A Capital Asset Pricing Model with Time-varying Covariances

Tim Bollerslev
Northwestern University

Robert F. Engle
University of California, San Diego

Jeffrey M. Wooldridge
Massachusetts Institute of Technology

The capital asset pricing model provides a theoretical structure for the pricing of assets with uncertain returns. The premium to induce risk-averse investors to bear risk is proportional to the nondiversifiable risk, which is measured by the covariance of the asset return with the market portfolio return. In this paper a multivariate generalized autoregressive conditional heteroscedastic process is estimated for returns to bills, bonds, and stocks where the expected return is proportional to the conditional covariance of each return with that of a fully diversified or market portfolio. It is found that the conditional covariances are quite variable over time and are a significant determinant of the time-varying risk premia. The implied betas are also time-varying and forecastable. However, there is evidence that other variables including innovations in consumption should also be considered in the investor's information set when estimating the conditional distribution of returns.

We wish to thank Douglas Breeden, Thomas MaCurdy, Ieuan Morgan, Michael Rothschild, and Larry Weiss for helpful comments. We would also like to thank seminar participants at the Fifth World Congress of the Econometric Society, University of Chicago, Northwestern University, New York University, University of Toronto, University of British Columbia, University of California at Berkeley, Irvine, and San Diego, and Yale University. An anonymous referee provided many useful suggestions in improving an earlier version of the paper. The research was supported by National Science Foundation grants SES 84-20680 and SES 80-08580 awarded to Engle.
I. Introduction

The capital asset pricing model (CAPM), originally proposed by Sharpe (1964) and Lintner (1965) following the suggestions of mean variance optimization in Markowitz (1952), has provided a simple and compelling theory of asset market pricing for more than 20 years. In its simplest form the theory predicts that the expected return on an asset above the risk-free rate is proportional to the nondiversifiable risk, which is measured by the covariance of the asset return with a portfolio composed of all the available assets in the market. The assumptions implicit in the version of the model discussed here are that (1) all investors choose mean-variance efficient portfolios with a one-period horizon, although they need not have identical utility functions; (2) all investors have the same subjective expectations on the means, variances, and covariances of returns; and (3) the market is fully efficient in that there are no transaction costs, indivisibilities, taxes, or constraints on borrowing or lending at a risk-free rate.

Empirical tests of the CAPM have tended to focus on assumption 1 while strengthening 2 to include the assumption that the common distributions are constant over time and that the entire market is the market for equities. These tests generally have found that the risk premium on individual assets can be explained by variables other than the estimated covariance. In particular, the own variance, firm size, and the month of January seem to be variables that help to explain expected returns. See, for example, Jensen (1972) for a survey of many of these early studies and Ross (1978), Roll and Ross (1980), Chen (1983), and Schwert (1983) for more recent surveys.

One interpretation for the failure of the CAPM to fully explain observed risk premia, due to Roll (1977), is that any empirical covariance is computed from an incomplete market for assets. Such an objection nearly makes the CAPM untestable. Another explanation is, of course, that alternative theories of asset pricing may be supportable such as the arbitrage pricing theory of Ross (1976) or the consumption beta formulation introduced by Breeden (1979).

In this paper we focus attention on the possibility that agents may have common expectations on the moments of future returns but that these are conditional expectations and therefore random variables rather than constants. For a discussion along these lines, see also Ferson (1985), Rothschild (1985), and Ferson, Kandel, and Stambaugh (1986).

Let $y_i$ be the vector of (real) excess returns of all assets in the market measured as the nominal return during period $t$, minus the nominal return on a risk-free asset, and let $\mu_i$ and $H_i$ be the conditional mean vector and conditional covariance matrix of these returns given infor-
mation available at time $t - 1$. Also let $\mathbf{\omega}_{t-1}$ be the vector of value weights at the end of the previous period so that the excess return on the market is defined as $y_{t} = y'_{t} \mathbf{\omega}_{t-1}$. Then the vector of covariances with the market is simply $\mathbf{H}_{t} \mathbf{\omega}_{t-1}$ and the CAPM requires

$$\mathbf{\mu}_{t} = \delta \mathbf{H}_{t} \mathbf{\omega}_{t-1}. \tag{1}$$

In this formulation, as derived by Jensen (1972), $\delta$ is a scalar constant of proportionality, which in equilibrium is an aggregate measure of relative risk aversion given by the harmonic mean of the agents' degree of relative risk aversion weighted by the agents' share of aggregate wealth (cf. Bodie, Kane, and McDonald 1983, 1984). Throughout the paper we assume $\delta$ to be constant.

The conditional variance of the market excess return is $\sigma_{M_{t}}^{2} = \mathbf{\omega}'_{t-1} \mathbf{H}_{t} \mathbf{\omega}_{t-1}$ and the conditional mean is $\mu_{M_{t}} = \mathbf{\omega}'_{t-1} \mathbf{\mu}_{t}$, which from (1) can be written as

$$\mu_{M_{t}} = \delta \sigma_{M_{t}}^{2}, \tag{2}$$

so that $\delta$ is seen to be the slope of the market trade-off between mean and variance. Defining the beta of an asset to be the covariance of that asset with the market divided by the variance of the market portfolio, $\beta_{t} = \mathbf{H}_{t} \mathbf{\omega}_{t-1} / \sigma_{M_{t}}^{2}$, and substituting in (1) and (2) yields the familiar expression

$$\mathbf{\mu}_{t} = \mathbf{\beta}_{t} \mu_{M_{t}}. \tag{3}$$

Because the covariance matrix of returns varies over time, the mean returns and the betas will in general also be time-varying.

We have stated the CAPM in terms of conditional moments since these reflect the information set available to agents at the time the portfolio decisions are made. But this model also implies a relation between unconditional moments. In the special case in which the value weights are fixed, the unconditional means are constant and are given by

$$E(\mathbf{y}_{t}) = \delta V(\mathbf{y}_{t}) \mathbf{\omega} - \delta^{3} V(\mathbf{H}_{t} \mathbf{\omega}) \mathbf{\omega}.$$

Only if $V(\mathbf{H}_{t} \mathbf{\omega}) = 0$ will the unconditional moments satisfy the same CAPM relations as the conditional moments. By a similar argument, if the econometrician uses only a subset of the relevant conditioning information, then the estimated conditional moments will not satisfy the CAPM.

In this paper the conditional covariance matrix of a set of asset returns is allowed to vary over time following the generalized autoregressive conditional heteroscedastic (GARCH) process (see Engle 1982; Bollerslev 1986). This essentially assumes that agents update their estimates of the means and covariances of returns each period.
using the newly revealed surprises in last period’s asset returns. Thus agents learn about changes in the covariance matrix only from information on returns. There may, of course, be additional information relevant to agents’ expectations that would lead to misspecification as mentioned above.

The approach is a multivariate generalization of Engle, Lilien, and Robins (1987), which treated a single asset, and therefore estimates the time-varying risk premium as a function of the conditional variance of that asset return alone. A similar idea was employed in the recent papers by French, Schwert, and Stambaugh (1986) and Poterba and Summers (1986). The approach can also be seen as a statistical implementation of the intertemporal CAPM of Bodie et al. (1983, 1984), in which they had no unknown parameters and no statistical test of the model performance. Finally, the paper can be viewed as a generalization of Frankel (1985), who assumed that $\omega_t$ may be time-varying but that $H_t$ is not, and of Friedman (1985a, 1985b), who allowed $H_t$ to be time-varying only because investors must learn about the unconditional variance $V(y_t)$.

II. Econometric Methods

According to the economic model (1), any explanation of time-varying expected excess holding yields should be built around a structure with a time-varying conditional covariance matrix. As mentioned above, a model ideally suited for this purpose is the multivariate GARCH in mean (GARCH-M) model. For $y_t, N \times 1$, the GARCH \((p, q) - M\) model takes the general form

$$ y_t = b + \delta H_t \omega_{t-1} + \epsilon_t, $$

$$ \text{vech}(H_t) = C + \sum_{i=1}^{q} A_i \text{vech}(\epsilon_{t-i} \epsilon_{t-i}^t) + \sum_{j=1}^{p} B_j \text{vech}(H_{t-j}), \quad (4) $$

$$ \epsilon_t | \psi_{t-1} \sim N(0, H_t), $$

where \text{vech}(\cdot) denotes the column stacking operator of the lower portion of a symmetric matrix, $b$ is an $N \times 1$ vector of constants, $\epsilon_t$ is an $N \times 1$ innovation vector, $C$ is a $\frac{1}{2}N(N + 1) \times 1$ vector, and $A_i, i = 1, \ldots, q$, and $B_j, j = 1, \ldots, p$, are $\frac{1}{2}N(N + 1) \times \frac{1}{2}N(N + 1)$ matrices. A nonzero $b$ vector might reflect a preferred habitat phenomenon or differential tax treatment of the assets. Of course the GARCH specification does not arise directly out of any economic theory, but as in the traditional autoregressive moving average time-series analogue, it provides a close and parsimonious approximation to the form of heteroscedasticity typically encountered with economic time-series data (cf. Bollerslev 1986; Engle and Bollerslev 1986).
The conditional log likelihood function for (4) for the single time period \( t \) can be expressed as

\[
L_t(\theta) = -\frac{N}{2} \log 2\pi - \frac{1}{2} \log |H_t(\theta)| - \frac{1}{2} \epsilon_t(\theta)'H_t^{-1}(\theta)\epsilon_t(\theta),
\]

(5)

where all the parameters have been combined into \( \theta' = (b', \delta, C', \text{vec}(A_1)', \ldots, \text{vec}(A_k)', \ldots, \text{vec}(B_p)') \), an \( m \times 1 \) vector. Thus, conditional on the initial values, the log likelihood function for the sample \( 1, \ldots, T \) is given by

\[
L(\theta) = \sum_{t=1}^{T} L_t(\theta).
\]

(6)

As is obvious from (4), (5), and (6), the log likelihood function \( L(\theta) \) depends on the parameters \( \theta \) in a highly nonlinear fashion, and the maximization of \( L(\theta) \) requires iterative methods. The approach taken here is to use the Berndt et al. (1974) algorithm along with numerical first-order derivatives to approximate \( \partial L_t(\theta)/\partial \theta \). These derivatives provide an added flexibility to changes in the specification.

Given standard regularity conditions (see Crowder 1976; Woolridge 1986), it follows that the maximum likelihood (ML) estimate for \( \theta \) will be asymptotically normal and unbiased with covariance matrix equal to the inverse of Fisher’s information matrix. Therefore, traditional inference procedures are immediately available. In particular, when equation (4) is tested versus a more general specification, the Lagrange multiplier (LM) test statistic takes the well-known form \( \text{LM} = T \cdot R^2_0 \), where \( R^2_0 \) is the uncentered coefficient of multiple correlation in the first Berndt et al. iteration for the augmented model starting at the ML estimates under the null (cf. Engle 1984).

As it stands, (4) is very general and involves a total of \( (N + 1) + \frac{1}{2}N(N + 1) + \frac{1}{4}N^2(N + 1)^2(p + q) \) parameters. A natural simplification is to assume that each covariance depends only on its own past values and surprises. Throughout this paper we shall therefore take \( p = q = 1 \) and impose diagonality on the matrices \( A_1 \) and \( B_1 \). With these simplifications, the GARCH(1, 1)–M model considered here becomes

\[
\begin{align*}
\eta_{it} &= b_i + \delta \sum_j \omega_{ij} h_{ijt} + \epsilon_{it}, \\
h_{ijt} &= \gamma_{ij} + \alpha_{ij} \epsilon_{i,t-1} \epsilon_{jt-1} + \beta_{ij} h_{ijt-1}, & i, j = 1, \ldots, N, \\
\epsilon_{it} | \psi_{t-1} &\sim N(0, H_t),
\end{align*}
\]

(7)

\[\text{In practice the presample values are set equal to their expected value, zero.}\]
where subscript $i$ refers to the $i$th element of the corresponding vector and $ij$ to the $ij$th element of the corresponding matrix. Thus only the own lagged moments and cross products appear in each of the conditional covariance equations.

Model (7) extends the univariate ARCH model introduced in Engle (1982) in several directions by allowing for multiple time series, conditional covariance terms in the mean, and own past conditional covariances in each of the covariance equations.

III. Data Description

The market portfolio studied in the present paper is composed of bills (6-month Treasury bills), bonds (20-year Treasury bonds), and stocks. In broad terms, these three assets account for a good part, but certainly not all, of the liquid investment opportunities available. The data are quarterly percentage returns from the first quarter of 1959 through the second quarter of 1984, for a total of 102 observations. The return on 3-month Treasury bills is taken to represent the risk-free return. For a detailed description of the data sources and data transformations, see the Appendix.

Two data sets have been analyzed for these three returns series. In the previous draft of this paper the Standard and Poor's 500 equity series was used with Citibase interest rates. In this version New York Stock Exchange value-weighted equity returns are used with Salomon Brothers bill and bond yields. The results are quite similar, so only the second data set will be discussed here. The original results are available from the authors.

The mean of the excess holding yield on 6-month bills over the sample is 0.142 percent at a quarterly rate while the standard deviation is 0.356. For bonds the average excess holding yield is −0.761 percent with standard deviation 6.255, and for stocks the excess holding yield and the standard deviation are −0.995 percent and 2.225. All the excess holding yield series, however, tend to be somewhat erratic. The maximum return on a 3-month balanced portfolio obtained by borrowing at the 3-month rate and lending at the 6-month rate was 2.046 percent at a quarterly rate in the second quarter of 1980. On the other hand, the three worst returns occurred in the first, third, and fourth quarters of 1980 with −0.462, −0.777, and −0.515 percent. For bonds, the best return on a balanced portfolio was also in the second quarter of 1980 with 22.274 percent, whereas the two worst returns were in the previous and subsequent quarters with −18.461 and −14.422 percent, respectively. Stocks did best in the first quarter of 1975 with 3.746 percent, but two quarters before, the return was as poor as −8.642 quarterly percentage rates. This sug-
gests that not only do the conditional mean excess holding yields vary over time, but also the conditional variances seem to be changing through time.

IV. Model Estimates

In this section we present model estimates for a trivariate CAPM. The econometric specification of the model is as in (7). The ML estimates (with corresponding standard errors in parentheses) are

\[
\begin{align*}
\begin{array}{c|c|c|c|c}
 y_{1t} & .070 & h_{1jt} & \epsilon_{1t} \\
(0.032) & & & \\
y_{2t} & -4.342 & + 0.499 \sum_{j} \omega_{jt-1} & h_{2jt} & + \epsilon_{2t} \\
(1.030) & (.160) & & & , (8a) \\
y_{3t} & -3.117 & h_{3jt} & \epsilon_{3t} \\
(0.710) & & & \\
\end{array}
\end{align*}
\]

\[
\begin{align*}
\begin{array}{c|c|c|c|c}
 h_{11t} & .011 & .445 \cdot \epsilon_{1t-1}^{2} & .466 \cdot h_{11t-1} \\
(0.004) & (.105) & (0.056) & \\
h_{12t} & .176 & .233 \cdot \epsilon_{1t-1} \epsilon_{2t-1} & .598 \cdot h_{12t-1} \\
(0.062) & (0.092) & (0.052) & \\
h_{22t} & 13.305 & + .188 \cdot \epsilon_{2t-1}^{2} & + .441 \cdot h_{22t-1} \\
(6.372) & (0.113) & (0.215) & , (8b) \\
h_{13t} & .018 & .197 \cdot \epsilon_{1t-1} \epsilon_{3t-1} & - .362 \cdot h_{13t-1} \\
(0.009) & (0.132) & (0.361) & \\
h_{23t} & 5.143 & .165 \cdot \epsilon_{2t-1} \epsilon_{3t-1} & - .348 \cdot h_{23t-1} \\
(2.820) & (0.093) & (0.338) & \\
h_{33t} & 2.083 & .078 \cdot \epsilon_{3t-1}^{2} & .469 \cdot h_{33t-1} \\
(1.466) & (0.066) & (0.333) & \\
\end{array}
\end{align*}
\]

where \( i = 1, 2, 3 \) refers to bills, bonds, and stocks, respectively.

The estimates for the model are appealing. The estimated value for \( \delta = 0.499 \) is reasonable and highly significant, lending some support for the theory presented here.

The intercept terms vary substantially across the three assets. Although the theoretical model does not include constant terms in the risk premia, these effects are of some importance. The large negative intercepts for bonds and stocks are not surprising since reduced capital gains taxes on long-term assets provide incentives to hold these assets even at otherwise unfavorable rates of return. It is also well known that bond and equity holders did consistently worse than other asset holders over the sample period. The intercepts -4.3 and -3.1 reflect this fact.
The dynamic structure in the second moments for bills and bonds is apparent as reflected by the significant variance and covariance parameters. Even though none of the six variance or covariance parameters for stocks is individually significant at the usual 5 percent level, it is interesting to note that the likelihood ratio test statistic for absence of dynamics in the second-order moments for stocks equals 18.639, which exceeds the .995 value of a \( \chi^2_6 \) distribution, thus soundly rejecting the null hypothesis. Any correctly specified intertemporal asset pricing model ought to take this observed heteroscedastic nature of asset returns into account. The same point has also been made in the recent papers by Ferson (1985) and Ferson et al. (1986). In particular, tests of the CAPM that treat the conditional covariance matrix as constant over time invariably falter.

The estimated risk premia from the model, \( b_i + \delta \sum h_{ijt} \omega_{jt-1} \), are plotted in figures 1–3 along with the excess holding yields for each of the three assets. Figures 1 and 2 show that the estimates for bills and bonds are fairly similar except for a difference in scale. Both assets have rising risk premia during the volatile post-October 1979 period. It is reasonable to believe that, on average, investors were paid a positive premium for holding bills or bonds during this period. Note that the negative premia observed for bonds and equities in some
Fig. 2.—Risk premia for bonds

Fig. 3.—Risk premia for stocks
Fig. 4.—Beta for bills

Fig. 5.—Beta for bonds
periods could be due to the preferential tax treatment as previously mentioned.

Plotted in figures 4–6 are the estimated betas. Not surprisingly, the beta for stocks is close to one, that for bonds is slightly above one, and that for bills is close to zero. There is, however, substantial movement over the sample period.

V. Diagnostic Tests

Given the evidence above, the trivariate CAPM seems to fit the data reasonably well. However, in order to assess the general validity of the model, a series of LM tests were performed. We shall consider here only a very small subset of these.

The first LM test involves the inclusion of the own conditional variances in each of the three equations for the conditional expectation of the excess holding yields. The test statistic equals 1.148, which is asymptotically a $\chi_3^2$ random variable if the null hypothesis is true. Thus the null cannot be rejected at any level under 25 percent, lending further support to the model. This test is of particular interest since in tests of the time-invariant CAPM the own variance is often found to be highly significant. This might also provide an explanation
for the empirical findings in French et al. (1986) and Poterba and Summers (1986), where a time-varying measure of the own conditional variance or volatility is found to have little explanatory power for the expected return on the stock market. Our results suggest that a better measure might be the nondiversifiable risk as given by the conditional covariance with the market.

The next test considers the lagged excess holding yields as explanatory variables for each of the three risk premia. This test rejects the formulation of the CAPM given in (8). The value of the test statistic is 18.311 and is highly significant at any reasonable level in the corresponding $\chi^2$ distribution. Thus agents may use information in addition to past innovations in forming their expectations. This ability of the lagged dependent variable to help forecast returns is not all that surprising in view of other recent results in the literature (see, e.g., Campbell 1987).

One of the competing theories of the intertemporal CAPM presented here is the consumption beta formulation mentioned in the Introduction. It is therefore interesting to note that the test for inclusion of innovations in the logarithm of per capita consumption in the conditional mean equals 14.027, the value of a $\chi^2$ random variable in the absence of any correlation. This surprising correlation of innovations in consumption and innovations in asset returns suggests that reformulation along the lines of a consumption beta model might deserve some consideration.\(^2\) However, the results in Hansen and Singleton (1982, 1983) do not lend much support to a strict formulation of that model. Also, Mankiw and Shapiro (1986) reported findings in favor of the traditional CAPM versus the consumption beta formulation.

VI. Conclusions

In summary, the results reported in this paper support several conclusions. First, the conditional covariance matrix of the asset returns is strongly autoregressive. The data clearly reject the assumption that this matrix is constant over time. The expected return or risk premia for the assets are significantly influenced by the conditional second moments of returns. There is also some evidence that the risk premia are better represented by covariances with the implied market than by own variances. However, information in addition to past innovations

---

\(^2\) The test probably has too small a size since, even if the CAPM were true, there might be consumption out of portfolio wealth, which would lead to a rejection. The rejection occurs only when future consumption is included, and this is, of course, not in the agents' information set. The LM test statistic for innovations in current consumption takes the value 3.289 corresponding to the .65 fractile in a $\chi^2$ distribution.
in asset returns is important in explaining premia and heteroscedasticity. In particular, lagged excess holding yields and innovations in consumption appear to have some explanatory power for the asset returns.

Probably even better econometric models with a richer specification for the risk premia, not necessarily derived directly from any economic theory, can therefore be constructed. See Hansen and Hodrick (1983) for a discussion along these lines. Other interesting questions that remain are the sensitivity of the results to the choice of the "market portfolio" and a quarterly one-period horizon. It is possible that wider definitions of the market would allow the model to do better. We leave the answer to all these questions for future research.

Data Appendix

The yields on 3-month Treasury bills, 6-month Treasury bills, and 20-year Treasury bonds were all taken from the Salomon Brothers' Analytic Record of Yields and Yield Spreads. The yields are percentages per annum for the first trading day in January, April, July, and October. The returns were converted

---

![Chart](chart.png)

**Fig. A1.**—Market weights
to quarterly rates $r^i_t$, $r^{\text{bill}}_t$, and $r^{\text{bond}}_t$, where $(1 + R_t)^{1/4} = 1 + r_t$. From these rates the one-quarter excess holding yields were calculated as

$$y^{\text{bill}}_t = 100 \left[ \frac{(1 + r^{\text{bill}}_t)^2}{1 + r^i_{t+1}} - 1 - r^i_t \right]. \quad (A1)$$

$$y^{\text{bond}}_t = 100 \left[ \left( \frac{r^{\text{bond}}_t}{r^{\text{bond}}_{t+1}} \right) + r^{\text{bond}}_t - 1 - r^i_t \right]. \quad (A2)$$

The stock market yields were based on the return on the value-weighted New York Stock Exchange index including dividends obtained from the Center for Research in Security Prices at the University of Chicago. From the quarterly flow of returns, $r^{\text{stock}}_t$, the one-quarter excess holding yield was simply calculated by

$$y^{\text{stock}}_t = 100(r^{\text{stock}}_t - r^i_t). \quad (A3)$$

The maturity distribution of the interest-bearing public debt held by private investors was taken from the Federal Reserve Bulletin and, for 1976 until the present, from the Treasury Bulletin. To get from par values to market values, we multiplied the outstanding debt in the different maturity categories within 1 year, 1–5 years, 5–10 years, and 10 years and over by the price indices reported in Cox (1985). The categories were then added together to get the market values within 1 year, bills, and longer than 1 year, bonds. The total market value of corporate equities was obtained from a special tabulation of the balance sheets of the flow of funds accounts by the Board of Governors of the Federal Reserve System. The relative market values are illustrated in figure A1.

Finally, the data for personal consumption expenditure on nondurables in 1972 dollars, $C_o$, were obtained from Citibank Economic Database and from the Survey of Current Business.

References


