The Likelihood Ratio Test of CAPM: A New Approach

Yong-Ho Baek*

According to capital-asset pricing model, the risk premium on a security is proportional to the beta coefficient. In the traditional model (Sharpe 1964), the market beta determines the risk premium; according to the consumption-based model (Ereeden 1979), the consumption beta determines the risk premium. In this paper, working with the consumption-based model, we present a likelihood-ratio test of the model, which tests whether indeed the mean rate of return is a linear function of the beta.

In part 1 we review the econometric procedures previously used to test the standard capital-asset pricing model, and we discuss problems with these procedures. One methodology is the "one-pass regression test for zero intercept," and another is the "two-pass regression test." At present there are no published tests of the consumption-based CAPM. In part 2, after we state the assumptions in the model and describe the null hypothesis we wish to test, we demonstrate the test procedures. In part 3 we draw the conclusion.

I. Review of CAPM Tests

a) One-pass Regression Test for Zero intercept

The Standard capital-asset pricing model developed by Sharpe (1964) and Lintner (1965) states.

$$E_t\ (R_{ti}) - R_{t} = \infty Cov_t(R_{ti}, R_{tm}). \tag{1}$$

Herde α denotes the relative risk aversion; $E_t(R_{ti})$ is the expected of return during period t on asset i;R_{to} is the risk-free interest rate; R_{ti} is the rate of return on asset i;R_{tm} is the rate of return on the market

* Ehwa Women's University

portfolio of all assets, which includes human wealth as well as non-human wealth. The subscript t on the expected value and the covariance signifies that they are conditional on all information available at the beginning of period t. The model (1) states that the risk premium is proportional to the covariance of the rate of return with the rate of return on the market portfolio.

Jensen (1968) derives the "one-pass regression test for zero intercept." From (1), it is straightforward to show.

$$E_t(R_{ti}-R_{to})=\beta_i E_t(R_{tm}-R_{to}). \qquad (2)$$

,Where β_i is $Cov_t(R_{ti},R_{tm})/Var_t(R_{tm})$.

Consider the least—squares linear regression of $R_{ti}-R_{to}$ on $R_{tm}-R_{to}$:

$$R_{ti}-R_{to}=\beta_i+\beta_i(R_{tm}-R_{to})+e_{ti}. \qquad (3)$$

By(2),∞_i=o. Hence one can test the standard CAPM by regressing $R_{ti}-R_{to}$ on $R_{tm}-R_{to}$ in a time-series analysis. The null hypothesis is that the coefficient of the constant term is zero. Jensen (1968), Friend and Blume (1970), and Black, Jensen and Scholes (1972) perform this test.

Using data on annual period returns for various mutual fund shares, Jensen (1968) finds that the intercepts differ quite a bit from zero.

Friend and Blume (1970) use an average of the returns on common stocks listed on the New York Stock Exchange as R_{tm} and base the time series for R_{to} on the returns on Treasury bills. Analyzing the monthly holding period return R_{ti} on portfolios of large numbers of common stocks listed on NYSE, they regress $R_{ti}-R_{to}$ in $R_{tm}-R_{to}$ for many portfolios. They find a linear relationship between the estimated intercepts and the estimated β_i. When $\beta_i>1$, the intercepts tend to be negative; and when $\beta_i<1$, the intercepts tend to be positive. They interpret this finding as evidence against the hypothesis(2).

They note that mismeasurment of R_{tm} might account for their findings. They also note that correlation among the residuals of the different regressions related in a systematic way to the intercepts and the slope would reconcile their findings with the model (2), but they conjecture that such a relationship does not hold. However their graphs do suggest that such a relationship does exist.

If the residuals among different securities are cerrelated we have a non-diagonal variance-covariance matrix of the error terms for the model (2) One cannot apply the usual T-test to test the zero intercept; instead, one must use an F-test.

Unfortunately one cannot adopt the one-pass regression test for zero intercept to test the consumption-based model.

b) Two-pass Regression Test

Several investigators Sharpe (1965), Miller and Scholes the traditional CAPM by a two-pass regression procedure. They test whether.

$$\mu_i = f + g\beta_i \tag{4}$$

is a good model for the holding period returns on securities, where the parameters f and g are stable through time. Here μ_i is $E_t(R_{ti})$ and β_i is $Cov_t(R_{ti},R_{tm})/Var_t(R_{tm})$. They assume that f,g, $Var_d t(R_{ti})$, $Var_t(R_{tm})$, and the correlation of R_{ti} and R_{tm} are the same for all time periods.

Since g and β_i are not observable, they adopt the "two-pass regression test." At the first pass, they regress R_{ti} on the R_{tm} to obtain the least-squares estimates μ^* and β^*. At the second stage, one runs a least-squares regression of μ^* or β^* to obtain estimates of f and g. One accepts the capital-asset pricing model if the following are true. The estimate of f must be reasonable value for the riskless rate of return. The estimate of g must be positive, a reasonable price of risk. And the second stage R^2 is high (it will be high if we have many years of data and if the model is valid).

Using data on annual hold king period returns on various mutual fund shares, Sharpe (1964) notes that the correlation of these returns with an average of the returns on these common stocks used to compute the Dow Jones Industrial Average is near one. Estimating the means and standard deviations of these returns over a ten year period, he regresses the estimated means for the various shares on the estimated standard deviations. If the model (4) is valid and the correlation of the returns with R_{tm} is near one, the constant term in the regression is an estimate of R_{to}; and the coefficient of the standard deviation provides an estimate of the gk. Sharpe finds that the constant term is a plausible value for a riskless rate of return and that the coefficient of the standard deviation is positive. Noting graphically that the relationship between the estimated means and the standard deviations is roughly linear, he concludes that his result supports the model(4).

Fama and MacBeth (1973) include β_i^2 to test for a non-linear relation between μ and β. They find that non-linear terms are significantly different from zero over some sample periods and offer a possible explanation for the significance of the non-linear beta term; viz, they suggest that there are omitted variavles from the theory for which the non-linear terms act as a proxies. If some omitted variables help explain the error term, the variance-covariance matrix of the error terms V may not be diagonal matrix. Miller and Scholes (1972) also have results similar to Fama and MacBeth.

Levy (1982) argues that even if beta is constant over time and there are no biases in the estimation of β in the first-pass regression, we cannot reject the standard CAPM (4), especially because of the uncertainty of β.

Since this two-pass methodology can be adapted to test the consumption-based model, here we will discuss several problemes. At the second stage, one treats the β_i as known exactly. One then uses least-squares to estimate f and g. Although the estimate is consistent under very general assumptions, the estimates are inefficient, unless the variance-covariance matrix of the R_{ti} is proportional to the identity matrix. More important, the computed errors are extremely misleading.

The important problem is that actually the β_i are estimated, not known. (One must take the error into account when one estimates f and g.) Allowing for the error will of the estimates and will affect the calculation of their standard error.

Users of the two-pass procedure demand a high R^2 at the second pass. However the estimates μ^* and β^* will not perfectly correlated even if there does exist a perfect linear relationship between μ and β. The size of the correlation will depend on the error of measurement. One requires a proper test whether μ is a linear function of β, in which one takes into account the uncertainty in β.

Even if the sizes of the estimates f and g may be "reasonable", equation (4) yet man not hold. The two-pass procedure does not relly test for a linear relation between the μ_i and the β_i.

To overcome these deficiencies, in the next part, we derive the likelihood-ratio test of CAPM.

Ⅱ. Maximum-Likelihood Estimation

a) Assumptions

Let R_{ti} denote the n-vector of the rates of return on asset i observed at time t. Define C_{pt} as an unexpected change and R_{ti} are joint white noise, normally distributed with constant mean ($E(c_{pt})=0$ and $E(R_{ti})=\mu_i$) and a constant variance-covariance matrix. Also we assume that each period we can observe an unexpected change in permanent consumption c_{pt} describing the aspect of the state of the world.

Let

$$R_{ti}=\mu_i+\beta_i c_{pt}+e_{ti} \tag{5}$$

denote the least-squares regression of R_{ti} on c_{pt}. Since c_{pt} has mean zero, the intercept is μ_i. We have $\beta_i/\mathrm{Cov}(R_{ti},c_{pt})= \mathrm{Var}(c_{pt})$, which is independent of time. The vector e_t of error terms is white noise, independent of c_{pt} with $E(e_t)=0$ and $\mathrm{Var}\ (e_t)=V$.

According to the permanent-consumption-based capital asset pricing model, the mean rate of return μ_i on security i equals the risk-free rate of return f plus the price of risk g multiplied by its beta coefficient β_i,

$$\mu_i=f+g\beta_i, \tag{6}$$

for coefficients f and g which are independent of i. We test whether μ is a linear function of β, by a maximum likelihood-ratio test. The linear constraint (6) embodies the hypothesis that the capital-asset pricing model is valid. We test the null hypothesis against the alternative hypothesis that there is no particular relationship between mean and beta. We perform such a test based on the assumption that μ_i and β_i are the same for all periods.

b) Maximum Likelihood Estimation

Let us define Y as the T×n matrix of the observed rates of return. Then the matrix form of (5) is

$$Y=1_t\mu'+c_p\beta_i'+E, \tag{7}$$

where μ and β are n-vector of unknown parameters; E is the T×n matrix of normally distributed error terms; i_t and c_p are T-vector of ones and the change in consumption.

We define X as the T×2 matrix of independent variables,

$$X=[1_t \ c_p], \tag{8}$$

in which 1_T is a T-vector of ones and c_p is the T-vector of the c_{pt}. We assume rank(X)=2. We define β as the 2×n matrix of regression coefficients.

$$B=[\mu \ \ \beta]'. \tag{9}$$

Then we rewrite (7) as

$$Y=XB+E. \tag{10}$$

We rewrite the linear constraint(6) in the symmetric form

$$pB'+q1_n'=0. \tag{11}$$

Here p is an unknown column 2-vector $p=[p_1, \ p_2]'$ and q is an unknown scala; 1_n denotes a n-vector of ones. We have $f=-q/p_1$ and $g=-p_2/p_1$. We add the constraint

$$p'(X'X)^{-1}p=1, \tag{12}$$

since otherwise p and q would be identified only up to some multiple; (12)identifies p and q up to the sign. We assume $n\geq3$; otherwise the constraint on B would be meaningless.

We find B to maximize the likelihood. We treat c_p as exogenous, with sample mean O ($1_T'c_p=0$) and sample variance $\sigma^2(c_p'c_p=T\sigma^2)$. Treating c_p as exogenous, we work with the likelihood conditional on c_p. The variance-covariance matri×V is known. Since

$$\sum_{t=1}^{T} e_t V^{-1} e_t' = \mathrm{tr} E V^{-1} E' = \mathrm{tr} E'EV^{-1},$$

the value of the log of the likelihood function is

$$\ell nL=(-nT/2)\ \ell n2\Pi+(T/2)\ell n\ |V^{-1}|-(1/2)\ \mathrm{tr}\ E'EV^{-1}. \tag{13}$$

Therefore the maximum-likelihood estimate minimizes tr $E'EV^{-1}$.

If β were unconstrained, it is well known that one would obtain the maximum-likelihood estimate β^* of B by applying ordinary least squares to each of the n equations in (10).

We have

$$B^*=(X'X)^{-1}X'Y. \tag{14}$$

The log of the likelihood is

$$\ell nL^{*}=(-nT/2)\ell n\ 2\pi+(T/2)\ell n\ |\ V^{-1}\ |\ -(1/2)\mathrm{tr}\ E^{*\prime}E^{*}V^{-1}, \tag{15}$$

in which $E^{*}=Y-XB^{*}$ denotes the estimated error matrix. We define

$$\phi^{*}=p'B^{*}+q1_{n}', \tag{16}$$

which measures the extent to which the unconstrained estimate B^{*} fails to satisfy the linear constraint(11).

We wish to maximize the likelihood(13), subject to the linear constraint(11) and the identifying constraint(12). We use Lagrangian maximization. One could define the Lagrangian

$$H=\ell nL-(p'B+q1_{n}')\gamma-\omega(p'(X'X)^{-1}p-1), \tag{17}$$

in which γ is an n-vector of Lagrange multipliers and ω is a scalar Lagrange multiplier for the identifying constraint.

This technique divides the maximization into two steps. At the first step, we assume that p and q are know and satisfy the identifying constraint(12). We find B(p, q) to maximize the likelihood, subject to the linear constraint(11). Substituting B(p, q) into the likelihood gives the concentrated likelihood, a function of p and q. At the second step, we find p and q to maximize the concentrated likelihood, subject to the identifying constraint(12).

For step one, we define the Lagrangian

$$J=\ell nL-(p'B+q1_{n}')\gamma. \tag{18}$$

Differentiation with respect to the parameters yields the firstorder necessary condition;

$$\partial J/\partial B=(X'Y-X'XB)V^{-1}-p\gamma'=0, \tag{19}$$

Premultiplying by $(X'X)^{-1}$ gives

$$(B^{*}-B)V^{-1}-(X'X)^{-1}p\gamma'=0. \tag{20}$$

Premultiplying by p' and using the (11) and (12) yields

$$\gamma=V^{-1}(B^{*\prime}p+1_{n}q). \tag{21}$$

We can then rewrite (21) as

$$B=B^{*}-(X'X)^{-1}p[p'B^{*}+q1_{n}']. \tag{22}$$

The expression in the bracket [ ] in (22) measures the extent to which the unconstrained estimate B^{*} fails to satisfy the linear constraint (11). Equation (22) says that $B=B^{*}$ if the unconstrained estimate B^{*} happens to satisfy the constraint exactly.

For step two, we first derive and then maximize the concentrated likelihood function. By (22),

$$\begin{aligned}E'E&=(Y-XB)'\ (Y-XB)\\&=[E^*+X(X'X)^{-1}p(p'B^*+q1_n')]'[E^*+(X'X)^{-1}p(p'B^*=q1_n')]\\&=E^{*\prime}E^*+(B^{*\prime}p+1_nq)(p'B^*+q1_n'),\end{aligned}$$

since $E^*X=0$ and $p'(X'X)^{-1}p=1$. Therefore,

$$\begin{aligned}\text{tr }E'EV^{-1}&=\text{tr }DE^{*\prime}E^*V^{-1}+\text{tr}(B^{*\prime}p+1_nq)(p'B^*=q1_n')V^{-1}\\&=\text{tr }E^{*\prime}E^*V^{-1}+(p'B^*+q1_n')V^{-1}(B^{*\prime}p+1_nq).\end{aligned}\tag{23}$$

Substituting this expression into (13) gives the concentrated likelihood.

$$\begin{aligned}\ell nL=&(-nT/2)\ell n2\pi+(T/2)\ell n\,|V^{-1}|-(1/2)\text{tr }E^{*\prime}E^*V^{-1}\\&-(1/2)(p'B^*+q1_n')V^{-1}(B^{*\prime}p+1_nq).\end{aligned}\tag{24}$$

At the second step, we maximize ℓnL, subject to the identifying constraint(12).

Because, in (24), tr $E^{*\prime}E^*V^{-1}$ is independent of p and q, our problem is the following:

$$\min_{p,q}(p'B^*+q1_n')V^{-1}(B^{*\prime}p+1_nq),\tag{25}$$

subject to the identifying constraint (12). We minimize a quadratic function subject to a quadratic constraint. We define the Lagrangian

$$H^*=(p'B^*+q1_n')V^{-1}(B^{*\prime}p+1_nq)-\lambda[p'(X'X)^{-1}p-1],\tag{26}$$

in which λ is a scalar Lagrange multiplier for the constraint. Differentiation with respect to the p and q yields the firstorder necessary conditions:

$$\partial H^*/\partial p=(p'B^*=q1_n')V^{-1}B^{*\prime}-\lambda p'(X'X)^{-1}=0,\tag{27}$$

and

$$\partial H^*/\partial q=(p'B^*+q1_n')V^{-1}1_n=0\tag{28}$$

From(28),

$$q=-p'B^*V^{-1}1_n/1_n'V^{-1}1_n.\tag{29}$$

Substituting this value for q into(27) yields

$$p'B^*MB^{*\prime}-\lambda p'(X'X)^{-1}=0\tag{30}$$

in which

$$M=V^{-1}-V^{-1}1_n1_n'V^{-1}/1_n'V^{-1}1_n.\tag{31}$$

Equivalently,

$$[(X'X)^{1/2}B^*MB^{*\prime}(X'X)^{1/2}](X'X)^{-1/2}p=\lambda(X'X)^{-1/2}p.\tag{32}$$

We can rewrite the matrix in brackets as

$$T\begin{bmatrix} \mu^{*}M\mu^{*} & \mu^{*\prime}M\sigma\beta^{*} \\ \sigma^{*\prime}M\mu^{*} & \sigma^{2}\beta^{*\prime}M\beta^{*} \end{bmatrix} \tag{33}$$

where μ^* and β^* are the unconstrained OLS estimates. Equation (32) says that $(X'X)^{-1/2}p$ is an eigenvector of $(X'X)^{1/2}B^*MB^*(X'X)^{1/2}$, with eigenvalue λ.

We have reduced the minimization to an eigenvalue problem : we simply find the eigenvalues and eigenvectors of the matrix. There are two solutions to the first-order conditions, one for each eigenvalue. Which eigenvalue yields the optimum ? Since the matrix M is positive definite, the matrix(33) is positive semi-definite. So both eigenvalues are non-negative. Postmultiplying (27) by p and multiplying (28) by q and adding yields

$$\lambda=(B^*p+q1_n')V^{-1}(B^{*\prime}p+1_nq); \tag{34}$$

λ therefore equals the value of the objective function in (25). Consequently the optimum is dthe smaller eigenvalue.

Given any eigenvalue λ in equation (32), one can determine parameter estimates which satisfy the first-order necessary conditions. The vector $(X'X)^{-1/2}p$ must be the eigenvector for λ, normalized so that he identifying constraint(12) holds. One can then determine B and λ from the equations(21) and (22).

One can understand this result better by considering a special case. Suppose that by coincidence there exists an exact linear relationship between the unconstrained estimates μ^* and β^*:

$$\mu^*=f1_n+g\beta^*. \tag{35}$$

Then the matrix (33) equals the following matrix with rank one,

$$T\beta^{*\prime}M\beta \begin{pmatrix} g \\ \sigma \end{pmatrix}\begin{pmatrix} g \\ \sigma \end{pmatrix}'.$$

Therefore the smaller eigenvalue is $\lambda=0$, and one can calculate

$$\begin{pmatrix} g \\ p_2 \\ q \end{pmatrix} \alpha \begin{pmatrix} -1 \\ g \\ f \end{pmatrix}.$$

Conseqently the coefficients f and g in (35) are the maximum-likelihood estimates.

Finally we test the linear relation (11) by a likelihood-ratio test. The equation (11) embodies the null hypothesis that capital-asset pricing model is valid. We test the null hypothesis against the alternative hypothesis that there is no particular relationship between μ and β. The null hypothesis places n—2 restriction on the parameters. Define $\phi = L^*/L$, the ratio of the unconstrained likelihood to the constrained likelihood. Comparing (13) and (15) reveals

$$\begin{aligned} -2\ell n\phi &= \mathrm{tr}\ E'EV^{-1} - \mathrm{tr}\ E^{*\prime}E^*V^{-1} \\ &= \lambda . \end{aligned} \tag{36}$$

If the null hypothesis is valid in a large sample, the smaller eigenvalue λ therefore has a chi-square distribution with n—2 degrees of freedom; one accepts the capital-asset pricing model as long as λ is not too big.

Ⅲ. Conclusion

In this paper we derive a maximum-likelihood estimate of return is a linear function of the beta coefficient. We estimate the model subject to this constraint, using Lagrangian maximization. Calculating the likelihood-ratio for the constrained versus the unconstrained model tests whether the CAPM is valid. Especially we show that the likelihood-ratio is the smaller eigenvalue of the matrix consisting of the unconstrained OLS estimates.

References

Aitchison, J. and S.D. Silvey (1958), "Maximum-Likelihood Estimation of parameters Subject to Restraints," *Annals of Mathematical Statistics,* 29(3), pp.813-828.

Black, F., M.C. Jensen and M. Scholes (1972), "The Capital Asset Pricing Model; Some Empirical Tests," in *Studies in the Theory of Capital Markets,* edited by M.C. Jensen, New York:Praeger, pp. 79-121.

Breeden, D.T. (1979), "An Intertemporà Asset Pricing Model with Stochastic Consumption and Investment Opportunities,"*Journal of Financial Economics,* 7(3), pp.265-296.

Fama, E.F. and J.D. MacBeth (1973), "Risk, Return and Equilibrium; Empirical Tests," *Journal of Political Economics,* 81, pp.607-636.

Friend, I. and M.E. Blume (1970), "Measurment of Portfolio Performance under Uncertainty," *American Economic Review,* 60(4), pp.561-575.

Gibbons, M. (1982), "Multivariate Tests of CAPM; A new Approach," *Journal of Financial Economics,* 10(1), pp.327.

Goldberger, A.S. (1964), *Econometic Theory New York*:John Wiley & Sons, Inc.

Jensen, M.C. (1968), "The Performance of Mutual Funds in the Period 1945-1964," *Journal of Finance,* 23(2), pp.389-416.

Levy,H. (1982), "A Test of the CAPM via a confidence Level Approach," *Journal of Portfolio Management,* 9, pp.56-61.

Lintner, J. (1965), "The Valuation of Risk Assents and the Selection of Risky Investment in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics,* 47, pp.13-37.

Markowitz, H. (1952), "Portfolio Selection," *Journal of Finace,* 7, pp.77-91.

____________ (1959), Portfolio Selection; Efficient Diversification of Investments, New York: John Wiley.

Miller, M. and M. Scholes (1972), "Rates of Return in Relation to Risk; A Re-Examination of Some Recent Findings," in *Studies in the Theory of Capital Market,* edited by M.C. Jensen New York; Praeger.

Roll, R. (1977), "A Critique of Asset ricing Theory's Tests," *Journal of financial Economics,* 4, pp.129–176.

Roll, R. and S.A. Ross (1980), "An Empirical Investigation of the Arbitrage Pricing Theory," *Journal of Finance,* 35, pp.1073-1103.

____________ (1978), "The Current Status of the Capital Asset Pricing Model," *Journal of Finance,* 33(3), pp.885-901.

Sharpe, W.F.(1964), "Capital Asset Prices; A Theory of Market Equilibrinm under Conditions of Risk," *Journal of Finance,* 19(2). pp.425-442.

____________ (1965),"Risk-Aversion in the Stock Market; Some Empirical Evidence," *Journal of Finance,* 20(3), pp.416-422.

Strang, G. (1976), Linear Algebra and Its Applications, New York; Academic Press, Inc.