Sensitivity of Multivariate Tests of the Capital Asset-Pricing Model to the Return Measurement Interval
Author(s): Puneet Handa, S. P. Kothari, Charles Wasley
Source: *The Journal of Finance*, Vol. 48, No. 4 (Sep., 1993), pp. 1543-1551
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2329051
Accessed: 25/11/2009 11:35

THE JOURNAL OF FINANCE • VOL. XLVIII, NO. 4 • SEPTEMBER 1993

Sensitivity of Multivariate Tests of the Capital Asset-Pricing Model to the Return Measurement Interval

PUNEET HANDA, S. P. KOTHARI, and CHARLES WASLEY*

ABSTRACT

The capital asset-pricing model's (CAPM) primary empirical implication is a positively sloped linear relation between a security's expected rate of return and its relative risk (beta). Recent research indicates that inferences about the risk-return relation are sensitive to the choice of the return measurement interval. We perform multivariate tests of the Sharpe-Lintner CAPM using monthly and annual returns on market-value-ranked portfolios. The CAPM is rejected using monthly returns, a result consistent with previous research. In contrast, we fail to reject the CAPM when annual holding period returns are used.

THE CAPITAL ASSET-PRICING model's (CAPM) primary empirical implication is a positively sloped linear relation between a security's expected rate of return and its relative risk (beta):

$$E(R_i) = \gamma_0 + \gamma_1 \beta_i, \tag{1}$$

where $E(R_i)$ is the expected rate of return on security i, β_i is the security's relative risk (beta) defined as $\mathrm{cov}(R_i, R_m)/\mathrm{var}(R_m)$, γ_0 is the risk-free rate of return or the expected rate of return on a zero-beta profile, and γ_1 is the positive market risk premium.

The CAPM relation (1) has been tested extensively; see, for example, Fama and MacBeth (1973), Gibbons (1982), Stambaugh (1982), Shanken (1985), and MacKinlay (1987). In testing relation (1), researchers use beta estimates and realized returns measured over an arbitrary measurement interval (typically one month). The effect of measurement error in proxies for "true" betas on tests of the CAPM has been examined by Miller and Scholes (1972), Shanken (1985), Chan and Chen (1988), and Gibbons, Ross, and Shanken (1990), among others. The choice of the return measurement interval in testing relation (1), however, has largely been unexplored. This is perhaps because theory does not provide explicit guidance on this issue, and

* The authors are from New York University, the University of Rochester, and Washington University, respectively. We are grateful to Yakov Amihud, Jay Shanken, and an anonymous referee for many suggestions. Puneet Handa acknowledges support from the Taggart Foundation and the Stern School of Business, NYU, New York. S. P. Kothari acknowledges support from the Managerial Economics Research Center at the Simon School and from the John M. Olin Foundation. Charles Wasley acknowledges the support of the Olin School of Business at Washington University, St. Louis.

previous research generally assumes assets' "true" betas are invariant to the return measurement interval.[1]

There has been recent interest in examining the implications of varying the return measurement interval on tests of the risk-return relation. Handa, Kothari, and Wasley (1989) provide evidence that cross-sectional dispersion in betas is an increasing function of the return measurement interval and that the univariate tests of CAPM are sensitive to the choice of the interval. In a related area, Campbell and Clarida (1987) and Lewis (1988) fail to reject the international CAPM using three-month holding period returns on foreign exchange and bonds whereas others (e.g., Hodrick and Srivastava (1984) and Campbell (1987)) reject it using return observations measured over shorter intervals.

There are at least two reasons for the sensitivity of beta to the return measurement interval. First, Levhari and Levy (1977) and Handa, Kothari, and Wasley (1989) show that for buy-and-hold returns the covariance of an asset's return with the market and the variance of the market return do not increase proportionately, implying that "true" betas are sensitive to the return interval. Second, changes in risk and changes in the expected rate of return on the market (French, Schwert, and Stambaugh (1987), Fama and French (1988), and Ferson and Harvey (1990)) induce negative serial correlation in returns. If the degree of serial correlation in returns is not the same across subsets of the market portfolio, relative risk estimates are affected by the return measurement interval.[2] Another implication of autocorrelated returns is that return variability, and hence the investors' perception of risk, is a function of investment horizon. This suggests that the marginal investor's investment horizon (i.e., return measurement interval) may be of relevance.

In this paper we test whether results from multivariate tests of the CAPM are sensitive to the choice of return measurement interval. We examine whether the risk-return behavior is consistent with the Sharpe-Lintner CAPM using monthly and annual return data on 5, 10, and 19 market-value-ranked portfolios. The CAPM using monthly returns is rejected, a result consistent with previous research. In contrast, we fail to reject the CAPM when annual holding period returns are used. Because the difference in results could be due to low power when annual returns are used, we perform simulations to assess power. The simulations reveal that, when the number of cross-sectional assets is 5 or 10, the test can detect fairly small deviations of the Treasury bill rate from the true risk-free rate. The power is relatively low when 19 assets are employed. Overall, our analysis suggests that investment horizon and beta sensitivity to the return measurement interval are two important factors affecting the risk-return relation.

[1] If returns are serially uncorrelated, asset betas estimated using continuously compounded returns are invariant to the measurement interval. However, the return measurement interval affects the cross-sectional dispersion in betas when buy-and-hold returns are used (see e.g., Levhari and Levy (1977)).

[2] Evidence in Fama and French (1988) and Lo and MacKinlay (1990), among others, suggests the degree of serial correlation in returns varies across market-value-ranked stocks or portfolios.

Section I outlines a multivariate test of the Sharpe-Lintner CAPM. Section II provides details of our test methodology. Finally, Section III contains our empirical results based on an analysis of monthly and annual portfolio returns.

I. A Multivariate Test of the CAPM

In the presence of a riskless asset, the CAPM relation (1) can be written in excess return form as

$$E = \beta\gamma_1 \tag{2}$$

where

E = an N-vector of expected excess returns on N assets,
β = an N-vector of "true" betas of the N assets,

and

γ_1 = positive market risk premium.

In testing the CAPM relation (2), the null hypothesis is that expected excess returns (over an unspecified measurement interval) on the N assets are linear in their "true" betas. Since the "true" and estimated betas are a function of the return measurement interval and the beta-expected return relation is affected by the investment horizon, the return measurement interval can potentially have a significant role in tests of the CAPM.

To test relation (2), excess return market model regressions for the N assets each with T observations are defined:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it} \qquad \text{for } i = 1, \ldots, N \tag{3}$$

where R_{it} is the excess return on security i in period t, R_{mt} is the period t excess return on the market portfolio, α_i and β_i are intercept and slope parameters, and ϵ_{it} is the normally distributed disturbance term. Let Σ be the covariance matrix of ϵ_{it} and S be the sample covariance matrix of the N time series of ordinary least squares (OLS) market model residuals. Under the null hypothesis, all the assets' α's are zero. MacKinlay (1987) provides a test statistic for this:

$$\theta_1 = K\left[1 + \mu_m^2/\sigma_m^2\right]^{-1} \hat{\alpha}' S^{-1} \hat{\alpha} \tag{4}$$

where $K = (T - N - 1)T/(T - 2)N$, μ_m and σ_m^2 are the sample mean and variance of the excess return on the market over T periods, and $\hat{\alpha}$ is the vector of estimated intercepts from the market model regressions in equation (3). The distribution of θ_1, conditional on the market return, is a noncentral F with N and $T - N - 1$ degrees of freedom (MacKinlay (1987)). The noncentrality parameter λ is given by

$$\lambda = T\left[1 + \mu_m^2/\sigma_m^2\right]^{-1} \alpha' \Sigma^{-1} \alpha \tag{5}$$

Under the null hypothesis that the Sharpe-Lintner CAPM holds, the distribution of θ_1 is a central F with N and $T - N - 1$ degrees of freedom and the noncentrality parameter is zero. To assess statistical significance, however, we rely on an empirical distribution of the test statistic under the null, because large sample theory may not be applicable in our context.

II. Methodology

Empirical analysis is based on monthly and annual buy-and-hold return data from 1927 to 1988 (62 years) for the New York Stock Exchange–American Stock Exchange (NYSE-AMEX) stocks. We initially construct 20 portfolios (MV1–MV20) each year by ranking all securities on their beginning of the year market values of equity. Securities with at least one year's return data available on the Center for Research in Security Prices (CRSP) monthly returns tape are included. MV1 is an equally weighted portfolio consisting of the 5 percent smallest market capitalization stocks whereas MV20 is an equally weighted portfolio consisting of the 5 percent largest market capitalization stocks. Portfolio MV10 is excluded from the analysis to avoid singularity and the tests are performed using return data on the remaining 19 portfolios, and on 5 and 10 portfolios constructed from the 19 portfolios, over three subperiods of 20 years and eight months each. The starting months of these subperiods are January 1927, September 1947, and May 1968.

For tests of the CAPM using annual returns we use overlapping return data. Fama and French (1988), and Poterba and Summers (1988), among others, have recently used overlapping return data to increase power and to (make it feasible to) examine security return behavior over longer measurement intervals. One consequence of using overlapping data is that returns and residuals are serially correlated, so we use the Newey and West (1987) correction to obtain an autocorrelation and heteroskedasticity-consistent covariance matrix of the residuals. Adjacent annual returns have an overlap of 11 since monthly observations of annual returns are used. We conservatively allow 18 lags to correct for autocorrelation in the residuals.[3]

While the use of overlapping annual returns yields a more powerful test, two issues remain: (1) using large sample theory may not be appropriate because the (effective) number of independent observations using annual return data is low compared to the number of assets; and (2) the power of the test may be low since the (effective) number of independent observations is low. The first issue is addressed by comparing the test statistic to its empirical distribution obtained by imposing the null hypothesis. To address the second issue, we examine the power of the test using simulations.

The empirical distribution of the test statistic under the null is generated from the following bootstrapping experiment performed separately for 5, 10, and 19 portfolios, and for monthly and overlapping annual return observations. In each simulation, the excess return vector for period t is defined as

[3] We find that results are insensitive to using 12, 18, or 24 lags for autocorrelation correction.

the sum of bR_{mt} and a market model residual vector that is drawn with replacement from the matrix of estimated market model residuals (i.e., equation (3)). The constructed excess return vectors thus have two characteristics: (1) they are obtained by imposing $\alpha_i = 0$ for all assets; and (2) they retain the cross-correlation and heteroskedasticity characteristics of the original data. The θ_1 test statistic in equation (4) is calculated using the above excess return matrix. The procedure is repeated one-thousand times to generate an empirical distribution of the test statistic. This distribution is used to assess the empirical p-value implied by the test statistic for the original asset return data.

We assess the tests' power by examining their ability to detect deviations of the Treasury bill rate from the true risk-free rate. Under the null hypothesis that the CAPM holds, the Treasury bill rate equals the true risk-free rate, and the noncentrality parameter λ defined in equation (5) equals zero. The alternative hypothesis is that the Treasury bill rate equals the true risk-free rate plus a constant, and the corresponding noncentrality parameter λ is nonzero. In our simulations, we obtain 200 samples of excess returns imposing the null hypothesis that λ equals 0 and estimate θ_1 for each sample. This gives us the empirical distribution of θ_1 under the null hypothesis and allows us to assess power.

III. Empirical Results

Table I reports the results of the empirical analysis. Panel A contains the test statistics for three subperiods using 5, 10, and 19 cross-sectional assets and monthly returns. Corresponding results using overlapping annual returns are provided in Panel B. Empirical p-values and, for purposes of comparison with past research, significance based on the F-distribution are reported. Using monthly returns, in all but the first subperiod and 19 assets, the null hypothesis of the Sharpe-Lintner CAPM is rejected at p-value of 0.05 based on both the F-statistic and empirical distribution. The overall significance level, calculated as in Shanken (1985) by determining the standard-normal z corresponding to the significance level in each subperiod, is less than 0.01 for 5, 10, and 19 cross-sectional assets.

The empirical significance levels when annual returns are used are dramatically different from those using monthly return data. The significance level in each subperiod and using 5, 10, and 19 assets is 0.16 or higher. The overall significance levels are 0.382, 0.172, and 0.156 using 5, 10, and 19 assets.[4] Thus, we fail to reject the null hypothesis that the risk-return relation is linear, using annual overlapping return data and an empirical significance level. Of course, low power could be an issue, which we next explore.

[4] p-Values based on the F-statistics indicate reliable rejection of the CAPM. Differences between the empirical and F-statistic-based p-values suggest the latter overstate significance levels.

Table I

Multivariate Tests of the Sharpe-Lintner CAPM: Monthly versus Overlapping Annual Market-Value-Ranked Portfolio Returns on NYSE-AMEX Stocks from 1926 to 1988

All the NYSE-AMEX securities having market value and return data available for at least one year on the CRSP monthly return tape from 1926 to 1988 are included in the analysis. Twenty market-value-ranked portfolios are formed each year by ranking all the available securities on their beginning-of-the-year market values of equity and assigning them in equal numbers to 20 portfolios. MV1 is the portfolio of the smallest 5 percent of firms and MV20 is the portfolio of the largest 5 percent of the firms. Tests are performed using either 19 market-value-ranked portfolios (MV10 is excluded), or 5 or 10 portfolios constructed using the 19 portfolios. Monthly overlapping observations of annual returns are used. An equally weighted portfolio consisting of the 20 market-value-ranked portfolios is the market portfolio proxy.

The following test statistic is calculated:

$$\theta_1 = K[1 + \mu_m^2/\sigma_m^2]^{-1} \hat{\alpha}' S^{-1} \hat{\alpha}$$

where $K = (T - N - 1)T/(T - 2)N$, μ_m and σ_m^2 are the sample mean and variance of the excess return on the market over T periods, and $\hat{\alpha}$ is the vector of estimated intercepts from the market model regressions, and S is the sample covariance matrix of the N times series OLS market model residuals (see equations (3) and (4) in the text). The p-value is calculated assuming the test-statistic is distributed F. The empirical p-value is based on an empirical distribution of the test statistic generated via one-thousand simulations in which the null hypothesis is imposed.

Results for the overall period are based on aggregating the results for the three subperiods.

Number of Portfolios	Test Statistic p-Value Empirical p-Value Subperiod 1	Subperiod 2	Subperiod 3	Overall p-Value Emp. p Value
	Panel A: Monthly Returns			
5	2.270	2.099	2.532	
	0.015	0.025	0.006	0.000
	0.046	0.047	0.026	0.001
10	2.277	2.741	2.548	
	0.014	0.003	0.006	0.000
	0.015	0.002	0.003	0.000
19	1.455	1.888	1.955	
	0.103	0.016	0.012	0.006
	0.045	0.007	0.002	0.000
	Panel B: Overlapping Annual Returns			
5	3.809	1.304	0.668	
	0.000	0.229	0.733	0.001
	0.285	0.308	0.707	0.382
10	4.331	0.993	2.861	
	0.000	0.450	0.002	0.000
	0.220	0.541	0.164	0.172
19	9.035	3.920	6.652	
	0.000	0.000	0.000	0.000
	0.227	0.322	0.294	0.156

Table II
Power of the Multivariate Tests of CAPM: Empirical Rejection Frequencies Based on 200 Simulations Using Annual Overlapping Return Data on 5, 10, and 19 Market-Value-Ranked Portfolios over Three Subperiods

All the NYSE-AMEX securities having market value and return data available for at least one year on the CRSP monthly return tape from 1926 to 1988 are included in the analysis. Twenty market-value-ranked portfolios are formed each year by ranking all the available securities on their beginning-of-the-year market values of equity and assigning them in equal numbers to 20 portfolios. MV1 is the portfolio of the smallest 5 percent of firms and MV20 is the portfolio of the largest 5 percent of the firms. Two hundred simulations are performed in each subperiod using either 19 market-value-ranked portfolios (MV10 is excluded) or 5 or 10 portfolios constructed using the 19 portfolios. Monthly overlapping observations of annual returns are used. An equally weighted portfolio consisting of the 20 market-value-ranked portfolios is the market-portfolio proxy.

γ_0 is the deviation of the Treasury bill rate from the true risk-free rate.

The cell frequencies are the fraction of the 200 simulations for which the test statistic below exceeded the 95th percentile of the empirical distribution of the test statistic generated by imposing the null hypothesis:

$$\theta_1 = K[1 + \mu_m^2/\sigma_m^2]^{-1}\hat{\alpha}'S^{-1}\hat{\alpha}$$

where $K = (T - N - 1)T/(T - 2)N$, μ_m and σ_m^2 are the sample mean and variance of the excess return on the market over T periods, and $\hat{\alpha}$ is market model residuals (see equations (3) and (4) in the text).

	5 portfolios			10 portfolios			19 portfolios		
γ_0	Subperiod 1	Subperiod 2	Subperiod 3	Subperiod 1	Subperiod 2	Subperiod 3	Subperiod 1	Subperiod 2	Subperiod 3
0.00	0.05	0.65	0.90	0.04	0.86	0.01	0.05	0.04	0.04
0.02	0.22	0.93	1.00	0.10	1.00	0.66	0.07	0.10	0.06
0.04	0.76	1.00	1.00	0.55	1.00	0.99	0.33	0.38	0.26
0.06	0.95	1.00	1.00	0.93	1.00	1.00	0.74	0.80	0.79
0.08	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	0.98
0.10	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00

We assess the power of the tests by allowing the annual Treasury bill rate to deviate from the true risk-free rate by zero through 10 percent in increments of 2 percent. For each level of deviation, we perform 200 simulations in each subperiod and using 5, 10, and 19 assets. The test statistic in equation (4) is calculated for each simulation and compared against the 95th percentile of the empirical distribution of the test statistic generated by imposing the null hypothesis. The frequency with which the null is rejected at five percent significance level, based on 200 simulations in each subperiod, is reported in Table II.

The test is powerful using 5 and 10 portfolios. For example, for 5 portfolios a deviation of the true risk-free rate from the Treasury bill rate of more than 2 percent per year leads to a rejection of the null hypothesis 76 or higher percent of the time and for 10 portfolios the corresponding rejection rate is 55 percent. Deviations in excess of 2 percent per year are detected very frequently. In contrast, power using 19 portfolios is quite low, unless the true risk-free rate substantially deviates from the Treasury bill rate. MacKinlay (1987) also finds that the power is greater using fewer cross-sectional assets. Overall, the results in Table II suggest that the failure to reject the CAPM using overlapping annual returns and ten or five assets is unlikely to be due to low power of the test. Low power could, however, be the reason for not rejecting the null when we use 19 assets.

In summary, the results in this study show that the inferences from multivariate tests of the risk-return relation are sensitive to the return measurement interval, possibly because betas are not invariant to the return measurement interval and the investment horizon of a marginal investor is longer than a month.

REFERENCES

Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373–399.

——— and R. Clarida, 1987, The term structure of Euromarket interest rates: An empirical investigation, *Journal of Monetary Economics* 19, 25–44.

Chan, K. C., and Nai-Fu Chen, 1988, An unconditional asset pricing test and the role of firm size as an instrumental variable for risk, *Journal of Finance* 43, 309–325.

Fama, Eugene F., and Kenneth R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

Fama, Eugene F., and James MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Ferson, Wayne E., and Campbell R. Harvey, 1990, The variation of economic risk premiums, *Journal of Political Economy* 99, 385–415.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Gibbons, Michael R., 1982, Multivariate tests of financial models: A new approach, *Journal of Financial Economics* 10, 3–27.

———, Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Handa, Puneet, S. P. Kothari, and Charles E. Wasley, 1989, The relation between return interval and betas: Implications for the size-effect, *Journal of Financial Economics* 23, 79–100.

Hodrick, Robert J., and Sanjay Srivastava, 1984, An investigation of risk and return in forward foreign exchange, *Journal of International Money and Finance* 3, 5–29.

Levhari, David, and Haim Levy, 1977, The capital asset pricing model and the investment horizon, *Review of Economics and Statistics* 59, 92–104.

Lewis, Karen K., 1988, The behavior of Eurocurrency returns across different holding periods and monetary regimes, *Journal of Finance* 45, 1211–1236.

Lo, Andrew, and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreaction? *Review of Financial Studies* 3, 175–205.

MacKinlay, A. Craig, 1987, On multivariate tests of the CAPM, *Journal of Financial Economics* 18, 341–371.

Miller, Merton H., and Myron S. Scholes, 1972, Rates of return in relation to risk: A reexamination of some recent findings, in M. C. Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger, New York).

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Poterba, James M., and Lawrence H. Summers, 1988, Mean reversion in stock prices: Evidence and implications, *Journal of Financial Economics* 22, 27–60.

Shanken, Jay, 1985, Multivariate tests of the zero-beta CAPM, *Journal of Financial Economics* 14, 327–348.

Stambaugh, Robert F., 1982, On the exclusion of assets from tests of the two parameter model, *Journal of Financial Economics* 10, 235–268.