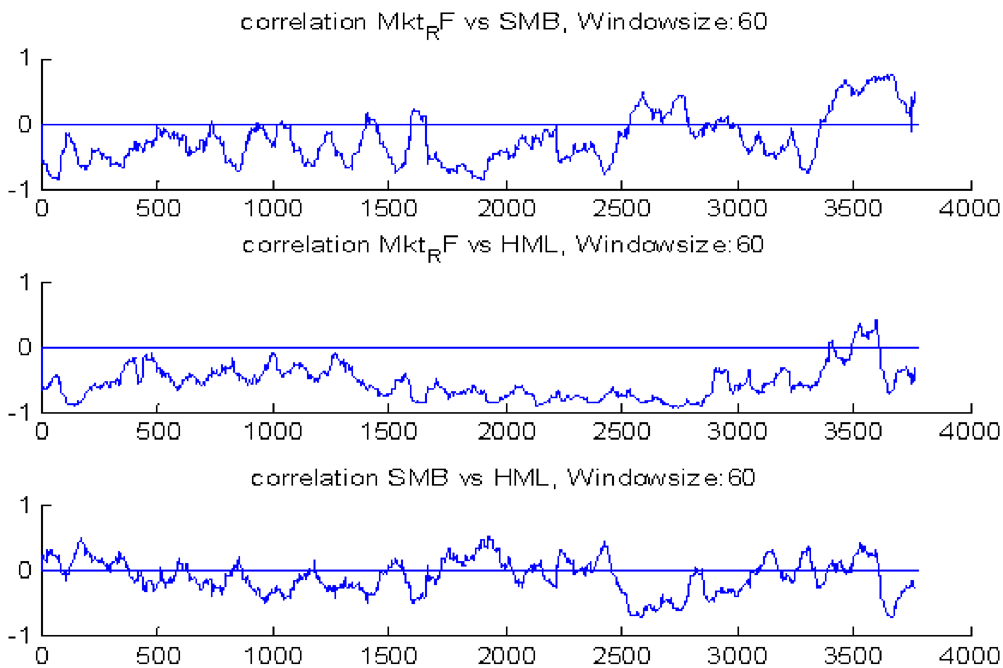


Figure 7. Correlation changes of the Fama French Factors

In this test, we compared two different rebalancing methods, both of them using the Markowitz optimization and choosing the time of rebalancing at the time the state changes. But one method takes the HMM regime into consideration when calculating the expected return and covariance, while the other method only uses the most recent data to do the calculation.

The regimes are defined by segmenting the S&P500 index return with an Hidden Markov Model. We used a two-state Markov switching model in the test.

We compared two ways of portfolio rebalancing. Both of them are mean variance portfolio rebalancing, one uses the last N days data to calculate the mean and variance, the other uses the last N days within the same state that the index is entering. Some restrictions was applied to guarantee there are enough data to calculate the estimated mean variance-covariance matrix .

Other than regimes of market index, we also tested the portfolio rebalancing under the regime switching of the Fama French Factor and the Momentum Factor.

We use the 49 Industry Portfolios from Fama French Data Library (Fama and French, 2012) to construct a portfolio, with weights calculation based on 2 choices. Method 1: without regime consideration, only use the past N days return to do the training and calculate the estimated return and covariance matrix and apply the optimized weight on the next M days. Method 2: with the regime switching in consideration, use the data from the same regime in the past for N days, to do the weight optimization and apply the weight for the next M day.

At each optimization point, we will calculate a true optimized weight set from the testing period of the next M days, which assumes we know the return and variance-covariance, what the weight will be. We will compare the optimized weights and expected performances of the above 2 methods with the true performance.

2.4.1 Data Description

The 49 Industry Portfolios we used is from Fama French Data Library. The data includes the periods from 1990/01/02 to 2011/12/30. The construction of the portfolio, as described on Fama French data library, is Assign each NYSE, AMEX, and NASDAQ stock to an industry portfolio at the end of June of year t based on its four-digit SIC code at that time. Compustat SIC codes is used for the fiscal year ending in calendar year t-1. Whenever Compustat SIC codes are not available, CRSP SIC codes is used for June of year t. Then, the returns from July of t to June of t+1 is computed. The components and descriptive statistics of the 49 industry Portfolio is in Table V.

We also applied HMM on S&P500 index data and Four factors data from 1990/01/02 to 2011/12/30 and estimated the regime probabilities for each day within the periods.

2.4.2 Portfolio Rebalance Points

With the regime labels from S&P 500 Index and each of the Fama French 3 factors and the momentum factor, we first define the rules for the time points to carry out the rebalance. Rules for rebalance point:

TABLE V

Statistics of 49 industry portfolios daily return from 1990/01/03 to 2011/12/30

Industry	Mean	Std.Dev.	Kurtosis	Skewness
Agriculture	0.048	1.63	16.34	0.55
Food Products	0.042	0.99	6.93	-0.01
Candy & Soda	0.056	1.65	8.02	-0.22
Beer & Liquor	0.050	1.26	4.75	0.19
Tobacco Products	0.062	1.65	8.99	-0.04
Recreation	0.029	1.52	6.10	-0.24
Entertainment	0.049	1.85	7.55	-0.02
Printing and Publishing	0.028	1.40	14.62	0.52
Consumer Goods	0.044	1.12	16.77	-0.76
Apparel	0.045	1.48	5.50	0.03
Health care	0.030	1.38	6.69	-0.73
Medical Equipment	0.047	1.24	5.07	-0.23
Pharmaceutical Products	0.047	1.25	4.31	-0.09
Chemicals	0.045	1.45	7.14	-0.20
Rubber and Plastic Products	0.044	1.29	5.18	-0.20
Textiles	0.022	1.66	11.83	0.58
Construction Materials	0.041	1.42	6.31	-0.05
Construction	0.046	1.88	5.69	-0.04
Steel Works Etc	0.037	2.01	8.63	-0.10
Fabricated Products	0.029	1.82	5.90	-0.06
Machinery	0.048	1.59	6.30	-0.08
Electrical Equipment	0.060	1.61	5.21	-0.01
Automobiles and Trucks	0.032	1.74	5.37	-0.03
Aircraft	0.056	1.50	9.29	-0.36
Shipbuilding, Railroad Equipment	0.053	1.67	3.80	-0.13
Defense	0.050	1.52	6.36	0.04
Precious Metals	0.040	2.56	6.63	0.73
Non-Metallic and Industrial Metal Mining	0.058	1.89	10.48	0.06
Coal	0.086	2.79	6.59	-0.07
Petroleum and Natural Gas	0.054	1.53	12.34	0.08
Utilities	0.040	1.06	16.00	0.25
Communication	0.029	1.34	8.48	0.19
Personal Services	0.034	1.40	4.89	-0.42
Business Services	0.036	1.24	5.54	-0.28
Computers	0.059	2.00	6.99	0.41
Computer Software	0.064	1.79	3.82	0.12
Electronic Equipment	0.055	2.00	4.36	0.24
Measuring and Control Equipment	0.044	1.57	4.66	0.01
Business Supplies	0.035	1.23	5.07	-0.07
Shipping Containers	0.042	1.47	4.41	-0.10
Transportation	0.042	1.37	5.84	-0.18
Wholesale	0.036	1.12	6.34	-0.21
Retail	0.047	1.35	4.61	0.13
Restaurants, Hotels, Motels	0.048	1.27	3.49	-0.01
Banking	0.045	1.83	15.30	0.46
Insurance	0.042	1.39	14.67	0.15
Real Estate	0.017	1.69	13.69	0.12
Trading	0.056	1.91	10.16	0.32
Others	0.009	1.55	9.29	-0.13

1. The regime of the market index or factors has just switched, and it has been in current regime for at least 5 days, so a confirmed regime switching is established.

2. It has been at least 60 days of data for the same regime as the current one, so to ensure we have data that's long enough to calculate expected return and variance-covariance matrix.

3. The number of days in the last opposite regimes is greater than 20,

After applying the above rules, we have around 30 rebalance points over the whole time period. The number of rebalances varies depend on the index or factors used for the regime switching model.

2.4.3 Evaluation of the Methods

At each point of the rebalance, we have the 2 training data sets selected by the two strategies. With the 2 training data sets, we have 2 set of weights calculated by the Markowitz optimization method. Each of the weight set consist of 10 sets of weights assign to each of the 49 industry portfolios. The Markowitz optimization also gives the best selected weights based on the expected return and risk.

We use the following measurements to evaluate the performance of the two methods:

1. Root mean squared error (RMSE) of the optimized weights of the two methods against the true best optimized portfolio, in form of

$$\sqrt{\frac{1}{P}\sum_{i=1}^{P}\sum_{j=1}^{N}\sum_{k=1}^{M}(w_{ijk}-\widetilde{w}_{ijk})^2}$$

where P represents the number of rebalancing that occurred, N is the number of assets constructed the portfolio, M the number of weight sets when calculated the efficient frontier, w_{ijk} represents the estimated weights by one of the methods at rebalance point i, for the jth industry portfolio, and of the jth set of optimized weight on the efficient frontier curve. $\widetilde{w}_{ijk}$ represents the best possible optimized weights.

2. RMSE of each method's estimated return-risk against the true return-risk.

$$\sqrt{\frac{1}{P}\sum_{i=1}^{P}((r_i-\widetilde{r}_i)^2+(\sigma_i-\widetilde{\sigma}_i)^2)}$$

, where P represents the number of rebalancing, r_i and σ_i represent the estimated return-risk optimized by Markowitz optimization.

3. The sharpe ratio of the portfolio.

We calculated the above measurements for each of two methods, using the regimes of the S&P500 index, 3 Fama French factors and momentum factor. Table VI shows the results of the measurements calculated. We can see the method considering the regime effect has better performance on RMSE(Risk/Return), RMSE(weights), Ratio of closer to the true weights. As for sharpe ratio (Sharpe, 1965) and (Carhart, 1997), when the regime of momentum factor is used, the regime considered method has a better performance than the non-regime considered model. In the test, when using SMB, HML, Mkt-RF and S&P500 as regime factor, both Non-regime and regime considered method have worse performance than all equal weighted portfolio. But again, when momentum factor is considered, both methods have better performance than equal weighted portfolio on sharpe ratio. The test shows the momentum factor is a good choice to be considered when apply the HMM regime switching portfolio rebalancing strategy. And for other factors, there are still many ways to improve the strategy to beat the equal weighted portfolio, such as increase the rebalance points rather than depends on regime changing points alone, or use the regime swithing combined with other traditional methods.

With these weight sets, we have drawn the efficient frontiers to show their relative risk-return position, we also drawn the efficient frontier from the next M days data so to compare with the best possible constructed portfolios.

The evaluation of the efficient frontier: When comparing the two efficient frontier curves from the two methods at one rebalance point, we do not test which method gives better expected return/risk curve. Since this is only the expected return/risk, the curve that's closer to the best possible curve generated from the next M days data will be the better one. Because it means the "expected ones" is close to the "truth". For example, if strategy 1 tells you it would expect a very high return with low risk, strategy 2 gives an expectation of a very return with high risk, and the true value indeed shows a low expected return with high risk, that means strategy 1 gives a worse prediction than strategy 2, even it has a better expectation.

To illustrate the comparison of the weight curve, we show 3 examples of efficient frontier when the index entering regime 2 from regime 1 in Figure 8.

- Figure on the left side :
  - Dots on top row are the time points where data used for method 2, the data from the same regime as start of test period.
  - Dots on the 2nd and 4th rows are the regime marks, 2nd row means regime 1, 4th row means regime 2.
  - Dots on the 3rd and 5th rows is the time points where data are used for method 1, last 60 days without considering regime.
  - The bottom is the index return series, where we can see the highly volatile periods associated with regime 2 marks.
- Figure in the middle:
  - Red circle represents the estimated efficient frontier curve from method 1, considering the regimes.
  - Blue square represents the estimated efficient frontier curve from method 2, without considering the regimes.

  - Green diamond represents the efficient frontier curve optimized using next 30 days data, which is the best optimized curve possible.
  - In each curve, the straight line point to the best selected weights under Markowitz optimization.

- Figure on the right side:

  - Red circle represents the realized efficient frontier curve using weights from method 1, considering the regimes.
  - Blue square represents the realized efficient frontier curve using weights from method 2, without considering the regimes.
  - Green diamond represents the efficient frontier curve optimized using next 30 days data, which is the best optimized curve possible. It is the same as the diamond curve in the middle figure.
  - The straight lines point to the realized risk-return using the weights selected by Markowitz optimization.

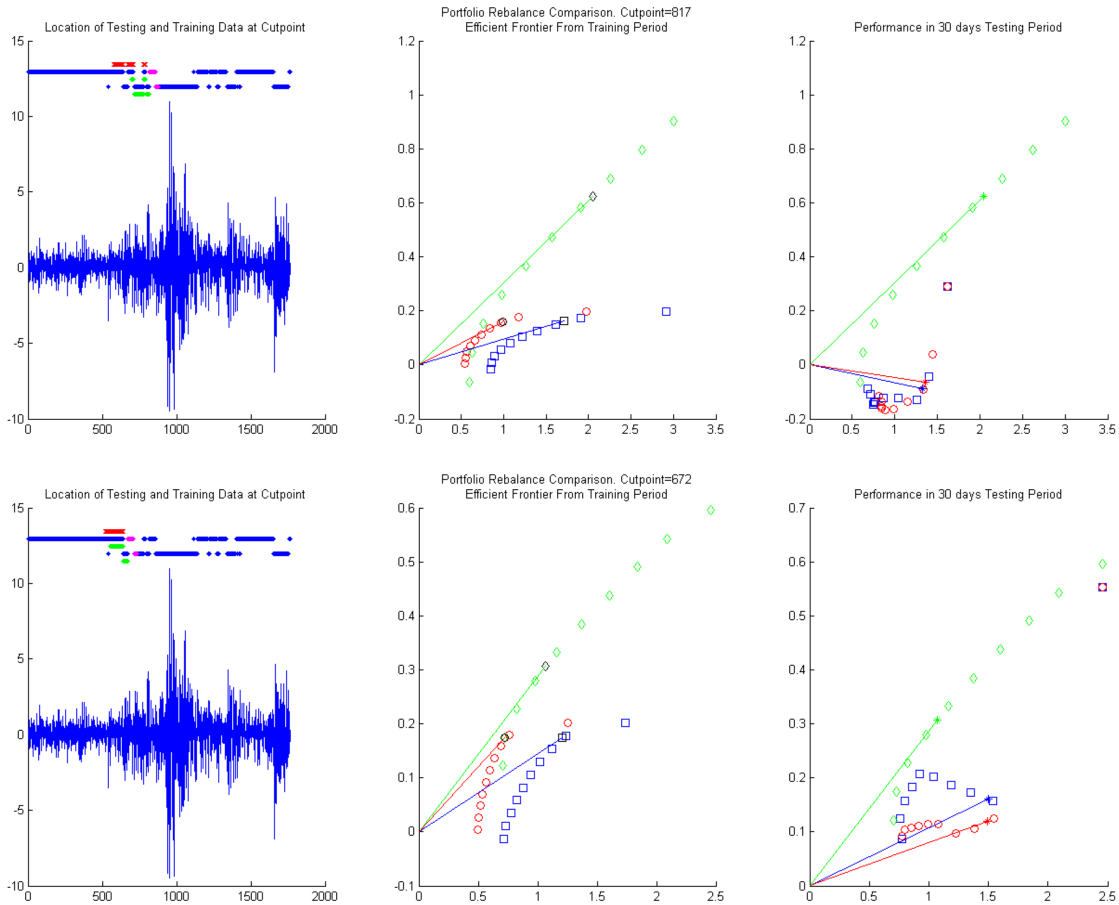


Figure 8. Efficient frontier of 2 methods comparing with best optimized weights

TABLE VI

Evaluation of the portfolio rebalance with the two methods.

Factor to Define Regime		SMB	HML	MNT	Mkt-RF	S&P500
Days to calculate weights		60	60	60	60	60
Days to apply weights		30	30	30	30	30
RMSE(Risk-Return Distance)	Non-Regime	0.2269	0.1584	0.1601	0.257	0.2631
RMSE(Risk-Return Distance)	Regime	0.2242	0.1715	0.1611	0.2318	0.2434
Ratio Closer Return/Risk		0.6	0.7391	0.5313	0.8	0.7667
RMSE(weights)	Non-Regime	0.7844	0.8649	0.8146	0.75	0.7516
RMSE(weights)	Regime	0.783	0.8626	0.8143	0.6902	0.7099
Ratio Closer Weight		0.56	0.5217	0.4063	0.6857	0.6
Sharpe Ratio	Non-Regime	0.0372	0.0142	0.0426	0.0362	0.0247
Sharpe Ratio	Regime	0.0365	0.0141	0.0457	0.0238	0.0239
Sharpe Ratio	Best Possible	0.0639	0.065	0.0763	0.0686	0.06
Sharpe Ratio	Equal Weight	0.0387	0.0387	0.0387	0.0387	0.0382
Number of re-balance		25	23	32	35	30

CHAPTER 3

REGIME SWITCHING IN U.S. GDP

3.1 Introduction

Regime switching models have long been a tool available to empirical economists. (Hamilton, 1989b) presented an approach of analyzing business cycles based on time series analysis with a hidden Markov model (HMM). During the past decade, much research has focused on defining economic recession as a statistical issue rather than a subjective qualitative assessment. The HMM applied to Gross Domestic Product (GDP) has since been adopted in many academic studies. Our purpose in this study is first to fit the HMM to U.S. GDP (mixed regular, quarterly difference of ln GDP), and the components of GDP. Also we fit the HMM to economic indicators such as non-farm employment, industrial product and study the relationship between the regime changes of these different indicators.

3.1.1 Business Cycle

The business cycle is the periodic up and downs in the economy. It is defined as a sequence of four states of the economy:

- Peak : the end of the fast growth period, the start of a contraction period.
- Trough: the end of a contraction period, before an expansion.
- Contraction: a slow or negative growth period of the economy.

- Expansion: high growth period of the economy.

Period of severe contraction usually becomes a recession.

The business cycle represents the status of the economy. It helps people to understand what is going on in the economy. It provides important guidance to the government on making economic policies. And because of that, the determination and prediction of the business cycle is important.

The US economic cycle is determined and announced by National Bureau of Economic Research (NBER) Business Cycle Dating Committee.

The announcements in the past few years from the NBER are as follow (from NBER web site):

- The June 2009 trough was announced September 20, 2010.
- The December 2007 peak was announced December 1, 2008.
- The November 2001 trough was announced July 17, 2003.
- The March 2001 peak was announced November 26, 2001.
- The March 1991 trough was announced December 22, 1992.

The period from 1991 to 2001 is the longest expansion in the U.S. history according to NBER.

NBER uses GDP and other macroeconomic variables including unemployed rate, index of industrial production, etc. to determine a business cycle. Although a conventional definitions of recession is two consecutive quarters of negative growth rate of GDP, NBER does not use a fixed definition to determine recession. The NBER's definition of a recession is "A period between a peak and a trough, and an expansion is a period between a trough and a peak. During a recession, a significant decline in economic activity spreads across the economy and can last from a few months to more than a year".

3.1.2 GDP

GDP is the most important economic indicator. It is used by the government to improve policy making and to prepare the federal budget. It's also closely watched by Wall Street as the indicator of economic activity. It's used by the businesses to prepare forecasts on the economy and make decisions on production, investments, spending and other activities. GDP is a quarterly figure, defined as the value of goods and services produced by labor and property located in the United States, whether the labor and property are domestically or foreign owned, and it reflects income as well as expenditure flows. It is produced by the national income and product account(s) (NIPA), and released by Bureau of Economic Analysis (BEA), on the last day of each quarter and reflects the previous quarter. After the initial release, it will be revised twice before the final figure.

The growth rate of GDP is the main figure people paying attention to. In a normal or healthy economic condition, the U.S. GDP growth rate should be around 3 percent yearly. If the growth rate is too low, or even be negative, it will lead to increased unemployment and lower spending, and lower productivity. The government will take action to stimulate the economy, by various policies. However, if the growth rate is too high, the economy is considered overheat, and will be unsustainable and high inflation may occur. The government will also take action to slow down the economy.

3.2 HMM on GDP And Estimation and Prediction of Business Cycle

3.2.1 Description of GDP Growth Rate Data

We use quarterly GDP from first quarter of 1947 to 3rd quarter of 2011 ($n = 255$ quarters) from the Bureau of Economic Affairs.

Continuous Growth Rate

Denote the series by $P_t, t = 1, 2, \ldots, n = 255$. Let $L_t = \ln P_t$. We analyze

- The first difference $V_t = L_t - L_{t-1} = \nabla L_t$, the "velocity"
- The second difference $A_t = V_t - V_{t-1} = (L_t - L_{t-1}) - (L_{t-1} - L_{t-2}) = L_t - 2L_{t-1} + L_{t-2} = \nabla^2 L_t$, the "acceleration."

The quantity $L_t - L_{t-1} = \ln P_t - \ln P_{t-1}$ is known as the *continuous rate of return,* for the following reason. Suppose that during the time interval $(t-1, t]$, a principal amount of P_{t-1} earned interest at rate r_t, compounded N times. Then the initial amount P_{t-1} has grown to an amount P_t given by

$$P_t = P_{t-1}(1 + r_t/N)^N.$$

As N tends to infinity, corresponding to continuous compounding, the quantity $(1+r_t/N)^N$ tends to e^{r_t}. This gives $P_t = P_{t-1}e^{r_t}$. The quantity r_t is be the continuous rate. To solve for r_t, take the natural logarithm of both sides, obtaining $\ln P_t = \ln P_{t-1} + r_t$; solving for r_t gives the continuous rate of return as

$$r_t = \ln P_t - \ln P_{t-1}.$$

Distribution of U.S. GDP Growth Rate

Figure 10 shows the distribution of U.S. GDP growth rate over all periods, in recession and in expansion. We can see from the histogram, in expansion regimes, the growth rate is slightly right skewed and has positive mean value. And in recession regimes, the growth rate is left skewed and has negative mean value.

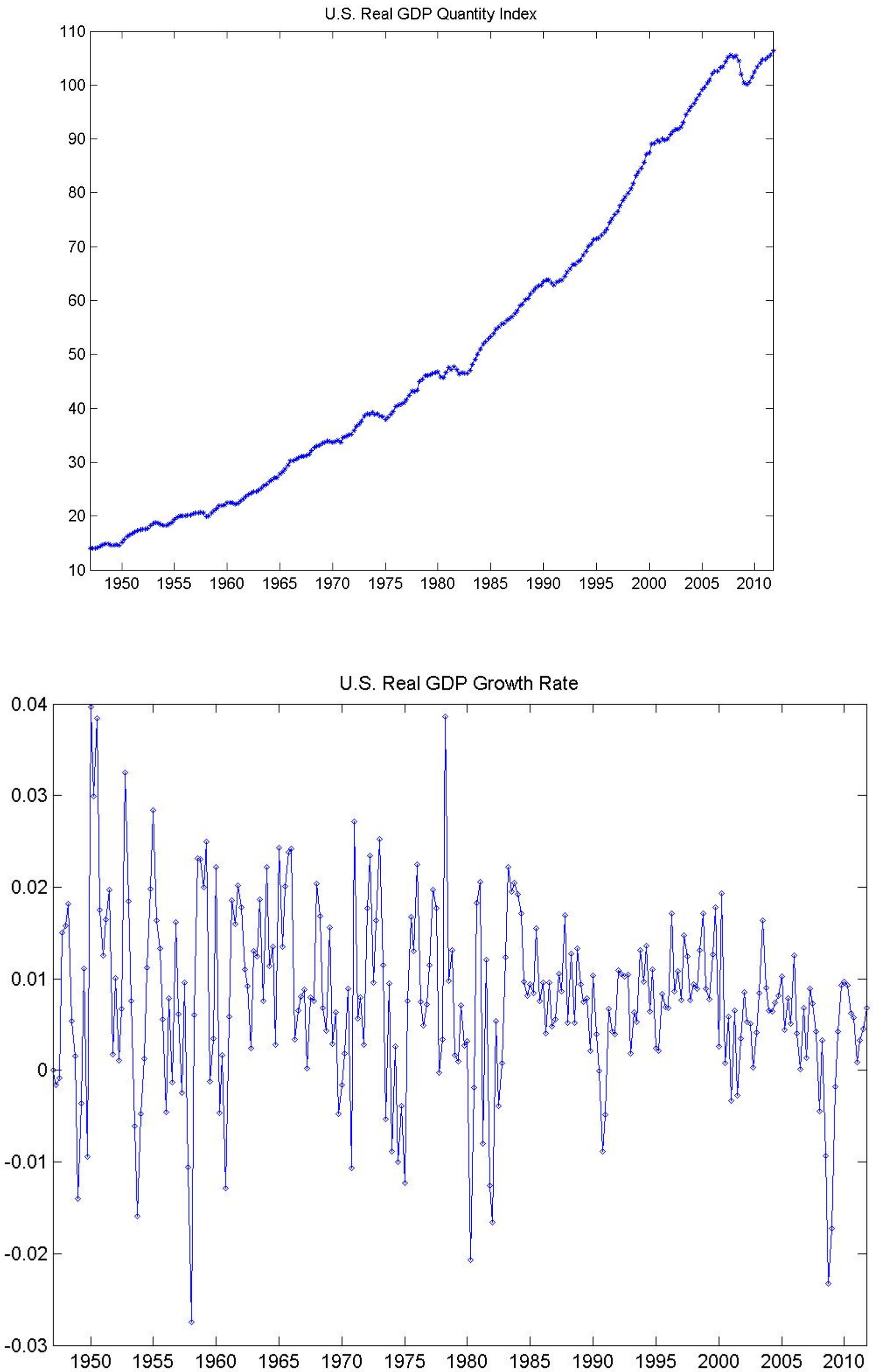


Figure 9. U.S. GDP index and growth rate (1990-2011)

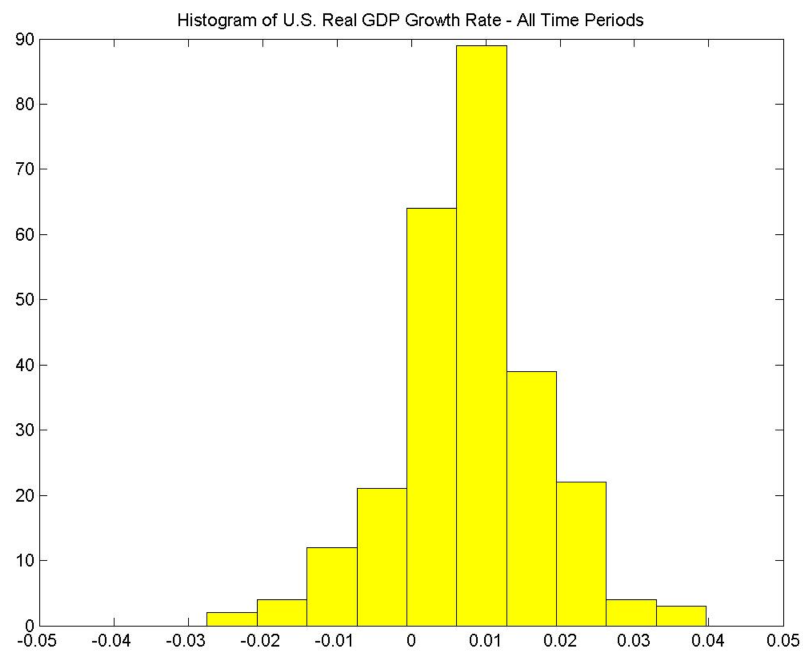


Figure 10. Histogram of U.S. GDP growth rate for all time periods.

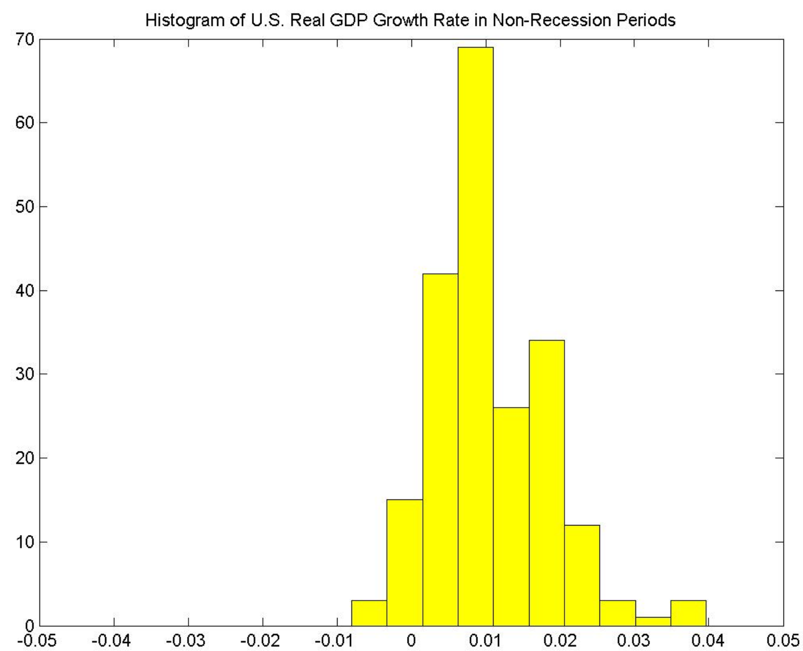


Figure 11. Histogram of U.S. GDP growth rate for non-recession periods

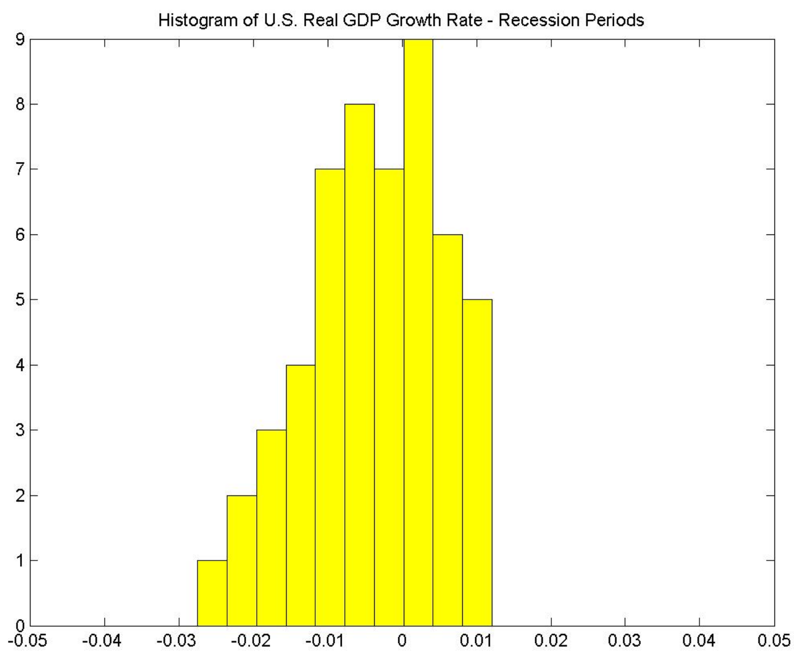


Figure 12. Histogram of U.S. GDP growth rate for recession periods

3.2.2 Model Estimation

3.2.2 shows the GDP HMM parameters estimated. We can see the estimated mean quarterly return for regime 1 is negative -0.059, which associated with non growth or negative growth periods. The mean return for regime 2 is positive 1.11, which represents normal positive growth periods. The transition probability from regime 1 to 1 is 0.757, and from regime 2 to 2 is 0.919, which means once the GDP enters a regime, it tends to stay in that regime. These are inline with the facts that one business cycle doesn't complete in a short period like one or two quarters. And the transition probability from regime 2 to 2 is larger than the probability from regime 1 to 1, associated with the fact the GDP stays in a positive growth regime longer than a non growth regime. Based on these facts, we can associate regime 1 from the HMM with the recession periods, and regime 2 for non-recession periods.

Figure 13 shows the probability of recession from Hidden Markov Model mapped with NBER announced recessions. We can see the HMM probabilities matches with the known recession periods in the past.

Figure 14 shows the GDP growth series, probability of each of the 2 states and NBER announced recession together. We can see mostly negative growth rate periods matches with NBER announced recession periods, but there are many quarters with negative growth rates that are not in the recession periods.

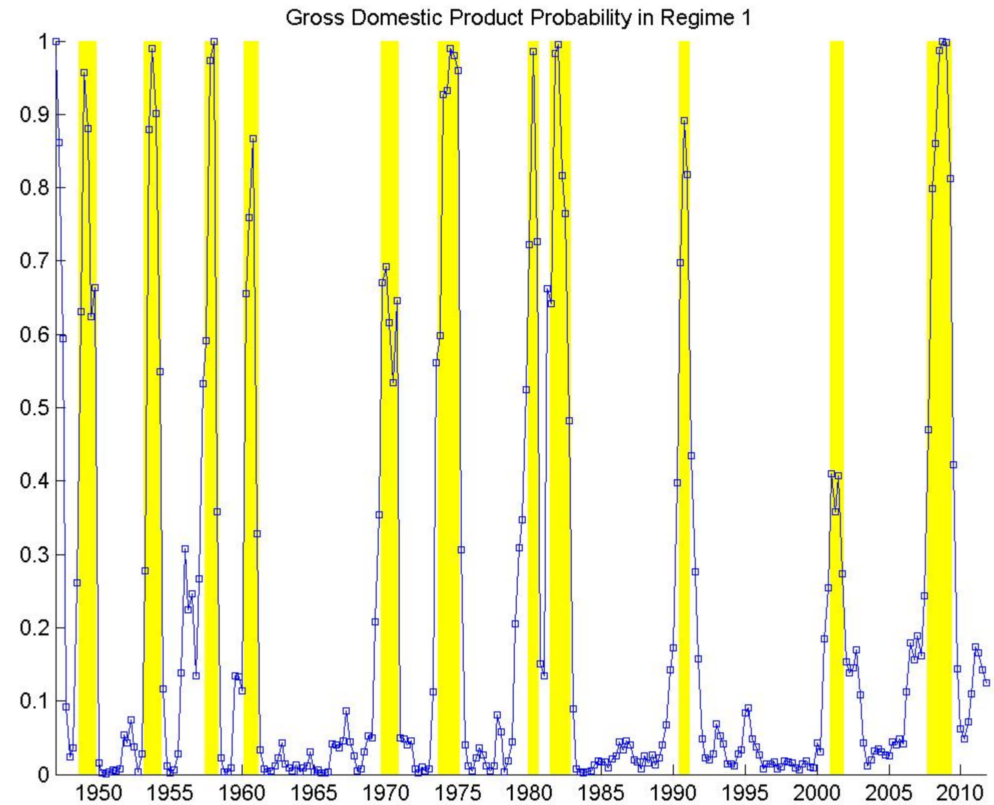


Figure 13. Probability of recession regimes and NBER recessions

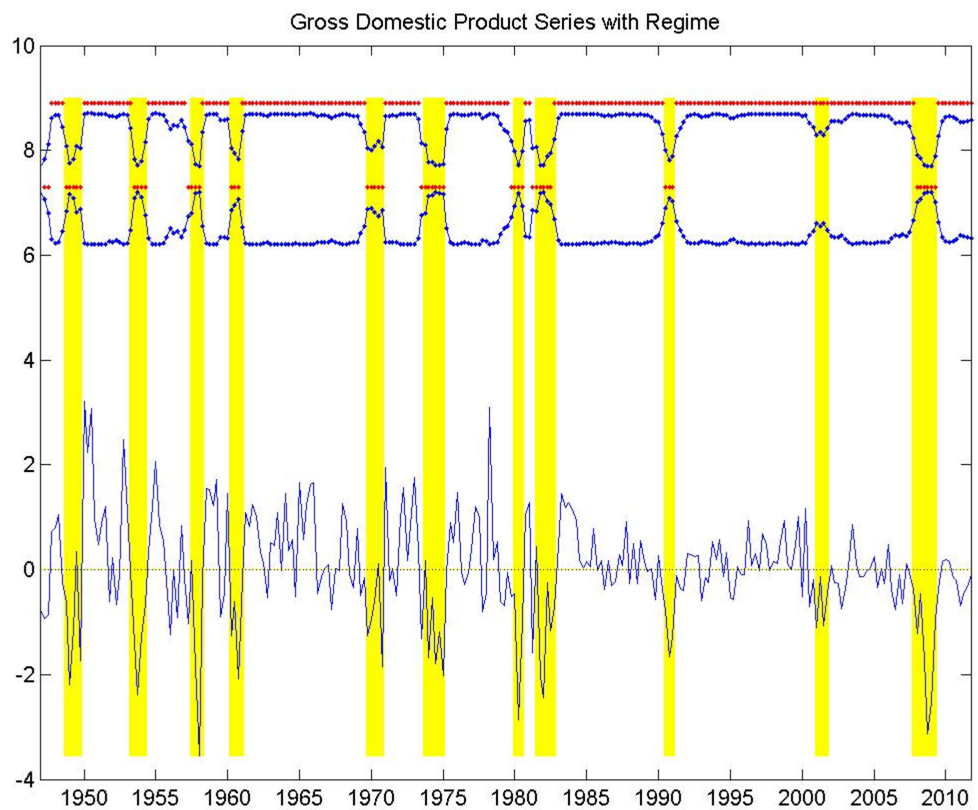


Figure 14. GDP Growth Rate, Recession Regimes, NBER recessions together

Regime	μ	σ
1	-0.059046	0.86736
	(-0.23511)	(3.6714)
2	1.1109	0.65882
	(11.3404)	(8.2043)
Trans. Prob.	P_{11}	P_{22}
	0.75721	0.91979
	(7.88)	(26.5528)

TABLE VII

HMM Parameters Estimated for U.S.GDP

3.2.3 Predict Business Cycle with Regime Probability of GDP

HMM Result and Prediction of Business Cycle

From the above results, we can see the HMM model does do a good job on labeling the past recession periods. And probability of regime 1 matches with the accounted recession periods. But Hidden Markov Model uses all past information to do the estimation, include the probability of regimes. That means for a time period at t, the information of $t+1$ to end of the time period T is used. So the prediction ability of HMM is a question, especially if the model within each HMM state gives no prediction to the future, for example, a mean mixture model instead than an AR model.

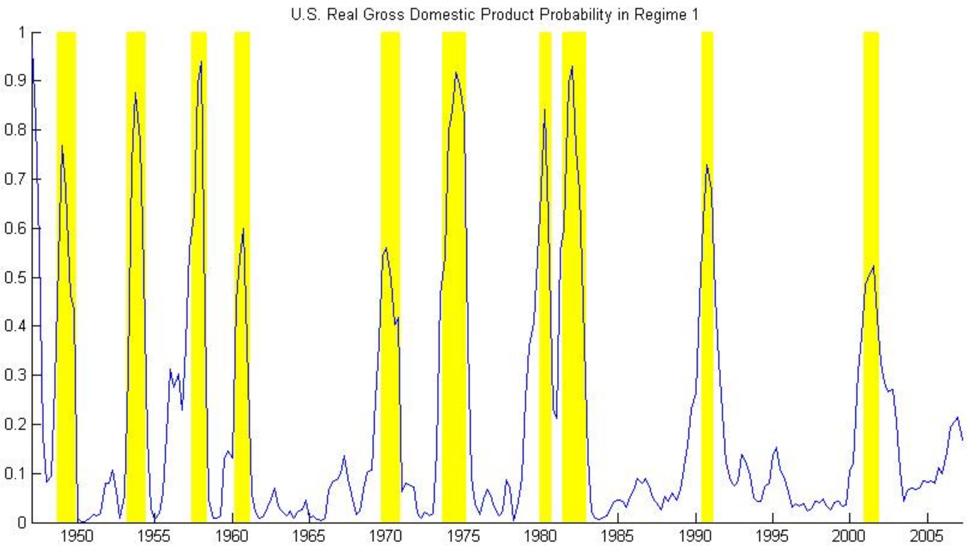
One possible solution for prediction from HMM is use the transition probability. An high transition probability to certain state S_{T+1} from current state means the value of interesting of the next time period $T+1$ can be predicted using model under state S_{T+1}.

Another solution we proposed is using the time series pattern of the estimated probability to do the prediction. In the HMM segmentation of US GDP, we can see at the beginning of a certain recession, even though the probability of being in a recession from the HMM model is below 0.5, which can be simply interpreted as not in an recession. But look back to the past probability patterns, the momentum of the regime probability value may tell us if a recession is on the way. The probability of regime 1 shows every time the probability rise above 0.2, an recession always occur within the next year. And the probability of regime 1 is usually rising for 3 quarters. Using the pattern of the probability can give us an way of prediction, other than depending on transition probability and predictive model within each state. Figure 15 shows at the end of 2007, the HMM gives probability of regime 1 up to 0.2, and after that, around year 2009, another recession happened. That is an example of such prediction of recession from Hidden Markov Model.

3.3 HMM on GDP Components and the Relationship with GDP Regimes

In this section we used Hidden Markov Model for each of the U.S. GDP components. The purpose is to test if these components also show regimes switching effect, and if any of the the components are leading or lagging the whole GDP on the regime changes. The results may help predicting the whole GDP, or growth of certain area of the economy.

Figure 15. HMM probability of recession at end of 2007



3.3.1 Data Description of U.S. GDP Components

The Structure of GDP components

Table VIII shows the U.S. GDP Components' names and structure. To save the space in the following tables, we used abbreviations for some of the components.

- FI stands for: Fixed Investment
- PCE stands for: Personal Consumption Expenditures
- GPDI stands for: Gross Private Domestic Investment
- GCEAGI stands for: Government Consumption

The abbreviations listed will be used in all the following sections.

Also, the components we showed here included sub-components in different level. Sub-components are labeled as (Main Component:Sub Component). For example: Durable goods component is under Goods component, who is under Personal Consumption Expenditures component, so it's labeled as "PCE:Goods:Durable goods" in the tables.

Descriptive Statistics of GDP components

To see each of the components' growth rate distribution under different business cycles. we calculated summary of statistics on several measures. Table IX shows the descriptive statistics summary of GDP growth rate over all periods. Table X shows the summary over non recession periods. Table XI shows the summary over recession periods. We can see that in general every components have a lower than usual growth rate in recession periods.

TABLE VIII

U.S. GDP Components Names and Structure

Personal Consumption Expenditures
PCE:Goods
PCE:Goods:Durable goods
PCE:Goods:Nondurable goods
PCE:Services
Gross Private Domestic Investment
GPDI:Fixed Investment
GPDI:FI:Nonresidential
GPDI:FI:NR:Structures
GPDI:FI:NR:Equipment and Software
GPDI:FI:Residential
GPDI:Exports
GPDI:Exp:Goods
GPDI:Exp:Services
GPDI:Imports
GPDI:Imp:Goods
GPDI:Imp:Services
Government Consumption Expenditures And Gross Investment
GCEAGI:Federal
GCEAGI:Fed:National Defense
GCEAGI:Fed:Nondefense
GCEAGI:State and Local

FI is short for Fixed Investment.
PCE is short for Personal Consumption Expenditures.
GPDI is short for Gross Private Domestic Investment.
GCEAGI is short for Government Consumption Expenditures And Gross Investment

TABLE IX

Descriptive Statistics of GDP Growth Rate For All Periods (1947Q1-2011Q3)

All Periods	Mean	Stdev	Skewness	Min	Max
Personal Consumption Expenditures	0.828	0.849	-0.330	-3.011	5.135
PCE:Goods	0.832	1.418	-0.477	-5.268	7.487
PCE:Goods:Durable goods	1.342	3.791	0.369	-14.350	23.702
PCE:Goods:Nondurable goods	0.619	0.803	-0.231	-2.219	3.711
PCE:Services	0.856	0.525	-0.217	-0.741	2.731
Gross Private Domestic Investment	1.037	5.470	0.059	-17.564	25.092
GPDI:Fixed Investment	0.895	2.718	-0.366	-9.259	9.282
GPDI:FI:Nonresidential	1.058	2.742	-0.442	-9.041	9.001
GPDI:FI:NR:Structures	0.484	3.024	-0.541	-9.618	9.779
GPDI:FI:NR:Equipment and Software	1.328	3.432	-0.618	-14.801	12.560
GPDI:FI:Residential	0.547	5.102	0.179	-18.501	19.327
GPDI:Exports	1.279	4.345	0.233	-13.438	22.410
GPDI:Exp:Goods	1.284	4.957	0.375	-14.711	26.641
GPDI:Exp:Services	1.558	6.596	1.524	-23.797	53.963
GPDI:Imports	1.545	4.044	0.695	-11.315	23.373
GPDI:Imp:Goods	1.620	4.964	0.941	-13.626	28.964
GPDI:Imp:Services	1.411	4.156	0.829	-15.087	20.896
Government Consumption Expenditures And Gross Investment	0.716	1.822	2.992	-3.540	13.751
GCEAGI:Federal	0.642	3.114	2.655	-5.393	21.407
GCEAGI:Fed:National Defense	0.573	3.700	2.988	-7.069	26.755
GCEAGI:Fed:Nondefense	1.054	6.589	0.666	-27.621	34.965
GCEAGI:State and Local	0.813	1.088	0.639	-2.705	5.701

TABLE X

Descriptive Statistics of GDP Growth Rate in Non-recession Periods (1947Q1-2011Q3)

Non Recession Periods	Mean	Stdev	Skewness	Min	Max
Personal Consumption Expenditures	0.992	0.771	-0.150	-3.011	5.135
PCE:Goods	1.084	1.297	-0.369	-5.268	7.487
PCE:Goods:Durable goods	1.846	3.601	0.691	-14.350	23.702
PCE:Goods:Nondurable goods	0.766	0.733	-0.052	-2.219	3.711
PCE:Services	0.948	0.461	0.216	-0.741	2.731
Gross Private Domestic Investment	2.362	4.642	0.839	-17.564	25.092
GPDI:Fixed Investment	1.631	2.233	0.233	-9.259	9.282
GPDI:FI:Nonresidential	1.753	2.306	-0.128	-9.041	9.001
GPDI:FI:NR:Structures	0.905	2.856	-0.480	-9.618	9.779
GPDI:FI:NR:Equipment and Software	2.174	2.943	-0.586	-14.801	12.560
GPDI:FI:Residential	1.444	4.482	0.654	-18.501	19.327
GPDI:Exports	1.710	4.120	0.369	-13.438	22.410
GPDI:Exp:Goods	1.742	4.766	0.405	-14.711	26.641
GPDI:Exp:Services	1.956	6.719	1.860	-23.797	53.963
GPDI:Imports	2.257	3.851	0.996	-11.315	23.373
GPDI:Imp:Goods	2.434	4.825	1.198	-13.626	28.964
GPDI:Imp:Services	1.757	4.075	1.005	-15.087	20.896
Government Consumption Expenditures And Gross Investment	0.745	1.896	3.322	-3.540	13.751
GCEAGI:Federal	0.705	3.274	2.885	-5.393	21.407
GCEAGI:Fed:National Defense	0.649	3.951	3.085	-7.069	26.755
GCEAGI:Fed:Nondefense	1.013	6.452	0.747	-27.621	34.965
GCEAGI:State and Local	0.731	0.954	0.099	-2.705	5.701

TABLE XI

Descriptive Statistics of GDP Growth Rate in Recession Periods (1947Q1-2011Q3)

Recession Periods	Mean	Stdev	Skewness	Min	Max
Personal Consumption Expenditures	0.175	0.837	-0.489	-2.270	1.824
PCE:Goods	-0.170	1.449	-0.492	-4.121	3.645
PCE:Goods:Durable goods	-0.662	3.898	-0.148	-10.786	8.688
PCE:Goods:Nondurable goods	0.034	0.808	-0.266	-1.785	1.633
PCE:Services	0.490	0.606	-0.138	-0.741	1.729
Gross Private Domestic Investment	-4.236	5.372	-0.253	-17.564	8.765
GPDI:Fixed Investment	-2.032	2.511	-0.807	-9.259	3.590
GPDI:FI:Nonresidential	-1.708	2.608	-0.448	-8.971	2.556
GPDI:FI:NR:Structures	-1.191	3.122	-0.616	-9.618	7.356
GPDI:FI:NR:Equipment and Software	-2.041	3.181	-0.501	-10.749	3.515
GPDI:FI:Residential	-3.020	5.856	0.376	-18.501	12.280
GPDI:Exports	-0.438	4.815	0.256	-13.438	14.381
GPDI:Exp:Goods	-0.538	5.320	0.590	-13.673	17.926
GPDI:Exp:Services	-0.028	5.873	-0.581	-18.380	12.758
GPDI:Imports	-1.286	3.551	0.093	-10.027	8.977
GPDI:Imp:Goods	-1.618	4.156	0.139	-11.651	9.975
GPDI:Imp:Services	0.035	4.228	0.440	-9.532	14.225
Government Consumption Expenditures And Gross Investment	0.598	1.504	-0.087	-3.540	4.527
GCEAGI:Federal	0.391	2.381	-0.537	-5.355	5.421
GCEAGI:Fed:National Defense	0.274	2.464	-0.787	-6.705	4.996
GCEAGI:Fed:Nondefense	1.220	7.176	0.408	-19.115	21.937
GCEAGI:State and Local	1.143	1.475	0.798	-1.405	5.701

3.3.2 Model Estimation and Empirical Study

The same as using HMM on the whole GDP series, we are interested in the growth rate under different regime, and expect one regime associated with the recession periods. So we use 2 states Hidden Markov Model, and a mean-variance model under each of the states.

3.3.2.1 Parameter Estimation

Table XII and Table XIII show the parameters estimated for each of the GDP components. We can see the mean value of the growth rate under regime 1 are much smaller than those under regime 2. And the mean growth rate estimated tend to be higher than the mean value from the descriptive statistics in Table XI. That can be explained by the hard cut of the NBER recession periods.

To confirm the probability of regime 1 matching with recession periods, we again graph the probabilities with the NBER announced recession periods which will be shown in later paragraphs.

TABLE XII

Parameters estimated for components of GDP Part 1.

	$\mu 1$	$\mu 2$	$P11$	$P22$	$\sigma 1$	$\sigma 2$	loglik
Personal Consumption Expenditures	0.35087	0.91406	0.49604	0.95113	3.8445	0.35552	-44.35
	(0.73034)	(19.1328)	(2.0439)	(35.1669)	(2.0956)	(7.3491)	
PCE:Goods:Durable goods	-0.18834	1.5069	0.36634	0.91209	73.427	7.0583	-421.46
	(-0.069272)	(7.2107)	(2.0135)	(16.6689)	(2.3826)	(5.5623)	
PCE:Goods:Nondurable goods	0.62698	0.70237	0.93403	0.93992	0.9846	0.20795	-38.563
	(6.2959)	(14.2405)	(22.7922)	(27.0037)	(5.8344)	(4.5917)	
PCE:Services	0.62368	1.1566	0.92302	0.93859	0.18422	0.18467	58.347
	(9.8602)	(20.3775)	(18.1639)	(23.0588)	(5.891)	(5.0824)	
Gross Private Domestic Investment	0.71797	1.1199	0.94249	0.96898	64.511	10.265	-497.54
	(0.7696)	(3.8393)	(27.378)	(58.5459)	(4.8923)	(6.1368)	
GPDI:Fixed Investment	-0.07622	2.1097	0.91514	0.90107	8.7063	2.0578	-329.64
	(-0.14507)	(9.5623)	(24.8347)	(26.1715)	(6.4818)	(4.3804)	
GPDI:FI:Nonresidential	-0.084693	2.206	0.88588	0.88617	9.4314	1.9576	-331.69
	(-0.23645)	(12.4277)	(20.4853)	(22.177)	(6.6746)	(4.2974)	
GPDI:FI:NR:Structures	-3.4411	1.1993	0.73838	0.9595	6.1286	4.8706	-347.09
	(-4.8587)	(6.7482)	(8.0718)	(52.289)	(3.2662)	(9.2967)	
GPDI:FI:NR:Equipment and Software	-0.23202	2.2074	0.86283	0.91522	20.216	3.7524	-387.66
	(-0.42711)	(7.6879)	(13.8136)	(25.6607)	(4.4742)	(5.4323)	
GPDI:FI:Residential	0.33578	1.1141	0.93573	0.92558	43.254	3.8939	-466.01
	(0.55828)	(3.6539)	(26.7906)	(31.6978)	(6.009)	(4.983)	

TABLE XIII

Parameters estimated for components of GDP Part 2.

	$\mu1$	$\mu2$	$P11$	$P22$	$\sigma1$	$\sigma2$	loglik
GPDI:Exports	0.69634	1.754	0.96345	0.9587	32.133	2.598	-422.58
	(1.385)	(10.0962)	(34.6316)	(36.6785)	(6.5937)	(6.1763)	
GPDI:Exp:Goods	0.482	1.7317	0.92051	0.9333	47.136	3.9292	-455.91
	(0.71885)	(8.6948)	(23.4843)	(33.0782)	(6.4436)	(6.5223)	
GPDI:Exp:Services	1.4462	1.1982	0.97513	0.94759	8.232	115.79	-497.84
	(6.1585)	(0.97651)	(63.0246)	(25.0457)	(6.9774)	(5.1575)	
GPDI:Imports	1.7455	1.2959	0.9439	0.92716	3.4315	31.051	-412.21
	(8.4213)	(2.2805)	(32.1464)	(20.2449)	(4.8201)	(5.2476)	
GPDI:Imp:Goods	1.7676	1.4018	0.94985	0.94638	3.8506	43.488	-452.85
	(8.1201)	(2.2734)	(28.6807)	(23.7967)	(4.9284)	(6.2442)	
GPDI:Imp:Services	1.0627	2.7111	0.99544	0.98758	6.7887	56.397	-410.61
	(5.5398)	(2.4792)	(182.8052)	(49.0823)	(9.216)	(4.8087)	
GCEAGI	0.52753	2.334	0.99486	0.95705	1.1647	15.968	-185.11
	(6.9052)	(3.0114)	(185.225)	(22.5117)	(9.7421)	(3.6189)	
GCEAGI:Federal	0.29375	2.9191	0.99528	0.95741	4.0613	45.916	-329.38
	(2.0567)	(1.9074)	(206.5627)	(21.1549)	(8.6401)	(3.235)	
GCEAGI:Fed:National Defense	0.11234	14.265	0.99566	0.83331	6.3178	50.207	-359
	(0.65625)	(4.1198)	(229.6284)	(5.5021)	(10.7123)	(1.6567)	
GCEAGI:Fed:Nondefense	0.65609	1.1794	0.9721	0.9457	5.5473	123.5	-478.23
	(3.4272)	(0.94055)	(60.4016)	(30.1219)	(6.554)	(6.0697)	
GCEAGI:State and Local	0.61676	1.1526	0.97236	0.97245	0.33166	1.7402	-104.2
	(9.1556)	(9.103)	(46.8429)	(45.9833)	(6.1602)	(7.2792)	

PCE is short for Personal Consumption Expenditures
GPDI is short for Gross Private Domestic Investment
GCEAGI is short for Government Consumption Expenditures And Gross Investment

3.3.2.2 Conditional Probability of Components Regimes and GDP Regimes

Table of the conditional probability 3.3.2.2 shows the conditional probabilities based on GDP components regimes and GDP regimes. Three measurements are calculated for each component.

1. $P1 = P(Irec(t) = 1|Sc(t-1) = 1)$, the probability a NBER recession happened on quarter t when Component is in regime 1 on quarter $t-1$.

2. $P2 = P(Sc(t-1) = 1|Irec(t) = 1)$, the probability that Component is in regime 1 on quarter $t-1$ on condition of NBER recession happened on quarter t.

3. $Cor1 = Correlation(Pgdp1, lag1(Pcomp1))$, where Pgdp1 is the regime 1 probability series from HMM estimated on GDP, lag1(Pcomp1) is the lagged regime 1 probability series from HMM estimated on the GDP component.

We use the above three measurement to discover the leading components for recession periods. The components that have high conditional probability and correlation as described above will be good candidates.

TABLE XIV

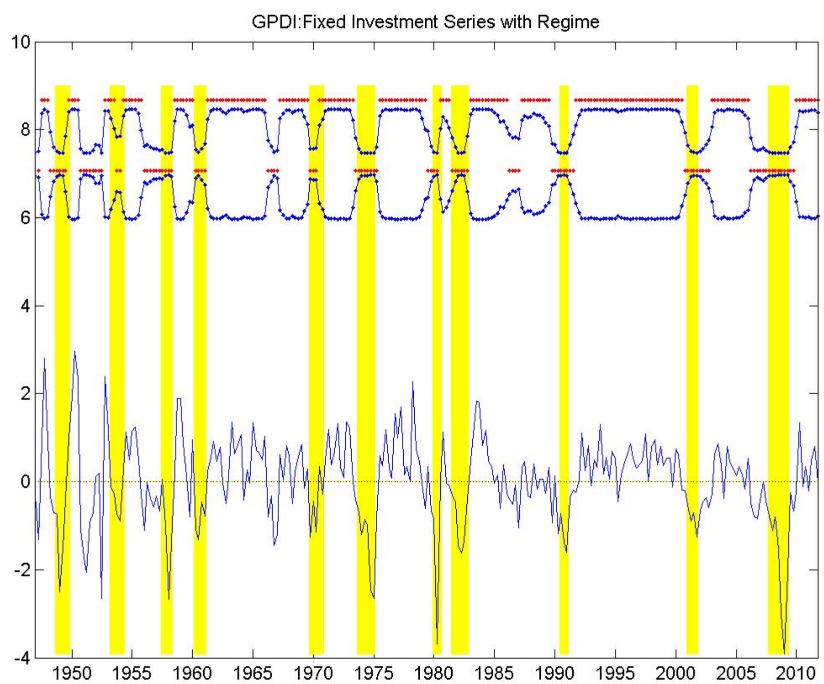
Conditional probability and correlation measurement for recession leading components

	$P1$	$P2$	$Corr1$
Personal Consumption Expenditures	0.714	0.385	0.498
PCE:Goods	0.500	0.096	0.335
PCE:Goods:Durable goods	0.294	0.096	0.140
PCE:Goods:Nondurable goods	0.263	0.673	0.254
PCE:Services	0.314	0.635	0.407
Gross Private Domestic Investment	0.380	0.577	0.373
GPDI:Fixed Investment	0.484	0.865	0.662
GPDI:FI:Nonresidential	0.368	0.827	0.408
GPDI:FI:NR:Structures	0.333	0.192	0.165
GPDI:FI:NR:Equipment and Software	0.411	0.750	0.464
GPDI:FI:Residential	0.617	0.558	0.641
GPDI:Exports	0.287	0.712	0.211
GPDI:Exp:Goods	0.315	0.673	0.245
GPDI:Exp:Services	0.229	0.365	-0.005
GPDI:Imports	0.345	0.731	0.368
GPDI:Imp:Goods	0.302	0.750	0.296
GPDI:Imp:Services	0.178	0.731	-0.144
Government Consumption Expenditures And Gross Investment	0.180	0.808	-0.080
GCEAGI:Federal	0.204	1.000	0.091
GCEAGI:Fed:National Defense	0.205	1.000	0.104
GCEAGI:Fed:Nondefense	0.188	0.654	-0.049
GCEAGI:State and Local	0.197	0.654	0.002

$P1 = P(Irec(t) = 1|Sc(t-1) = 1)$, $P2 = P(Sc(t-1) = 1|Irec(t) = 1)$,$Cor1 = Corr(Pgdp1(t+1), Pcomp1(t))$

Using 0.4 as the cut off value for all three measurements, from the table we can see 3 components are good at predicting the recession:"GPDI:Fixed Investment","GPDI:FI:NR:Equipment and Software" and "GPDI:FI:Residential". Because the last two are the sub components of the first one. We will keep checking the GPDI:Fixed Investment component alone. Figure 16 show the probability and time series graph with NBER recession periods, and we can see the periods of regime 1 for GPDI:Fixed Investment does tend to be in front of the NBER announced recession periods.

Figure 16. Regimes with growth rate series: GPDI:Fixed Investment



CITED LITERATURE

[Ang and Timmermann, 2011] Ang, A. and Timmermann, A.: Regime changes and financial markets. NBER Working Papers 17182, National Bureau of Economic Research, Inc, June 2011.

[Aray, 2008] Aray, H.: Effects of macroeconomic announcements on stock returns across volatility regimes. ThE Papers 08/17, Department of Economic Theory and Economic History of the University of Granada., December 2008.

[Baum et al. , 1970] Baum, L. E., Petrie, T., Soules, G., and Weiss, N.: A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. Annals of Mathematical Statistics, 41:164–171, 1970.

[Beine et al. , 2003] Beine, M., Laurent, S., and Lecourt., C.: Official central bank interventions and exchange rate volatility: Evidence from a regime-swichting analysis. European Economic Review, 47, 891-911., 2003.

[Bhar and Hamori., 2004] Bhar, R. and Hamori., S.: Empirical characteristics of the permanent and transitory components of stock return: Analysis in a Markov switching heteroskedasticity framework. Economics Letters, 82, 157-165., 2004.

[Bollen et al. , 2000] Bollen, N., Gray, S. F., and Whaley, R. E.: Regime-switching in foreign exchange rates: Evidence from currency option prices. Journal of Econometrics, 94:239–276, 2000.

[Brunetti et al. , 2003] Brunetti, C., Mariano, R. S., Scotti, C., and Tan., A. H. H.: Markov switching garch models of currency crises in southeast Asia. Working Paper 03-008, Penn institute for Economic Research., 2003.

[Carhart, 1997] Carhart, M. M.: On persistence in mutual fund performance. Journal of Finance, 52(1):57–82, March 1997.

[Chang and Feigenbaum, 2008] Chang, G. and Feigenbaum, J.: Detecting log-periodicity in a regime-switching model of stock returns. Quantitative Finance, 8(7):723–738, 2008.

[Chang, 2009] Chang, K.-L.: Do macroeconomic variables have regime-dependent effects on stock return dynamics? evidence from the Markov regime switching model. Economic Modelling, 26(6):1283–1299, November 2009.

[Chen, 2001] Chen, L.-H.: Inflation, real short-term interest rates, and the term structure of interest rates: a regime-switching approach. Applied Economics, 33(3):393–400, 2001.

[Chkili and Nguyen, 2011] Chkili, W. and Nguyen, D. K.: Modeling the volatility of Mediterranean stock markets: a regime-switching approach. Economics Bulletin, 31(2):1105–1113, 2011.

[Dempster et al. , 1977] Dempster, A. P., Laird, N. M., and Rubin, D. B.: Maximum likelihood from incomplete data via the E-M algorithm. Journal of the Royal Statistical Society, Series B, 39:1–38, 1977.

[Dueker, 1997] Dueker, M. J.: Markov switching in GARCH processes and mean-reverting stock-market volatility. Journal of Business and Economic Statistics, 15, 26-34., 1997.

[Engel, 1994] Engel, C.: Can the Markov switching model forecast exchange rates? Journal of International Economics, 36, 151-165., 1994.

[Engel and Hamilton., 1990] Engel, C. and Hamilton., J. D.: Long swings in the dollar: Are they in the data and do markets know it? American Economic Review, 80, 689-713., 1990.

[Erlandsson, 2002] Erlandsson, U.: Regime switches in Swedish interest rates. Working Papers 2002:5, Lund University, Department of Economics, February 2002.

[Fama and French, 1992] Fama, E. F. and French, K. R.: The cross-section of expected stock returns. Journal of Finance, XLVII:427–465, 1992.

[Fama and French, 2012] Fama, E. F. and French, K. R.: Fama/French data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data$_library.html$, 2012.

[Forney, 1973] Forney, G. D., J.: The Viterbi algorithm. Proceedings of the IEEE, 61:268–278, 1973.

[Gamerman, 1997] Gamerman, D.: Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. New York, Chapman and Hall, 1997.

[Gruss and Mertens, 2009] Gruss, B. and Mertens, K.: Regime switching interest rates and fluctuations in emerging markets. Technical report, 2009.

[Hamilton, 1988] Hamilton, J. D.: Rational-expectations econometric analysis of changes in regime : An investigation of the term structure of interest rates. Journal of Economic Dynamics and Control, 12(2-3):385–423, 1988.

[Hamilton, 1989a] Hamilton, J. D.: A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57, 357-384., 1989.

[Hamilton, 1989b] Hamilton, J. D.: A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57(2-3):357–384., 1989.

[Henry, 2009] Henry, l. T.: Regime switching in the relationship between equity returns and short-term interest rates in the UK. Journal of Banking & Finance, 33(2):405–414, February 2009.

[Hobbes et al. , 2007] Hobbes, G., Lam, F., and Loudon, G. F.: Regime shifts in the stockcbond relation in australia. Review of Pacific Basin Financial Markets and Policies (RPBFMP), 10(01):81–99, 2007.

[Kashyap, 1982] Kashyap, R. L.: Optimal choice of AR and MA parts in autoregressive moving average models. IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-4:99–104, 1982.

[Klaassen, 2002] Klaassen, F.: Improving garch volatility forecasts with regime-switching GARCH. Empirical Economics, 27, 363-394., 2002.

[Lunde and Timmermann, 2004] Lunde, A. and Timmermann, A.: Duration dependence in stock prices: an analysis of bull and bear markets. Journal of Business and Economic Statistics, 22:253–273, 2004.

[Maheu and McCurdy, 2000] Maheu, J. M. and McCurdy, T. H.: Identifying bull and bear markets in stock returns. Journal of Business and Economic Statistics, 18:100–112, 2000.

[Marcucci, 2005] Marcucci, J.: Forecasting stock market volatility with regime-switching garch models. Studies in Nonlinear Dynamics & Econometrics, 9(4):6, 2005.

[Nielsen and Olesen, 2001] Nielsen, S. and Olesen, J. O.: Regime-switching stock returns and mean reversion. Working Papers 11-2000, Copenhagen Business School, Department of Economics, July 2001.

[Pagan and Schwert., 1990] Pagan, A. R. and Schwert., G. W.: Alternative models for conditional stock volatility. Journal of Econometrics, 45, 267-290., 1990.

[Qiao et al. , 2008] Qiao, Z., Smyth, R., and Wong, W.: Volatility switching and regime interdependence between information technology stocks 1995-2005. Global Finance Journal, 19(2):139–156, 2008.

[Sclove, 1992] Sclove, S. L.: Mixture-model cluster analysis. CRIM Working Paper, 1992. Center for Research in Information Management, College of Business Administration, University of Illinois at Chicago.

[Sharpe, 1965] Sharpe, W. F.: Mutual fund performance. The Journal of Business, 39:119, 1965.

[Shyu and Hsia, 2008] Shyu, Y.-W. and Hsia, K.: Does stock market volatility with regime shifts signal the business cycle in Taiwan? International Journal of Electronic Finance, 2(4):433–450, January 2008.

[Susmel, 2000] Susmel, R.: Switching volatility in international equity markets. International Journal of Finance and Economics, 5, 265-283., 2000.

[Susmel and Kalimipalli, 2001] Susmel, R. and Kalimipalli, M.: Regime-switching stochastic volatility and short-term interest rates. CEMA Working Papers: Serie Documentos de Trabajo. 197, Universidad del CEMA, May 2001.

[Turner et al. , 1989] Turner, C. M., Startz, R., and Nelson., C. R.: A Markov model of heteroskedasticity, risk, and learning in the stock market. Journal of Financial Economics, 25, 3-22., 1989.

[Viterbi, 1967] Viterbi, A. J.: Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Transactions on Information Theory, IT-13:260–269, 1967.

[Zhang et al. , 2011] Zhang, C., Zhang, D., and Breece, J.: Financial crisis, monetary policy, and stock market volatility in china. Annals of Economics and Finance, 12(2):371–388, November 2011.