

Capital Asset Pricing Model and Changes in Volatility

André Oliveira SANTOS

Graduate Institute of International Studies, Geneva

Research Paper N° 4

INTERNATIONAL CENTER FOR FINANCIAL
ASSET MANAGEMENT AND ENGINEERING

Capital Asset Pricing Model and Changes in Volatility

André Oliveira Santos*

30 September 1998

*The author is a Ph.D. student at the Graduate Institute of International Studies, Geneva, Switzerland. The research is supported by a grant from the International Center for Financial Asset Management and Engineering (FAME), Geneva. The author would like to acknowledge Fred James and Eduardo do Couto e Silva, from Centre Europeen pour la Recherche Nuclaire, for their help in the implementation of the routines in FORTRAN.

Abstract

This article applies regime-switching models to assess the effects of different regimes of volatility in asset pricing. Different variance-covariance matrices for different regimes of volatility are introduced in the Capital Asset Pricing Model. They are scaled with respect to a conditional variance-covariance matrix that simply follows a GARCH process. The probabilities that U.S. financial markets were in a low, medium, or high regime of volatility from March 1958 to December 1995 are computed.

Capital Asset Pricing Model and Changes in Volatility

By André Oliveira Santos

Executive Summary

In the financial literature asset volatility is important since many of the asset pricing models require expected returns to be inversely proportional to the variance as in the case of the Capital Asset Pricing Model (CAPM) or prices of a derivative asset to be dependent on the variance of the underlying asset as in the case of options and futures. One approach to compute volatility is the Autoregressive Conditional Heteroskedasticity (ARCH) class of models. These models usually set the current variance-covariance matrix of asset returns equal to a projection on matrices of past squared error terms and on past variance-covariance matrices in a multivariate setting. This specification is suited to capture "the tendency for volatility clustering, i.e., for large (small) price changes to be followed by other large (small) price changes, but of unpredictable sign"¹.

Episodes of large and sudden shifts in volatility are rare. Examples are the Great Depression in the 1930s, the oil shock in the 1970s and the stock market crash of 1987. ARCH models are able to capture the sudden shifts in volatility but with lags due to the relevance of past information on projecting current volatility. The contribution of this paper is to extend ARCH models to capture the sudden changes in volatility with the use of regime switching models.

Regime-switching models allow the estimation of the probability that a certain observation comes from different probability distributions. A successful application has been, for instance, the evaluation of business cycles. Hamilton (1988) proposed a fourth order centered Autoregressive Moving Average (ARMA) process for changes in the U.S. GDP with a constant switching between negative and positive values. In periods of expansion, changes in GDP come from a probability distribution with a positive mean while, in periods of contraction, changes in GDP come from a probability distribution with a negative mean.

In the regime switching literature, volatility of stock returns, short term interest rates and spreads between long and short-term interest rates have been documented as being dramatically affected by the Federal Reserve Bank's change of monetary policy at the beginning of the 1980s. Researchers like Cai (1994), Hamilton and Susmel(1994), Gray (1996), Dueker (1996), and Hamilton and Lin (1996) incorporated these volatility changes by allowing some parameters in ARCH models to switch between regimes.

If financial markets are subject to sudden shifts in volatility and if asset pricing models are valid for all regimes of volatility, then economic agents price assets differently in different periods of volatility. Particularly, economic agents face different probability distributions for different states of volatility. To infer the existence of different regimes of volatility and the consequent effect on asset pricing, a regime-switching econometric model is proposed and implemented in this paper.

The Conditional Capital Asset Pricing Model (CAPM) with regime switches in volatility proposed in the paper is an extension of the work by Cai (1994), Hamilton and Susmel(1994),

¹See Bollerslev, Chou and Kroner (1992), p. 8.

Gray (1996), Dueker (1996), and Hamilton and Lin (1996). The variance-covariance matrix of asset returns is different for different states of volatility. In addition, the dynamics of volatility is also conditioned on past information, capturing the volatility clustering effect so often documented in the ARCH literature.

The combination of the ARCH and regime switching approaches in a multivariate setting introduces flexibility in the dynamics of volatility but at the cost of a large number of parameters which must be estimated. To decrease the number of parameters in the estimation, two simplifications are adopted in the paper. The first is to set all the off-diagonal elements in the parameter matrices of the BEKK representation for the dynamics equal to zero for all states. The second simplification is to choose one state and scale all of the parameter matrices for the other states with respect to that first state.

The changes in the scale of the variance-covariance matrices or in the regime of volatility are driven by a state variable that evolves according to a Markov process. Even though financial markets share simultaneous periods of high and low volatility, the degree of the response of the U.S. bond or stock markets to changes in volatility may be different. Different scaling factors for each market, driven by the same state variable, are allowed in the text.

If three regimes of volatility (low, medium, and high) are allowed, the estimates for the transition probabilities that drive the changes in volatility from state 1 at time t to state 1 at time $t+1$ and from state 2 at time t to state 2 at time $t+1$ are very high for U.S. financial markets during the period March, 1958 to December, 1995. However, the estimate of the probability of going from state 3 at time t to state 3 at time $t+1$ in the text is around 57.5%, which implies a probability of 42.5% for going from state 3 to state 2. This caused a reversion from the state of high volatility to the medium regime in U.S. financial markets.

The estimates for the scaling parameters within a three-regime framework implies that the conditional variance in state 2 is 25.2 times greater than the one in state 1 for bonds and 1.9 times higher for stocks. In state 3, the conditional variance in state 3 is 165.9 times greater than the one in state 1 for bonds and 7.5 times greater for stocks. These numbers suggest that bonds are much less volatile than stocks.

Once the estimates for the parameters in the model are obtained, the conditionally expected excess returns at time t based on a certain state of volatility and on past excess returns are computed. Since economic agents do not know exactly the current state of volatility, they infer the probability of being in certain state and weight the conditionally expected excess returns on that state by their respective probability, resulting in conditionally expected excess returns not based on any state but only on past excess returns.

From the filter used in the estimation, smoothed probabilities for the regimes 1, 2 and 3 were also obtained. Smoothed probabilities represent “the smoothed inference about the regime the

process was in at date t based on data obtained through some later date T .¹² The smoothed probabilities for the U.S. financial markets show that volatility stays in regime 2 during most of the period from March 1958 to December 1995. Periods of low volatility are the ones comprehended between December 1962 and August 1965, between November 1971 and November 1972, and between February 1977 and September 1977. Periods of high volatility include the sharp changes in market rates from June 1958 to July 1958, the oil shock and the Bankhaus Herstatt and Franklin National crises from September 1974 to October 1974, the changes in monetary policy by the Federal Reserve from October 1979 to April 1980, and the stock market crash in October 1987. Thus, periods of high volatility are brief, suggesting the existence of a mean reversion in volatility.

¹²Hamilton (1994), p. 694.

Capital Asset Pricing Model and Changes in Volatility

André Oliveira Santos*

30 September 1998

Abstract

This article applies regime-switching models to assess the effects of different regimes of volatility in asset pricing. Different variance-covariance matrices for different regimes of volatility are introduced in the Capital Asset Pricing Model. They are scaled with respect to a conditional variance-covariance matrix that simply follows a GARCH process. The probabilities that U.S. financial markets were in a low, medium, or high regime of volatility from March 1958 to December 1995 are computed.

1. Introduction

In the Finance literature, the estimation of asset volatility is important because many of the asset pricing models require expected returns to be inversely proportional to the variance - as in the case of the Capital Asset Pricing Model (CAPM) - or prices of a derivative asset to be dependent on the variance of the underlying asset - as in the case of options and futures. One approach to compute volatility is the Autoregressive Conditional Heteroskedasticity (ARCH) class of models, which projects the current variances of asset returns on past squared error terms and on past variances and covariances. ARCH specification is suited to capture

*The author is a Ph.D. student at the Graduate Institute of International Studies, Geneva, Switzerland. The research is supported by a grant from the International Center for Financial Asset Management and Engineering (FAME), Geneva. The author would like to acknowledge Fred James and Eduardo do Couto e Silva, from Centre Europeen pour la Recherche Nuclaire, for their help in the implementation of the routines in FORTRAN.

"the tendency for volatility clustering, i.e., for large (small) price changes to be followed by other large (small) price changes, but of unpredictable sign"¹.

Recently, financial markets around the world saw periods of small price changes interrupted by some sparks of large price changes. Yet such episodes of large and sudden price shifts are still more rare but include the Great Depression in the 1930s, the oil shocks in the 1970s and the stock market crash of 1987 among others. ARCH models are able to capture the sudden shifts in volatility but with some lags due to the relevance of past information on projecting current volatility. The contribution of this paper is to extend ARCH models to capture the sudden changes in volatility with the use of regime switching models.

Regime-switching models allow the estimation of the probability that a certain observation comes from different probability distributions. A successful application has been, for instance, the evaluation of business cycles. Hamilton (1988) proposed a fourth order centered Autoregressive Moving Average (ARMA) process for changes in the U.S. GDP with a constant switching between negative and positive values. In periods of expansion, changes in the GDP come from a probability distribution with a positive mean while, in periods of contraction, changes in GDP come from a probability distribution with a negative mean.

If financial markets are subject to sudden shifts in volatility and if asset pricing models are valid for all regimes of volatility, then economic agents price assets differently in different periods of volatility. Economic agents face different probability distributions for different states of volatility. To infer the existence of different regimes of volatility and the consequent effect on asset pricing, a model that combines both the ARCH and the regime switching methodologies is proposed and implemented in the paper.

The paper is organized as follows: subsection 2.1 briefly reviews the generalized autoregressive conditional heteroskedasticity in the mean (GARCH-M) literature in multivariate tests of the CAPM. The next subsection summarizes models in the regime-switching literature that allow the conditional variance-covariance matrix to change within regimes. In section 3, I propose a model that includes both volatility clustering and regime switches, but that is slightly different from those of the previous sections. Section 4 uses U.S. data and applies the methodology of the previous section to the CAPM. Finally, I briefly review the most important results in the article and suggest some directions for future research.

¹See Bollerslev, Chou and Kroner (1992), p. 8.

2. GARCH/ARCH-M and Regime-Switching Models

2.1. GARCH/ARCH-M in Asset Pricing Models

Before proceeding with the application of regime-switching models to the analysis of different regimes of volatility, I will briefly review the literature on conditional tests of the CAPM in a multivariate setting.

From the mean-variance analysis, the conditionally expected excess return on any asset is linearly proportional to its covariance with the market return:

$$E_{t-1}(R_{i,t} - R_{ot}1) = \delta cov_{t-1}(R_{i,t}, R_{m,t}) \quad (2.1)$$

where δ is the market price of risk (assumed to be constant in the whole article), E_{t-1} is the expectation operator based on information available at time $t-1$, $R_{i,t}$ is the gross return on a risky asset, R_{ot} is the risk-free interest rate, and $R_{m,t}$ is the gross return on the market portfolio. The market return is a weighted average of all returns on assets available in the economy:

$$R_{m,t} = \sum_{i=1}^N w_{i,t} R_{i,t} \quad (2.2)$$

where $w_{i,t}$ are the asset shares in the market portfolio.

Substitution of the previous definition of the market return in the asset pricing equation (2.1) above yields an expression where expected excess returns are proportional to a linear combination of variances and covariances of asset returns, with the weights given by the asset shares in the market portfolio. In matrix notation, this is equivalent to:

$$E_{t-1}(R_t - R_{ot}1) = \delta \Omega_t w_{t-1} \quad (2.3)$$

where R_t is an $(N \times 1)$ vector of gross returns on risky assets, 1 is an $(N \times 1)$ vector of ones, Ω_t is an $(N \times N)$ conditional variance-covariance matrix of excess returns, and w_{t-1} is a $(N \times 1)$ vector of asset shares.

Under rational expectations, current excess returns are equal to conditionally expected excess returns plus an error term:

$$R_t - R_{ot}1 = E_{t-1}(R_t - R_{ot}1) + \varepsilon_t \quad (2.4)$$

where ε_t is an $(N \times 1)$ vector of error terms defined as:

$$\varepsilon_t = R_t - E_{t-1}(R_t) \quad (2.5)$$

with a $N(0, H_t)$ distribution given the information set available to investors at time $t-1$.

Substitution of the equilibrium asset pricing equation (2.3) in expression (2.4) yields:

$$R_t - R_{ot}1 = \delta\Omega_t w_{t-1} + \varepsilon_t \quad (2.6)$$

or, for the ease of notation²:

$$r_t = \delta\Omega_t w_{t-1} + \varepsilon_t \quad (2.6.1)$$

where $r_t = R_t - R_{ot}1$ is the vector of excess returns.

Although the mean-variance analysis has a linear relationship between conditionally expected excess returns and their conditional variances and covariances as in expression (2.3), it does not predict how variances and covariances evolve through time. The Constrained Asset Share Estimation (CASE) Method of Engel, Frankel, Froot, and Rodrigues (1993) constrains the variance-covariance matrix of excess returns to be equal to the variance-covariance matrix of the error terms:

$$\Omega_t = E_{t-1}(\varepsilon_t \varepsilon_t') = H_t \quad (2.7)$$

In much of the empirical work in a multivariate setting, the latter matrix H_t is set as a linear function of a matrix of constants, of lagged squared error terms $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$, and of lagged variance-covariance matrices H_{t-1}, \dots, H_{t-q} ; that is, H_t follows a GARCH(p,q) process. Since the variance-covariance matrix H_t must be positive-definite, the conditions of the GARCH process that guarantee it are: (i) a positive definite matrix of constants and (ii) positive semi-definite matrices of coefficients that multiply the lagged squared error terms $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$ and the past variances and covariances H_{t-1}, \dots, H_{t-q} ³.

Among the many representations listed in Ding and Engle (1994), one that has attracted special interest is the BEKK representation:

$$H_t = C'C + A'_1 \varepsilon_{t-1} \varepsilon_{t-1}' A_1 + \dots + A'_p \varepsilon_{t-p} \varepsilon_{t-p}' A_p + B'_1 H_{t-1} B_1 + \dots + B'_q H_{t-q} B_q \quad (2.8)$$

where C is an $(N \times N)$ inferior triangular matrix of constants and $A_1, \dots, A_p, B_1, \dots, B_q$ are $(N \times N)$ matrices of coefficients. Its weak conditions for positive definiteness

²Equation(2.6) may be interpreted as the inverse of a demand equation where shares are a function of excess returns plus error terms. The CAPM restricts the way in which shares depend on their returns.

³See Attanasio (1991).

and its capacity to include other representations⁴ have made it one of the most used in a multivariate setting. Examples of recent applied work that test either a “closed economy” or an international version of the CAPM with a BEKK representation include Fornari (1995), Chan, Karolyi, and Stulz (1992), Engel (1994), De Santis and Gerard (1996), and De Santis and Gerard (1997).

Since the current conditional variance-covariance matrix H_t is a projection on the squared vectors of past error terms $\varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ and on the variance-covariance matrices H_{t-1}, \dots, H_{t-p} , GARCH/ARCH models are able to capture “the tendency for volatility clustering, i.e., for large (small) price changes to be followed by other large (small) price changes, but of unpredictable sign.”⁵ This is an important feature of many macroeconomic and financial time series. However, if one thinks of periods of volatility clustering as being interrupted by sudden shifts in the level of volatility, then the simple GARCH/ARCH framework is not flexible enough to capture the sudden change in regimes.

2.2. Regime Switching Models and Changes in Volatility

In the regime switching literature researchers have particularly focused attention on the behavior of excess returns on stocks, of short term interest rates, and of spreads between long and short-term interest rates. The variances of these variables seem to have been dramatically affected when the Federal Reserve Bank changed its monetary policy procedures at the beginning of the 1980s.

To allow some flexibility in capturing this change in regime, some researchers have proposed GARCH/ARCH models that have some parameters switching between regimes. For instance, Hamilton and Susmel (1994) introduced a scaling regime-switching ARCH model with returns on stocks following an AR(1) process⁶:

$$R_t = \alpha + \phi R_{t-1} + \varepsilon_t$$

where R_t is a vector of asset returns, ϕ is an $(N \times N)$ matrix of coefficients, and ε_t is an $(N \times 1)$ vector of error terms.

The latter vector of error terms ε_t was set equal to the product:

$$\varepsilon_t = \sqrt{g_{s_t}} u_t \tag{2.9}$$

⁴See Engle and Kroner (1995).

⁵See Bollerslev, Chou and Kroner (1992), p. 8.

⁶Although the articles reviewed in this section refer to a univariate setting, I adapt the models to a multivariate setting in order to keep consistency of notation.

where u_t is an $(N \times 1)$ vector of error terms with a $N(0, H_t)$ distribution and where g_{s_t} is a scaling factor. The latter scaling factor g_{s_t} is indexed by an unobservable state variable s_t , which assumes values from one to K and evolves according to a Markov process with a $(K \times K)$ transition probability matrix. This means that, when the state variable s_t is equal to one, the scaling factor g_{s_t} assumes a value equal to one; when s_t is different from one, g_{s_t} assumes other positive real numbers. With such a scaling variable g_{s_t} , changes in volatility from one state to another occur through changes in the scale of the process ε_t .

In Hamilton and Susmel (1994), the vector of error terms u_t was the usual one described by the product :

$$u_t = M_t v_t \quad (2.10)$$

where the term v_t is an $(N \times 1)$ vector of error terms with a $N(0, I)$ distribution and where the $(N \times N)$ matrix M_t stands for the Cholesky decomposition of the matrix H_t . The latter evolves according to an ARCH(p) process:

$$H_t = C' C + A'_1 u_{t-1} u'_{t-1} A_1 + \dots + A'_p u_{t-p} u'_{t-p} A_p \quad (2.11)$$

where use of the BEKK representation is made here.

The matrix H_t is made dependent on regimes by a substitution of the vector of error terms u_t in expression (2.11) by its definition in expression (2.9):

$$H_{t|s_{t-1}, \dots, s_{t-p}} = C' C + A'_1 (\varepsilon_{t-1} \varepsilon'_{t-1} / g_{s_{t-1}}) A_1 + \dots + A'_p (\varepsilon_{t-p} \varepsilon'_{t-p} / g_{s_{t-p}}) A_p \quad (2.12)$$

With this substitution, the conditional variance-covariance matrix of the error terms ε_t can be written as:

$$E_t(\varepsilon_t \varepsilon'_t | s_t, s_{t-1}, \dots, s_{t-p}) = g_{s_t} [C' C + A'_1 (\varepsilon_{t-1} \varepsilon'_{t-1} / g_{s_{t-1}}) A_1 + \dots + A'_p (\varepsilon_{t-p} \varepsilon'_{t-p} / g_{s_{t-p}}) A_p] \quad (2.13)$$

The error term ε_t is g_{s_t} times larger than the error terms u_t . This definition introduces a dependence of the conditional variance-covariance matrix $E_{t-1}(\varepsilon_t \varepsilon'_t | s_{t-1}, \dots, s_{t-p})$ on present and past states.

Cai (1994) also included regime switches in volatility in a ARCH(p) process. His model has an autoregressive process with a variance-covariance matrix $H_{t|s_t}$ equal to:

$$H_{t|s_t} = C' C_{s_t} + A'_1 \varepsilon_{t-1} \varepsilon'_{t-1} A_1 + \dots + A'_p \varepsilon_{t-p} \varepsilon'_{t-p} A_p \quad (2.14)$$

where the vector of error terms ε_t is also defined as in expression (2.10) and thus has a $N(0, H_{t|s_t})$ distribution. In the latter ARCH(p) process, the matrix of constants $C' C$ changes according to an unobservable state variable s_t .

Note that the researchers above dealt only with ARCH processes. On one hand, ARCH processes with changes in regime are simple and do not cause any problems in estimation, whereas GARCH processes are extremely cumbersome. The reason is that the current conditional variance-covariance matrix H_t depends not only on the history of the latest p states (as in Hamilton and Susmel (1994)), but on the whole history of states. For instance, after substitution of the matrix $H_{t-1|s_{t-2}}$ in a GARCH (1,1) process:

$$H_{t|s_{t-1}} = C'C + A'_1(\varepsilon_{t-1}\varepsilon'_{t-1}|g_{s_{t-1}})A_1 + B'_1H_{t-1|s_{t-2}}B_1 \quad (2.15)$$

I obtain the current conditional variance-covariance matrix as a function of the past states s_{t-1} and s_{t-2} . Further substitutions in the conditional variance-covariance matrices $H_{t-2|s_{t-3}}$, $H_{t-3|s_{t-4}}$, ..., introduce successive terms containing the scaling factor $g_{s_{t-3}}$, $g_{s_{t-4}}$, ..., g_{s_0} . Evaluation of the log-likelihood function at each point in time is performed with increasing difficulty due to the increasing number of possible histories of states.

Indeed, if coefficients in the mean equation assume different values according to current and past states of volatility, a similar recursion to the one in the ARCH-M of Engle, Lilien and Robins (1987) occurs through the error terms ε_t . A simple ARCH-M(1) in tests of the CAPM illustrates this point. Assume for the moment that the pricing equation (2.6.1) can be rewritten as:

$$r_t = \delta g_{s_t} H_t w_{t-1} + \varepsilon_{t|s_t} \quad (2.16)$$

when there exist regime shifts in volatility. If the pre-sample vector of error terms ε_0 is set equal to zero, then the variance-covariance matrix H_1 at time 1 is equal to $C'C$. Thus, the vector of error terms at time 1 can be calculated as $\varepsilon_{1|s_1} = r_1 - \delta g_{s_1} C'w_0$. Yet, at time 2 the variance-covariance matrix $H_{2|s_2}$ is:

$$H_{2|s_1} = C'C + A'_1\varepsilon_{1|s_1}\varepsilon'_{1|s_1}A_1 \quad (2.17)$$

In this case, the vector of error terms ε_2 at time 2 is:

$$\varepsilon_{2|s_2,s_1} = r_2 - \delta g_{s_2} H_{2|s_1} \quad (2.18)$$

which depends on the state variable s_t at times 1 and 2. Further substitution increases the dependence of error terms and of the conditional variance-covariance matrix on past states.

To overcome the recursions above and allow the use of GARCH with regime switches, Gray (1996) suggested the following procedure. Instead of past error

terms and variance-covariance matrices conditioned on past states, a GARCH(p,q) process would employ unconditioned (with respect to past states) vectors of error terms $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$ and variance-covariance matrices H_{t-1}, \dots, H_{t-q} . The current variance-covariance matrix $H_{t|s_t}$ would still depend on the current regime through the matrix of constants $C'C$ and the matrices of coefficients multiplying the past squared error terms and the past variance-covariances matrices. For instance, for a GARCH(1,1) process, Gray (1996) suggested defining the variance-covariance matrix $H_{t|s_t}$ as:

$$H_{t|s_t} = C'_{s_t} C_{s_t} + A'_{1|s_t} \varepsilon_{t-1} \varepsilon'_{t-1} A_{1|s_t} + B'_{1|s_t} H_{t-1} B_{1|s_t} \quad (2.19)$$

Clearly, the unconditioned vector ε_{t-1} and matrix H_{t-1} in the variance-covariance matrix $H_{t|s_t}$ break the dependence of the latter matrix on previous states s_{t-1}, s_{t-2}, \dots .

Dueker (1997) proposed a slightly different procedure. For a GARCH(1,1) process, the variance-covariance matrix $H_{t|s_t, s_{t-1}}$ depends on both current and lagged values of the state variable s_t . The current state variable s_t drives the constant matrix C while the lagged variable s_{t-1} drives the coefficient matrices A_1 and B_1 :

$$H_{t|s_t, s_{t-1}} = C'_{s_t} C_{s_t} + A'_{1|s_{t-1}} \varepsilon_{t-1} \varepsilon'_{t-1} A_{1|s_{t-1}} + B'_{1|s_{t-1}} H_{t-1} B_{1|s_{t-1}} \quad (2.20)$$

The unconditional vector of error terms ε_t in Gray (1996) is similar to the one in Dueker (1997), but the variance-covariance matrix in both papers are computed slightly differently. The unconditioned vector of error terms ε_t in Gray (1996) and in Dueker (1997) involves the multiplication the error term $\varepsilon_{t|s_t}$ in each state s_t by its correspondent regime probability $p(s_t|r_{t-1}, r_{t-2}, \dots; \theta)$ and then the sum of the product over all possible states s_t :

$$\varepsilon_t = \sum_{s_t=1}^K p(s_t|r_{t-1}, r_{t-2}, \dots; \theta) \cdot \varepsilon_{t|s_t} \quad (2.21)$$

where $p(s_t|r_{t-1}, r_{t-2}, \dots; \theta)$ stands for the regime probability; that is, the probability that a certain state occurs conditional to the information set at time $t-1$ ⁷.

In turn, Dueker (1997) computed the unconditional variance-covariance matrix H_t , first, by multiplying each conditional $H_{t|s_t}$ by its respective regime probability

⁷See next section for an example of its computation.

$p(s_t|r_{t-1}, r_{t-2}, \dots; \theta)$ and then, summing up all the products:

$$H_t = \sum_{s_t=1}^K p(s_t|r_{t-1}, r_{t-2}, \dots; \theta) \cdot H_{t|s_t} \quad (2.22)$$

while the unconditioned variance-covariance matrix H_t in Gray (1996) is simply found through the use of the definition of conditional variances:

$$H_t = \sum_{s_t=1}^K p(s_t|r_{t-1}, \dots; \theta) \varepsilon_{t|s_t} \varepsilon'_{t|s_t} - \left(\sum_{s_t=1}^K p(s_t|r_{t-1}, \dots; \theta) \varepsilon_{t|s_t} \right) \left(\sum_{s_t=1}^K p(s_t|r_{t-1}, \dots; \theta) \varepsilon_{t|s_t} \right)' \quad (2.23)$$

3. CAPM with Changes in Volatility

In this section I propose extensions to Hamilton and Susmel (1994), Gray (1996) and Dueker (1997) that allow different variance-covariance matrices in a multivariate GARCH-M specification for different states of volatility.

If the variance-covariance matrix of excess returns is different for each regime of volatility, then the equilibrium asset pricing equation (2.6.1) can be written as:

$$r_t = \delta \Omega_{t|s_t} w_{t-1} + \varepsilon_{t|s_t} \quad (3.1)$$

where the vector of error terms $\varepsilon_{t|s_t}$ has a $N(0, H_{t|s_t})$ distribution. As previously indicated, the use of the Constrained Asset Share Estimation (CASE) Method constrains the conditional variance-covariance matrix $\Omega_{t|s_t}$ of excess returns to be equal to the conditional variance-covariance matrix $H_{t|s_t}$ of the error terms $\varepsilon_{t|s_t}$:

$$\Omega_{t|s_t} = E_{t-1}(\varepsilon_{t|s_t} \varepsilon'_{t|s_t}) = H_{t|s_t} \quad (3.2)$$

With the latter constraint, the previous asset pricing expression (3.1) can be written as:

$$r_t = \delta H_{t|s_t} w_{t-1} + \varepsilon_{t|s_t} \quad (3.3)$$

In the variance-covariance matrix $H_{t|s_t}$ above, the state variable s_t drives the volatility in all asset markets; however, there could exist different state variables driving volatility in the different asset markets. An example would be different state variables driving the foreign currency exchange and stock markets in an

International CAPM. For the ease of presentation, I assume that the only state variable s_t that drives volatility in all asset markets evolves according to a Markov process with a ($K \times K$) transition probability matrix P ⁸:

$$P = \begin{bmatrix} p_{11} & p_{21} & \dots & p_{k1} \\ p_{12} & p_{22} & \dots & p_{k2} \\ & & \dots & \\ p_{1k} & p_{2k} & \dots & p_{kk} \end{bmatrix}$$

The conditions for stability of a Markov process are assumed to hold.

To complete the description of the model, the process for the variance-covariance matrix $H_{t|s_t}$ needs to be specified. In a BEKK representation where all the elements in the constant matrix C and in the coefficient matrices $A_1, \dots, A_q, B_1, \dots, B_p$ are subject to changes in regime (as in Gray (1996)), the variance-covariance matrix $H_{t|s_t}$ can be written as:

$$H_{t|s_t} = C'_{s_t} C_{s_t} + A'_{1|s_t} \varepsilon_{t-1} \varepsilon'_{t-1} A_{1|s_t} + \dots + A'_{p|s_t} \varepsilon_{t-p} \varepsilon'_{t-p} A_{p|s_t} + B'_{1|s_t} H_{t-1} B_{1|s_t} + \dots + B'_{q|s_t} H_{t-q} B_{q|s_t} \quad (3.4)$$

where both the error terms $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$ and the matrices H_{t-1}, \dots, H_{t-q} are unconditioned with respect to past state variables s_{t-1}, s_{t-2}, \dots . With a small number of assets and states, the number of parameters to be estimated in the variance-covariance matrix $H_{t|s_t}$ is large. For a GARCH(1,1) process with 2 assets and 2 states, the BEKK representation above has 22 parameters!

To decrease the number of parameters, I adopt two procedures. The first is to set all the off-diagonal elements in the coefficient matrices $A_{1|s_t}, \dots, A_{q|s_t}$ and $B_{1|s_t}, \dots, B_{p|s_t}$ equal to zero as in De Santis and Gerard (1996) and De Santis and Gerard (1997). The second procedure is to scale all the elements in the matrix $H_{t|s_t}$ with respect to a matrix H_t that does not switch with the state of volatility. For instance, if variances and covariances of excess returns change by the same proportion from one state to another, the variance-covariance matrix $H_{t|s_t}$ can be decomposed as:

$$H_{t|s_t} = g_{s_t} H_t \quad (3.5)$$

⁸Regarding the elements of the transition probability matrix, they may: (a) remain constant through all the sample size as assumed in the text; (b) be time-varying, that is, a function of observed economic fundamentals as in Filardo (1994); (c) be duration-depend, that is, not only a function of the inferred current state but also of the number of periods in which the state variable s_t has remained as in Durland and McCurdy (1994). The last two possibilities provide some flexibility to capture expected changes in the duration of phases of low and high volatility.

which scales the variance-covariance matrix $H_{t|s_t}$ with respect to matrix H_t . The latter matrix H_t simply follows a GARCH(p,q) process:

$$H_t = C'C + A'_1 \varepsilon_{t-1} \varepsilon'_{t-1} A_1 + \dots + A'_p \varepsilon_{t-p} \varepsilon'_{t-p} A_p + B'_1 H_{t-1} B_1 + \dots + B'_q H_{t-q} B_q \quad (3.6)$$

Note that all the past matrices H_{t-1}, \dots, H_{t-q} on the right hand side of the expression above are not conditioned on past states. This is not due to a pre-multiplication of past matrices $H_{t|s_{t-1}}, \dots, H_{t|s_{t-q}}$ by their respective regime probabilities as in Dueker (1997), but to a possible invertibility of the GARCH process. For a diagonal BEKK representation, the previous expression can be rewritten as:

$$\left(I - b_1 b'_1 \otimes L - \dots - b_q b'_q \otimes L^q \right) H_t = C'C + \left(a_1 a'_1 \otimes L + \dots + a_p a'_p \otimes L^p \right) \varepsilon_t \varepsilon'_t \quad (3.7)$$

where the a_i and b_j are vectors containing the diagonal elements in matrices A_i and B_j , respectively, and where L stands for lag operator. Invertibility means that the matrix H_t is an infinite sum of unconditioned past squared error terms:

$$H_t = C'^* C^* + (d_1 \otimes L + d_2 \otimes L^2 + \dots) \varepsilon_t \varepsilon'_t \quad (3.8)$$

where:

$$C'^* C^* = \left(I - b_1 b'_1 - \dots - b_q b'_q \right)^{-1} C'C \quad (3.9)$$

and where the term $(d_1 \otimes L + d_2 \otimes L^2 + \dots)$ is an (NxN) matrix representing the result of the product:

$$\left(I - b_1 b'_1 \otimes L - \dots - b_q b'_q \otimes L^q \right)^{-1} \left(a_1 a'_1 \otimes L + \dots + a_p a'_p \otimes L^p \right)$$

Expression (3.5) is equivalent to:

$$\varepsilon_{t|s_t} = \sqrt{g_{s_t}} M_t v_t \quad (3.10)$$

where v_t is an (Nx1) vector of error terms with a $N(0, I)$ distribution and where the matrix M_t stands for the Cholesky decomposition of the variance-covariance matrix H_t .

In this setting there exists only one scaling factor g_{s_t} for all the elements in the variance-covariance matrix H_t . However, different assets may have different scaling factors in different regimes. Even though financial markets share simultaneous periods of high and low volatility, the degree to which each market responds

may be different. This observation leads to a redefinition of the vector of error terms $\varepsilon_{t|s_t}$ as:

$$\varepsilon_{t|s_t} = G_{s_t}^{1/2} M_t v_t \quad (3.11)$$

where the (N×N) matrix $G_{s_t}^{1/2}$ of scaling factors contains the squared root of the scaling factors for each asset in its diagonal:

$$G_{s_t}^{1/2} = \begin{bmatrix} \sqrt{g_{1|s_t}} & 0 & \dots & 0 \\ 0 & \sqrt{g_{2|s_t}} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sqrt{g_{N|s_t}} \end{bmatrix} \quad (3.12)$$

With the latter redefinition, the conditional variance-covariance matrix $H_{t|s_t}$ of the error terms $\varepsilon_{t|s_t}$ becomes:

$$H_{t|s_t} = G_{s_t}^{1/2} H_t G_{s_t}^{1/2} \quad (3.13)$$

For the estimation of the parameters in the variance-covariance matrix $H_{t|s_t}$, of the elements in the transition probability matrix P, and of the price of risk δ (all comprised in the vector θ), I use the filter described in Hamilton (1994), chapter 22:

(i) given the past unconditional error terms $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$ and variance-covariance matrices H_{t-1}, \dots, H_{t-q} , I compute the conditional density function at time t, represented by a normal distribution:

$$f(r_t|s_t, r_{t-1}, r_{t-2}, \dots; \theta) = \frac{1}{\sqrt{2\pi|H_{t|s_t}|}} \exp[(r_t - \delta H_{t|s_t} w_{t-1})' (H_{t|s_t})^{-1} (r_t - \delta H_{t|s_t} w_{t-1})] \quad (3.14)$$

(ii) from the multiplication of the conditional density function of excess returns by the regime probability, I calculate the joint density function of excess returns and state s_t :

$$f(r_t, s_t|r_{t-1}, r_{t-2}, \dots; \theta) = f(r_t|s_t, r_{t-1}, r_{t-2}, \dots; \theta) \cdot p(s_t|r_{t-1}, r_{t-2}, \dots; \theta) \quad (3.15)$$

(iii) to find out the unconditional density distribution function of excess returns, I sum the joint density functions over all states s_t :

$$f(r_t|r_{t-1}, r_{t-2}, \dots; \theta) = \sum_{s_t=1}^K f(r_t, s_t|r_{t-1}, r_{t-2}, \dots; \theta) \quad (3.16)$$

(iv) to update the regime probabilities, I use the following Bayesian rule:

$$p(s_t|r_t, r_{t-1}, r_{t-2}, \dots; \theta) = \frac{f(r_t, s_t|r_{t-1}, r_{t-2}, \dots; \theta)}{f(r_t|r_{t-1}, r_{t-2}, \dots; \theta)} \quad (3.17)$$

(v) to obtain a forecast for the regime probabilities in period t+1 based on information available at time t, I multiply the updated probabilities by the transition probability:

$$p(s_{t+1}, s_t|r_t, r_{t-1}, r_{t-2}, \dots; \theta) = p(s_{t+1}|s_t).p(s_t|r_t, r_{t-1}, r_{t-2}, \dots; \theta) \quad (3.18)$$

and, then, sum over s_t :

$$p(s_{t+1}|r_t, r_{t-1}, r_{t-2}, \dots; \theta) = \sum_{s_t=1}^K p(s_{t+1}, s_t|r_t, r_{t-1}, r_{t-2}, \dots; \theta) \quad (3.19)$$

(vi) finally, I compute the unconditional error terms needed in the beginning of the next iteration:

$$\varepsilon_t = \sum_{s_t=1}^K p(s_t|r_{t-1}, r_{t-2}, \dots; \theta). \varepsilon_{t|s_t} \quad (3.20)$$

As a by-product of the use of the filter in steps (i) and (ii), the sample log-likelihood can be calculated:

$$L(\theta) = \sum_{t=1}^T \log(f(r_t|r_{t-1}, r_{t-2}, \dots; \theta)) \quad (3.21)$$

Once all the coefficients are estimated, the conditional expected excess returns at time t based on the state $s_t=k$ and on past excess returns can be easily computed as:

$$E(r_t|s_t = k, r_{t-1}, r_{t-2}, \dots; \theta) = \delta H_{t|s_t=k} w_{t-1} \quad (3.22)$$

In addition, conditionally expected excess returns based only on past excess returns are obtained by multiplying the expected excess returns conditioned on state s_t by the regime probability in state s_t and then summing over all possible states:

$$\begin{aligned} E(r_t|r_{t-1}, r_{t-2}, \dots; \theta) &= E(E(r_t|s_t, r_{t-1}, r_{t-2}, \dots; \theta)|r_{t-1}, r_{t-2}, \dots; \theta) \quad (3.23) \\ &= \sum_{s_t=1}^K p(s_t|r_{t-1}, r_{t-2}, \dots; \theta). \delta H_{t|s_t} w_{t-1} \end{aligned}$$

Economic agents are able to calculate precisely the variance-covariance matrix of excess returns on each state. However, they do not know exactly the current state of volatility. They infer the probability of being on certain state and weight the expected excess returns on that state by its respective probability. Therefore, conditionally expected excess returns are not based on any state but only on past excess returns.

Note that, to start the algorithm above, pre-sample values for the conditional variance-covariance matrix, for the error terms and for regime probabilities are necessary. A similar procedure as in De Santis and Gerard (1997) can be used for the pre-sample variance-covariance. In the first iteration each element in the sample variance-covariance matrix of excess returns is divided by the respective element in the matrix of scaling factors. In the other iterations the pre-sample variance-covariance matrix is set equal to sample variance-covariance matrix of the unconditional error terms of the previous iteration divided by the respective element in the matrix of scaling factors. Finally, the pre-sample values for the regime probabilities are simply set to equal to the ergodic probabilities, as suggested in Hamilton (1994).

4. Data Description and Interpretation of Empirical Results

The market portfolio studied in this section is composed of U.S. treasury bills, bonds and stocks. Treasury bills are the riskless asset. The data are monthly percentage excess returns and asset shares from March 1958 to December 1995, for a total of 454 observations. Monthly excess returns for all assets are taken from Ibbotson and Associates (1995) while assets shares are taken from publications of the Federal Reserve Bank. Asset shares are plotted in figure 4.1. For a detailed description of the data sources, definitions and transformations see the Appendix.

Table 4.1 provides summary statistics for the excess returns on bonds and stocks. In panel A, the computed statistics for the excess kurtosis for bonds and stocks are statistically different from zero at the 5% level. Since I use a "moment specification testing" approach (as in Cho and West (1995)) to estimate jointly the mean, the standard deviation, the skewness and the excess kurtosis coefficients, a test of the null hypothesis of normality of the excess returns simply corresponds to a Wald test of the skewness and of excess kurtosis coefficients being equal to zero. Since the Wald test statistics is greater than its critical value at a 5% level, I reject the null hypothesis of normality for the excess returns on bonds and

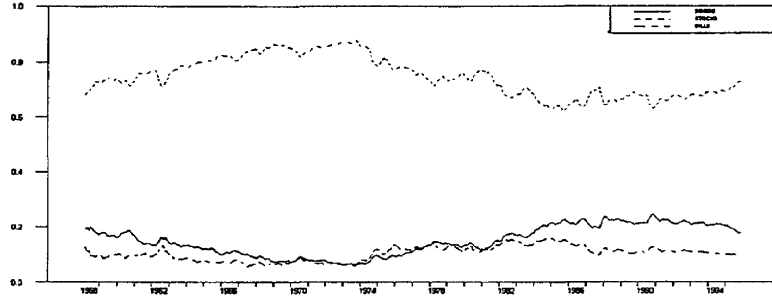


Figure 4.1: U.S. government bonds, stocks and treasury bills shares in the market portfolio, from March 1985 to December 1995.

stocks. I use a GARCH model with regime switches as an attempt to capture the non-normality of excess returns.

In panel A of table 4.1, I also compute a Modified Ljung-Box statistics adjusted to consider ARCH effects⁹. Diebold (1986) showed that the existence of ARCH effects may underestimate the standard errors of the autocorrelation coefficients and the Ljung-Box test statistics. Despite the adjustment for ARCH effects, the Ljung-Box test statistics show the existence of an autocorrelation of orders 1 and 3 in the excess returns for bonds¹⁰. Since parameter estimators are unbiased

⁹Cho and West (1995) suggested the modified Box-Ljung statistics below to account for ARCH effects:

$$N(N+2)\sigma_0^{\wedge 2} \sum_{j=1}^r (N-j)^{-1} (\rho_j^{\wedge 2} / k_j^{\wedge}) \sim \chi^2(r)$$

where $\sigma_0^{\wedge 2}$ is the unconditional variance, $\rho_j^{\wedge 2}$ is the calculated autocorrelation coefficient at lag j , and

$$k_j^{\wedge} = N^{-1} \sum_{t=j+1}^N (r_t - mean)^2 (r_{t-j} - mean)^2$$

¹⁰The cause of autocorrelation in the return on bonds may be, in part, due to the non-synchronous trading phenomenon discussed in Lo and MacKinlay (1988). Among the different maturities that compound our government bonds, one particular maturity is responsible for the autocorrelation: the U.S. Intermediate Government Bonds.

Table 4.1: Summary statistics for monthly excess returns on U.S. government bonds and stocks, from March 1958 to December 1995.

	Bonds	Stocks
Panel A: r_t		
Mean	0.0011 (0.0009)	0.0052 (0.0020)
Std.dev.	0.0175 (0.0014)	0.0413 (0.0025)
Skewness	0.5299 (0.4240)	-0.3508 (0.3404)
Excess kurtosis	4.8064 (2.3838)	2.4363 (1.2645)
Wald test for normality (H_0 : Skewness and excess kurtosis=0)	4.6067 [0.0999]	5.2052 [0.0741]
Modified L-B Q(1)	5.6032 [0.0179]	0.2844 [0.5938]
Modified L-B Q(2)	6.1072 [0.0472]	0.5049 [0.7769]
Modified L-B Q(6)	9.9549 [0.1265]	6.3847 [0.3815]
Panel B: r_t^2		
Mean	0.0003 (0.0000)	0.0017 (0.0002)
Std.dev.	0.0008 (0.0002)	0.0035 (0.0007)
L-B Q(1)	8.4334 [0.0037]	5.0634 [0.0244]
L-B Q(2)	32.5223 [0.0000]	7.2376 [0.0268]
L-B Q(6)	79.0749 [0.0000]	14.3975 [0.0255]

1) Standard errors computed according to Newey and West (1987), with an automatic lag selection as in Newey and West (1994), are shown in parenthesis.

2) Modified Ljung-Box statistics, adjusted for possible conditional heteroskedasticity effects, are computed according to Cho and West (1995).

3) p-values of the chi-squared statistics are shown in brackets.

but inefficient when there exists autocorrelation, I compute heteroskedasticity and autocorrelation-consistent standard errors - as in Newey and West (1987), with an automatic lag selection as in Newey and West (1994) - for the mean, the standard deviation and the skewness and excess kurtosis coefficients.

Panel B of table 4.1 reports summary statistics for the squared excess returns. The Ljung-Box test statistics is significantly different from zero even at high lags, which provides support for the use of a GARCH parametrization. This volatility clustering is exactly the reason why the GARCH framework was designed.

The estimation of the many statistical models in this section are undertaken with MINUIT¹¹. MINUIT is an optimization program widely employed by physicists and is specially suited to handling difficult problems such as these. The algorithm implemented in MINUIT is a stable variation of the DFP variable metric algorithm. However, MINUIT does not provide the scores, which are useful for economists in the computation of robust standard errors and in other specification tests. To overcome this drawback, I use a subroutine to calculate numerically the first derivative of the likelihood, at each point at time, with respect to each parameter once the maximum is obtained.

As a benchmark, I fit a GARCH(1,1) within a single regime. This captures only the time-varying nature of the second moments and corresponds to the statistical model:

$$r_t = \delta H_t w_{t-1} + \varepsilon_t \quad (4.1)$$

where

$$H_t = C'C + A_1' \varepsilon_{t-1} \varepsilon_{t-1}' A_1 + B_1' H_{t-1} B_1 \quad (4.2)$$

Table 4.2, panel A, first column, reports the Maximum Likelihood estimates and standard errors from the inverse of the Hessian matrix and from a procedure in Newey and West (1987), with an automatic lag selection as in Newey and West (1994). Except for the coefficient c_{21} in the constant matrix C , all the other coefficients in the mean equation (the relative risk aversion coefficient δ) and in the GARCH(1,1) process (the elements in matrices C , A_1 , and B_1) are statistically significant at a 5% level. Interestingly, the value for the risk aversion coefficient δ is equal to 4.75.

Table 4.3 provides summary statistics for the standardized residuals. Since the Modified Ljung-Box statistics for the residuals in the first equation are weakly different from zero at lags 1, 2, 3 and 4 (not shown) at the 5% level, heteroskedasticity

¹¹© Centre Européen pour la Recherche Nucleaire (CERN).

Table 4.2: Maximum likelihood estimates and specification tests of the relative risk aversion coefficient δ and of the elements in matrices A_1 , B_1 , and C in expressions (4.1), (4.2), (4.3) and (4.5), within a single, two and three regimes of volatility.

Panel A: Coefficients in Matrices A_1 , B_1 and C and δ within:			
	Single Regime	Two Regimes	Three Regimes
c_{11}	-0.0013 (0.0004) [0.0011]	0.0008 (0.0001) [0.0002]	0.0004 (0.0001) [0.0001]
c_{21}	0.0003 (0.0003) [0.0005]	-0.0002 (0.0001) [0.0001]	-0.0001 (0.0000) [0.0000]
c_{22}	0.0010 (0.0024) [0.0022]	-0.0067 (0.0015) [0.0013]	-0.0062 (0.0019) [0.0020]
a_{11}	0.4230 (0.0532) [0.1376]	0.0658 (0.0128) [0.0195]	0.0382 (0.0100) [0.0109]
a_{22}	0.2500 (0.0410) [0.0541]	0.1738 (0.0403) [0.0573]	0.1482 (0.0501) [0.0743]
b_{11}	0.9202 (0.0167) [0.0430]	-0.9273 (0.0216) [0.0339]	-0.9686 (0.0109) [0.0113]
b_{22}	0.9386 (0.0184) [0.0166]	-0.9399 (0.0162) [0.0135]	-0.9512 (0.0194) [0.0202]
δ	4.7587 (1.4584) [1.9346]	4.4417 (1.4693) [1.9462]	4.4253 (1.5263) [1.8834]
Loglikelihood	2506.102	2541.609	2560.827
Panel B: Wald Test Statistics of the Null Hypothesis $A_1=B_1=0$ within:			
	Single Regime	Two Regimes	Three Regimes
Hessian	31849	7186	19731
N.W.	11954	8085	17197

- 1) Standard errors from the Hessian matrix are shown in parenthesis.
- 2) Standard errors computed according to Newey and West (1987), with an automatic lag selection as in Newey and West (1994), are shown in brackets.

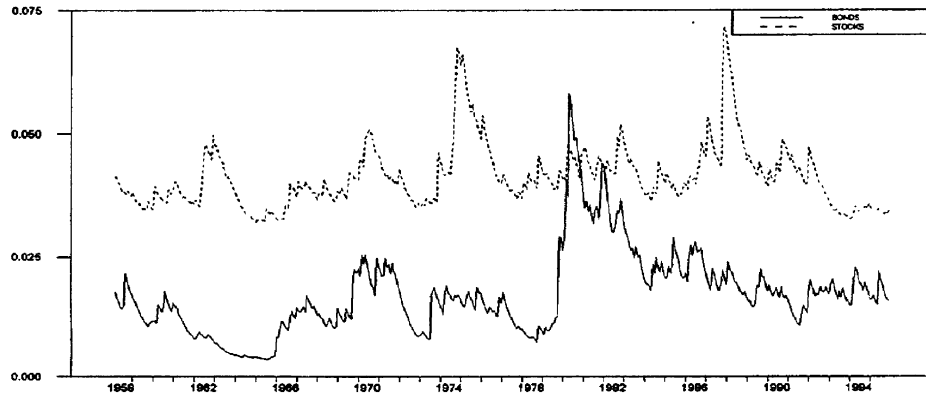


Figure 4.2: Conditional standard deviation for U.S. government bonds and stocks, computed according to expression (4.2), from March 1958 to December 1995.

and autocorrelation-consistent standard errors for the mean, the standard deviation and the skewness and excess kurtosis coefficients are also computed. The skewness and the excess kurtosis coefficients for the residuals in the first equation are statistically different from zero at the 5% level. A weakly significant Wald test statistics for the residuals implies that they may not be normally distributed. Since weak evidence of heteroskedasticity and autocorrelation was detected in the standardized residuals within a single regime, I also report heteroskedasticity and autocorrelation-consistent standard errors for all the estimates in tables 4.2 and 4.4 within a single, two and three regimes of volatility.

Conditional standard deviations and the risk premia, from March 1958 to December 1995, for bonds and stocks are plotted in figures 4.2 and 4.3, respectively. With the estimates of the parameters and pre-sample values for the error terms ε_0 and for the variance-covariance matrix H_0 , series for the conditional standard deviation and the risk premia are calculated recursively. Note that both figures also show that the simple regime GARCH(1,1) model is able to capture the oil shock of 1974, the changes in the monetary policy by the Federal Reserve during the period 1979 to 1982, and the stock market crash of October 1987.

In addition to the time-varying second moments, the inclusion of regime shifts

Table 4.3: Summary statistics for the standardized residuals in expressions (4.1), from March 1958 to December (1995).

	Bonds	Stocks
Panel A: ε_t		
Mean	-0.0251 (0.0529)	-0.0348 (0.0495)
Std.dev.	0.9850 (0.0526)	0.9876 (0.0463)
Skewness	-0.5714 (0.2857)	-0.4835 (0.2889)
Excess kurtosis	2.9465 (1.0051)	1.9091 (1.2051)
Wald test for normality (H_0 : Skewness and excess kurtosis=0)	10.1760 [0.0061]	2.8366 [0.2421]
Modified L-B Q(1)	7.5747 [0.0226]	0.3576 [0.5499]
Modified L-B Q(2)	7.5891 [0.0553]	0.4089 [0.8151]
Modified L-B Q(6)	10.2780 [0.1143]	7.5453 [0.2733]
Panel B: ε_t^2		
Mean	0.9704 (0.1045)	0.9766 (0.0930)
Std.dev.	2.1700 (0.3999)	1.9464 (0.4133)
L-B Q(1)	0.0654 [0.7982]	0.1322 [0.7161]
L-B Q(2)	0.0796 [0.9610]	0.1652 [0.9210]
L-B Q(6)	4.7844 [0.5712]	0.6620 [0.9953]

1) Standard errors computed according to Newey and West (1987), with an automatic lag selection as in Newey and West (1994), are shown in parenthesis.

2) Modified Ljung-Box statistics adjusted for possible conditional heteroskedasticity effects and computed according to Cho and West (1995).

3) p-values of the chi-squared statistics are shown in brackets.

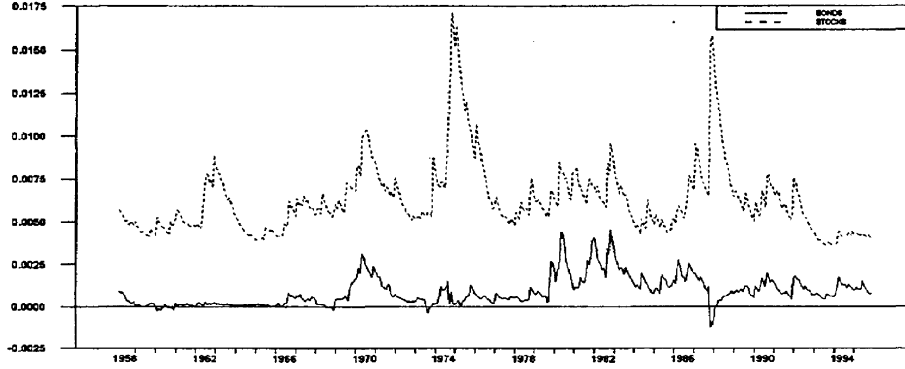


Figure 4.3: Risk premium for U.S. government bonds and stocks, computed according to expression (4.1), from March 1958 to December 1995.

in volatility in the Conditional Capital Asset Pricing Model yields the following statistical model:

$$r_t = \delta H_{t|s_t} w_{t-1} + \varepsilon_{t|s_t} \quad (4.3)$$

where:

$$H_{t|s_t} = G_{s_t}^{1/2} H_t G_{s_t}^{1/2} \quad (4.4)$$

and:

$$H_t = C'C + A_1' \varepsilon_{t-1} \varepsilon_{t-1}' A_1 + B_1' H_{t-1} B_1 \quad (4.5)$$

and:

$$G_{s_t}^{1/2} = \begin{bmatrix} \sqrt{g_{1|s_t}} & 0 \\ 0 & \sqrt{g_{2|s_t}} \end{bmatrix} \quad (4.6)$$

Indexing the scaling factors g_{s_t} , the state variable s_t is driven by the transition probability matrix P :

$$P = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \quad (4.7)$$

within a two regime framework, and by:

$$P = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} \quad (4.8)$$

within a three regime framework.

Table 4.2, panel A, also shows the estimates for the risk aversion coefficient and the coefficients in matrices A_1 and B_1 that do not vary despite the shifts in the regime of volatility. These coefficients in the GARCH(1,1) process characterize the time-varying nature of the second moments within regimes. When there exist two and three regimes of volatility, the risk aversion coefficient δ and all the diagonal elements in the matrices A_1 and B_1 are significantly different from zero. Interestingly, the diagonal elements a_{11} and a_{22} in matrix A_1 within two and three regimes are much smaller than within a single regime. This evidently decreases the impact of past shocks ε_{t-1} on the conditional variances and covariances. In turn, the persistence effect measured by the diagonal coefficients b_{11} and b_{22} is still high. The Wald test statistics in panel A for the null hypothesis that $a_{11}=a_{22}=b_{11}=b_{22}=0$ is statistically different from zero at the 5% level.

The estimates for the coefficients a_{11} , a_{22} , b_{11} and b_{22} also imply that the current conditional variance-covariance matrix is covariance stationary. The matrix H_t represents an infinite sum of past squared error terms and it is not conditioned on any present or past state. The possibility that the inclusion of regime switches in volatility implies covariance stationary in the GARCH process will be left for future research.

Characterizing the shifts in the level of volatility, the scaling parameters in matrix $G^{1/2}$ are shown in table 4.4, panel A. Within two regimes of volatility, the scaling parameters for bonds and stocks in state 2 are 24.4 and 2.2, respectively, and they are statistically significant. This means that the conditional variance in state 2 is 24.4 times higher than the one in state 1 for bonds and 2.2 times higher for stocks. Within a three-regime framework, the conditional variance in state 2 is 25.2 times higher than the one in state 1 for bonds and 1.93 times higher for stocks. In state 3, the conditional volatility for stocks is 165.86 and 7.54 times greater than the one in state 1 for bonds and stocks respectively. These numbers and figure 4.4 - which shows conditional standard deviations for bonds and stocks within a three-regime framework - suggest that bonds are much less volatile than stocks.

I next test the hypothesis that volatility could have remained constant within regimes and only changed with shifts in the regime of volatility. This hypothesis is equivalent to set the diagonal elements in matrices A_1 and B_1 in expression (4.5) equal to zero and involves 4 restrictions. Table 4.2, panel B, also presents the Wald statistics for this null hypothesis. I reject a constant variance-covariance matrix $H_{t|s_t}$ within regimes. This rejection favors the interpretation that volatility is not

Table 4.4: Maximum likelihood estimates of the coefficients in the scaling factor matrix G and in the probability matrix P in expressions (4.6), (4.7) and (4.8), within two and three regimes of volatility.

Panel A: Scaling Factor Matrix G within:					
	Two Regimes		Three Regimes		
	$s_t=1$	$s_t=2$	$s_t=1$	$s_t=2$	$s_t=3$
$g_{1 s_t}$	1	24.2170	1	25.1760	168.2700
	-	(6.5894)	-	(7.2997)	(91.7990)
	-	[9.9873]	-	[8.7506]	[84.2469]
$g_{2 s_t}$	1	2.2302	1	1.9919	7.5091
	-	(0.5854)	-	(0.6581)	(3.4443)
	-	[0.7450]	-	[0.9286]	[4.8070]

Panel B: Transition Probability Matrix P within:		
	Two Regimes	Three Regimes
	P_{11}	0.9335 (0.0404) [0.0385]
P_{12}	0.0665 (0.0404) [0.0385]	0.0697 (0.0395) [0.0369]
P_{21}	0.0107 (0.0068) [0.0083]	0.0090 (0.0058) [0.0059]
P_{22}	0.9893 (0.0068) [0.0083]	0.9667 (0.0185) [0.0214]
P_{23}	- - -	0.0243 (0.0243) [0.0273]
P_{32}	- - -	0.4180 (0.2082) [0.2806]
P_{33}	- - -	0.5820 (0.2082) [0.2806]

- 1) Standard errors from the Hessian matrix are shown in parenthesis.
- 2) Standard errors computed according to Newey and West (1987), with an automatic lag selection as in Newey and West (1994), are shown in brackets.

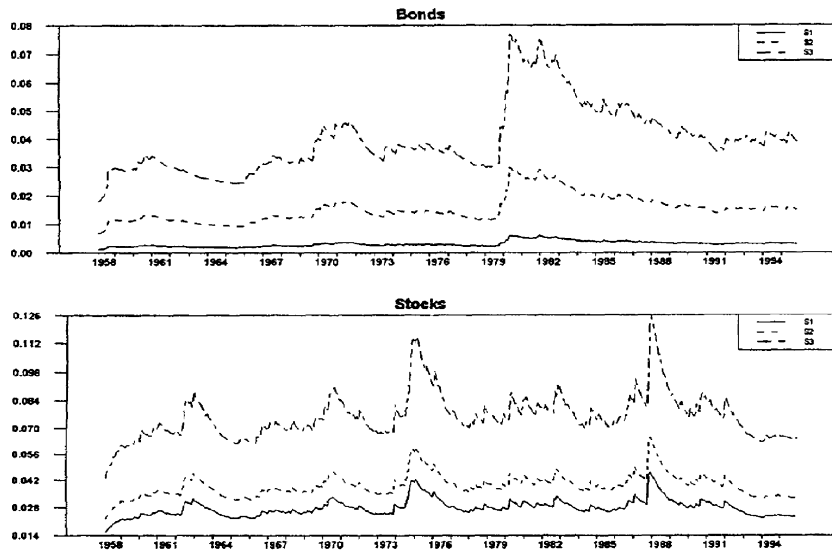


Figure 4.4: Conditional standard deviations for U.S. government bonds and stocks within three regimes of volatility, computed according to expression (4.4), (4.5) and (4.6), from March 1958 to December 1995.

a static concept in the sense that financial markets may be in a high volatility state today, but the same volatility level may be considered low in another period of time. “High” and “low” are, after all, relative terms.

The distinction of different states of volatility evidently implies more flexibility in capturing sudden changes in volatility than allowed by the single regime framework. Driving the changes in volatility, the estimates of the elements in the transition probability matrix P within two and three states are shown in table 4.4, panel B. Initially, no constraints were imposed within a three regime framework apart from the conditions¹² that $0 \leq p_{ij} \leq 1$ and $\sum_{j=1}^3 p_{ij} = 1$. However, many of the Maximum Likelihood estimations had the elements p_{13} and p_{31} on the boundary of zero. For purposes of finding the standard errors, these elements were set equal to zero and treated as constants. The intuition behind these two probabilities being set equal to zero is that markets do not jump from the lowest state of volatility to the highest and vice-versa.

A comparison between the values obtained for the elements p_{11} and p_{22} within two and three regimes of volatility shows that the probability of going from state 1 at time t to state 1 at time $t+1$ is very high in both frameworks, and the same remarks also apply to the probability of going from state 2 at time t to state 2 at time $t+1$. The probability of going from state 3 at time t to state 3 at time $t+1$ within a three regime framework is around 57.46%, and this implies a probability of 42.54% of going from state 3 to state 2. This is enough to cause a reversion from the state of high volatility to the medium regime.

Figures 4.5 plot the smoothed probabilities for the regimes 1, 2 and 3, respectively, within a three regime framework. Smoothed probabilities represent “the smoothed inference about the regime the process was in at date t based

¹² As in Hamilton and Susmel (1994), these conditions were guaranteed through a reparameterization of the log-likelihood function. The elements in the transition probability matrix were redefined as:

$$\begin{aligned} p_{11} &= \omega_{11}^2 / (1 + \omega_{11}^2); \\ p_{12} &= 1 / (1 + \omega_{11}^2); \\ p_{21} &= \omega_{21}^2 / (1 + \omega_{21}^2 + \omega_{22}^2); \\ p_{22} &= \omega_{22}^2 / (1 + \omega_{21}^2 + \omega_{22}^2); \\ p_{23} &= 1 / (1 + \omega_{21}^2 + \omega_{22}^2); \\ p_{32} &= \omega_{32}^2 / (1 + \omega_{32}^2); \\ \text{and } p_{33} &= 1 / (1 + \omega_{32}^2). \end{aligned}$$

Once the estimates for ω_{ij} are obtained, they can be substituted in the above expressions to compute the probabilities p_{ij} . The latter probabilities are then used in the routines to find the standard errors.

on data obtained through some later date T .¹³ Given the last regime probabilities $p(s_T|r_T, r_{T-1}, \dots; \theta)$, smoothed probabilities are obtained through the following algorithm:

$$P(s_t|r_T, \dots; \theta) = P(s_t|r_t, \dots; \theta) \odot \{P' \cdot \{P(s_{t+1}|r_T, \dots; \theta) (\div) P(s_{t+1}|r_t, \dots; \theta)\}\}$$

where $P(\cdot)$ may stand for an (3×1) vector of regime or smoothed probabilities and where the symbols \odot and (\div) stand for the multiplication and division of element by element of the vectors, respectively.

A simple examination of figure 4.5 shows that volatility stays in regime 2 during most of the period from March 1958 to December 1995. Periods of low volatility are the ones observed between December 1962 and August 1965, between November 1971 and November 1972, and between February 1977 and September 1977. Periods of high volatility include the sharp changes in market rates from June 1958 to July 1958¹⁴, the oil shock and the Bankhaus Herstatt and Franklin National crises from September 1974 to October 1974, the changes in monetary policy by the Federal Reserve from October 1979 to April 1980, and the stock market crash in October 1987.

According to figure 4.5, the periods of high volatility are brief, suggesting the existence of a mean reversion in volatility. However, Hamilton and Susmel (1994) obtained different results. In their work, the periods of high volatility in the stock market in a three regime framework last longer than the ones indicated here. They suggested that the periods of high volatility are associated with downturns in business cycles.

Figure 4.6 plots expected excess returns for bonds and stocks in all states within a three-regime framework. Expected excess returns for bonds and stocks are greater than zero from March 1958 to 1995 in all states of volatility. Expected excess returns in the highest state of volatility are greater than in the others, as one would expect from an inspection of the conditional standard deviations in figure 4.5.

Expected excess returns computed within one (based only on past information) and three regimes of volatility are plotted in figure 4.7. Expected excess returns within three regimes are higher than within one regime in the episodes of high

¹³Hamilton (1994), p. 694.

¹⁴"The sharp turnaround in market rates in July 1958 followed a large Treasury financing in mid-June and gave rise to outcries against speculation in government securities. The result was an extensive investigation by the Federal Reserve System and the Treasury". (Friedman and Schwartz (1963), p. 618)

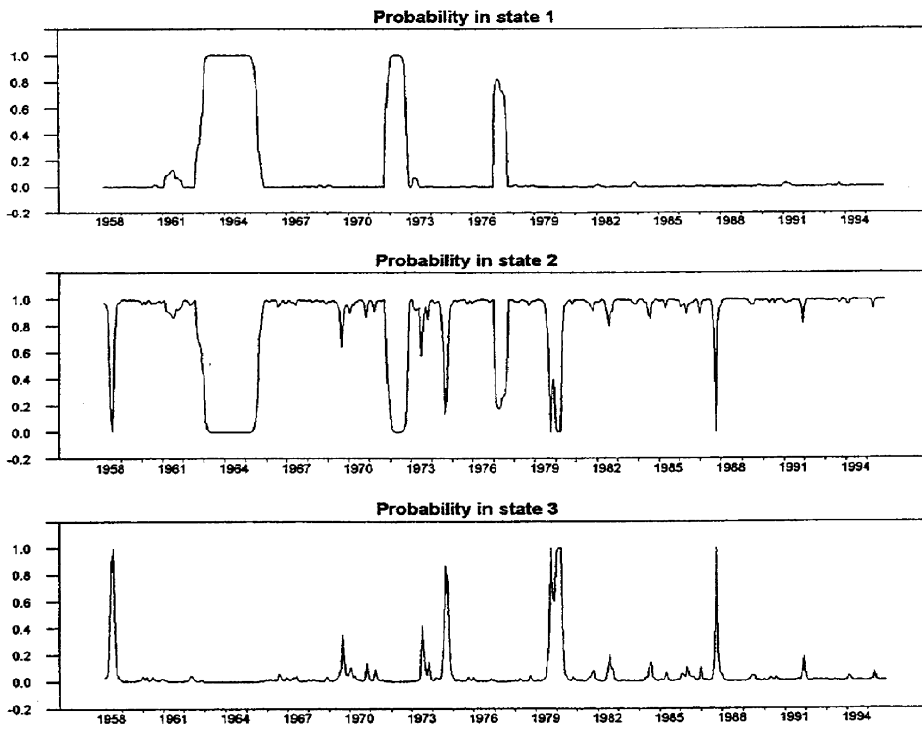


Figure 4.5: Smoothed probabilities that U.S. financial markets were in a state of low ($s_t=1$), medium ($s_t=2$) or high ($s_t=3$) volatility, from March 1958 to December 1995.

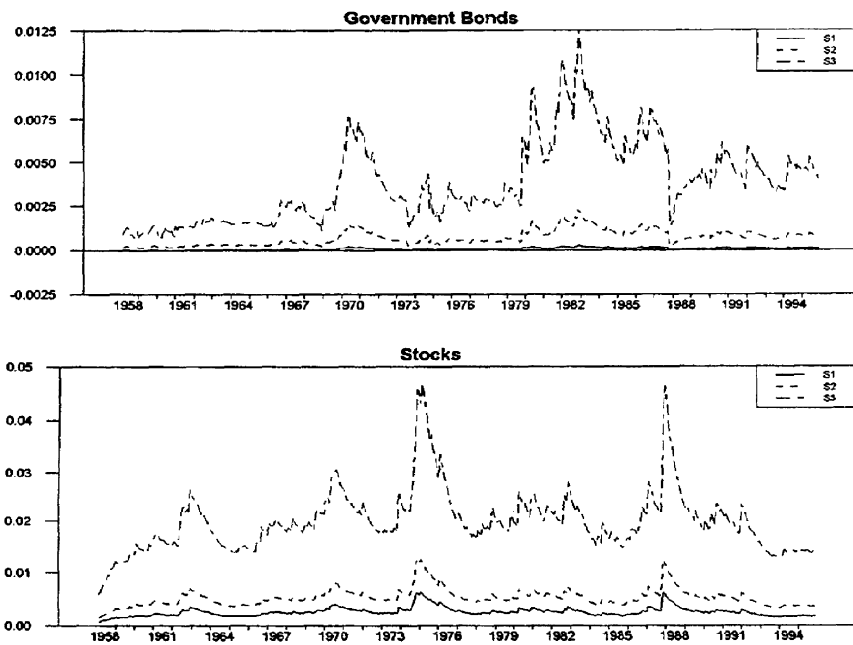


Figure 4.6: Expected excess returns for U.S. government bonds and stocks in all states, computed according to expression 4.3, within a three-regime volatility framework, from March 1958 to December 1995.

volatility. Indeed, expected excess returns are never negative within three regimes for both stocks and bonds.

5. Conclusion

In this paper I extended a multivariate GARCH-M model to the case where volatility is also subject to changes in regime. Essentially, I introduced different conditional variance-covariance matrices for different states of volatility in the Capital Asset Pricing Model (CAPM). The different conditional variance-covariance matrices were scaled with respect to a conditional variance-covariance matrix only following a GARCH(1,1) process.

Corresponding to the different conditional variance-covariance matrices, different expected excess returns were also computed. I observed that expected excess returns for bonds and stocks within a three-regime framework are higher than the ones computed within a single regime during the episodes of high volatility.

Moreover, I computed the probability that U.S. financial markets were in a low, medium, or high regime of volatility from, March 1958 to December 1995. The model implied a high probability of the U.S. financial markets being in high volatility state during the sharp changes in market rates in 1958; during the oil shock of the beginning of the 1970s; during the change in the monetary policy by the Federal Reserve in the late 1970s and early 1980s; and during the stock market crash in 1987. The periods of high volatility are brief, suggesting the existence of a mean reversion in volatility.

An important policy application concerns predictions of financial and exchange rate crises. If asset returns are really characterized by different probability distributions and financial and exchange rate crises correspond to periods of high volatility, then the econometric model in this paper could be used to compute the probability of such crises.

Appendix

The data used in this study includes the period from March 1958 to December 1996 and is similar to the one in Bollerslev, Engle, and Wooldridge (1988). Basically, the market portfolio is composed of U.S. Treasury Bills, U.S. Government Bonds and U.S. corporate equities. All excess returns over the riskless U.S. Treasury Bills are available in Ibbotson and Associates (1995). The asset shares are calculated from the amount of interest-bearing public debt held by private investors published in the Federal Reserve Bulletin/Treasury Bulletin, and from

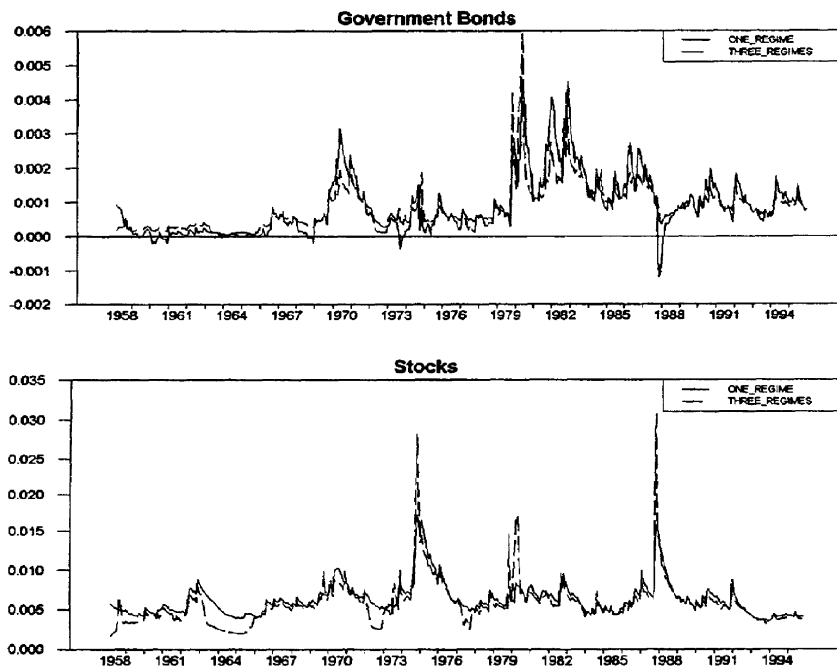


Figure 4.7: Expected excess returns for U.S. government bonds and stocks within one and three regimes of volatility, from March 1958 to December 1995.

the total market value of corporate equities available on-line from the Board of Governors of the Federal Reserve System.

Ibbotson and Associates (1995) publish monthly excess returns of Long and Intermediate-Term Government Bonds and of Large Company Stocks over Treasury Bills. However, in the text, I use aggregate excess returns for Government Bonds. The latter is calculated as follows: the amount of one to five-year and five to ten-year marketable interest-bearing debt (lagged one period) is added to obtain the outstanding amount of Intermediate-Term Government Bonds. The outstanding amount of maturities over the 10-year period is also added to calculate the outstanding amount of Long-Term Government Bonds.

Given the calculated outstanding amount of Long and Intermediate-Term Government Bonds, their excess returns can be weighted in order to calculate the excess returns of Government Bonds over U.S. Treasury Bills. The outstanding amount of Government Bonds is essentially the sum of the outstanding amount of Long and Intermediate-Term Government Bonds. The market values of the latter are obtained from a multiplication of their par values by a price index for marketable treasury debt described in Cox (1985) and kindly provided by the Federal Reserve Bank of Kansas.

Since the Flows of Funds tables contains only quarterly data, the monthly market value of corporate equities is calculated through the use of the command "Interpolate" in MATHEMATICA¹⁵. The behavior of the quarterly data for the total market value of equities is very smooth.

Finally, from the monthly outstanding amount of Government Bonds and Corporate Equities, asset shares can be calculated.

References

- [1] Attanasio, O. (1991). Risk, Time-Varying Second Moments and Market Efficiency. *Review of Economic Studies* 58, pp. 479-494.
- [2] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, pp. 307-327.
- [3] Bollerslev, T., Chou, R., and Kroner, K. (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics* 52, pp. 5-59.

¹⁵© Wolfram Research Inc.

- [4] Bollerslev, T., Engle, R., and Wooldridge, J. (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy* 96(11), pp. 116-131.
- [5] Cai, J. (1994). A Markov Model of Switching-Regime ARCH. *Journal of Business and Economic Statistics* 12(3), pp. 309-316.
- [6] Chan, K., Karolyi, G.A., and Stulz, R. (1992). Global Financial Markets and the Risk Premium on U.S. Equity. *Journal of Financial Economics* 32, 137-167.
- [7] Cho, K. and West, K. (1995), The Predictive Ability of Several Models of Exchange Rate Volatility. *Journal of Econometrics* 69, pp. 367-391.
- [8] Cox, W. (1985). The Behavior of Treasury Securities: Monthly, 1942-1984. *Journal of Monetary Economics* 16(2), pp. 227-40.
- [9] De Santis, G. and Gerard, B. (1996). How Big is the Premium for Currency Risk. University of Southern California. Manuscript.
- [10] De Santis, G. and Gerard, B. (1997). International Asset Pricing and Portfolio Diversification with Time-Varying Risk. *Journal of Finance* 52(5), pp. 1881-1912.
- [11] den Haan, W. and Levin, A. (1997), A Practitioner's Guide to Robust Covariance Matrix Estimation. in Maddala, G. and Rao, C. (eds.). *Handbook of Statistics Volume XV*, pp. 299-342.
- [12] Diebold, F. (1986). Testing for Serial Correlation in the Presence of ARCH. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 323-328.
- [13] Diebold, F. and Lopez, J. (1995). Modelling Volatility Dynamics. in Hoover, K. (ed.), *Macroeconometrics: Developments, Tensions, and Prospects*. Norwell: Kluwer Academic Publishers, pp. 437-466.
- [14] Ding, Z. and Engle, R. (1994). Large Scale Conditional Covariance Matrix Modeling, Estimation, and Testing. Manuscript
- [15] Dueker, M. (1997). Markov Switching in GARCH Processes and Mean-Reverting Stock Market Volatility. *Journal of Business and Economic Statistics* 15(1), pp. 26-34.

- [16] Durland, J. and McCurdy, T. (1994). Duration-Dependent Transitions in a Markov Model of US. GNP Growth. *Journal of business and Economics Statistics* 12(3), pp. 279-288.
- [17] Engel, C. (1994). Tests of CAPM on an International Portfolio of Bonds and Stocks. Tests of CAPM on an International Portfolio of Bonds and Stocks. in Frankel, J. (ed.). *The internationalization of Equity Markets*. Chicago: The University of Chicago Press, pp. 149-172.
- [18] Engel, C., Frankel, J., Froot, K., and Rodrigues, A. (1993). The Constrained Asset Share Estimation (CASE) Method: Testing Mean-Variance Efficiency of the U.S. Stock Market. *NBER Working Paper*, no. 4294.
- [19] Engle, R. and Kroner, K. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory* 11, pp. 122-150.
- [20] Engle, R., Lilien, D., and Robins, R. (1987). Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica* 55, pp. 391-407.
- [21] Filardo, A. (1994). Business-Cycle Phases and Their Transitional Dynamics. *Journal of Business and Economic Statistics* 12(3), pp. 299-308.
- [22] Fornari, F. (1995). The Conditional Capital Asset Pricing Model: An Application to the Italian Financial Market. *Journal of Foreign Exchange and International Finance* 8(4), pp. 418-436.
- [23] Friedman, M. and Schwartz, A. (1963). *A Monetary History of the United States: 1867-1960*. Princeton: Princeton University Press.
- [24] Gray, S. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process. *Journal of Financial Economics* 42, pp. 27-62.
- [25] Hamilton, J. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 11, 357-384.
- [26] Hamilton, J. (1994). *Time Series Analysis*. New Jersey: Princeton University Press.
- [27] Hamilton, J. and Lin, G. (1996). Stock Market Volatility and the Business Cycle. *Journal of Applied Econometrics* 11, pp. 573-593.

- [28] Hamilton, J. and Susmel, R. (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics* 64, pp. 307-333.
- [29] Ibbotson and Associates (1995). *Stocks, Bonds, Bills, and Inflation 1995 Yearbook*. Chicago: Ibbotson and Associates.
- [30] Lo, A. and MacKinlay, A. (1988), Stocks Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies* 1, pp. 41-66.
- [31] Newey, W. and West, K. (1987). A Simple, Positive Semidefinite, Heteroskedasticity, and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, pp. 703-708.
- [32] Newey, W. and West, K. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Econometric Studies* 61, pp. 631-654.

RESEARCH PAPER SERIES

Extra copies of all research papers will be available to the public upon request. In order to obtain copies of past or future works, please contact our office at the location below. As well, please note that all works will be available on our website, www.fame.ch, under the heading "Research" in PDF format for your consultation. We thank you for your continuing support and interest in FAME and look forward to serving you in the future.

N° 1: Enhancing portfolio performance using option strategies: why beating the market is easy.
François-Serge LHABITANT, HEC-University of Lausanne

N° 2: Application of simple technical trading rules to Swiss stock prices: Is it profitable ?
Dušan ISAKOV, HEC-Université de Genève
Marc HOLLISTEIN, Banque Cantonale de Genève

N° 3: Real Options as a Tool in the Decision to Relocate: An Application to the Banking Industry
Pascal BOTTERON, HEC-University of Lausanne



INTERNATIONAL CENTER FOR
FINANCIAL ASSET MANAGEMENT AND ENGINEERING

102, Bd. Carl-Vogt
PO Box 5543
1211 Geneva 11, Switzerland
Tel (+4122) 312 09 61
Fax (+4122) 312 10 26
<http://www.fame.ch>
E-mail: admin@fame.ch