Journal of Financial Economics 10 (1982) 237-268. North-Holland Publishing Company

ON THE EXCLUSION OF ASSETS FROM TESTS OF THE TWO-PARAMETER MODEL

A Sensitivity Analysis

Robert F. STAMBAUGH*

University of Pennsylvania, Philadelphia, PA 19104, USA

Received August 1981, final version received May 1982

This study investigates the sensitivity of tests of the CAPM to different sets of asset returns. Tests are conducted with market portfolios that include returns for bonds, real estate, and consumer durables in addition to common stocks. Even when stocks represent only 10% of the portfolio's value, inferences about the CAPM are virtually identical to those obtained with a stocks-only portfolio. In contrast, inferences are sensitive to the set of assets used in the tests.

1. Introduction

Many discussions about the Capital Asset Pricing Model (CAPM) in recent years have centered on problems of excluding assets from tests of the model. Ross (1978, p. 891) states: 'Undoubtedly the most exciting work on the CAPM in recent years has occurred on the empirical front.' Much of this excitement is attributed to Roll (1977), who maintains that a 'correct and unambiguous' test of the model has not appeared. Roll suggests that such a test is unlikely, since it has to use all individual assets included in the market portfolio. The impact of Roll's critique comes not from the statement that assets are excluded; rather it comes from the contention that inferences about the model's validity are *sensitive* to incorrect specification of the market index portfolio.

That one index portfolio can reverse inferences about the model made with another index portfolio is certainly true and is not an empirical question. The empirical question is whether such a reversal occurs with indexes that approximate returns on portfolios of aggregate wealth. It is the latter question that bears on the testability of the CAPM. If inferences about mean-variance efficiency differ for alternative market indexes, then there is

*I wish to thank my Ph.D. dissertation committee — Eugene Fama (chairman), Craig Ansley, Robert Hamada, Jonathan Ingersoll, Merton Miller, Myron Scholes and Arnold Zellner — as well as Mike Gibbons, Pat Hess, Roger Ibbotson, Marc Reinganum, Bill Schwert and Jerry Warner for valuable assistance and comments. I am also grateful for comments provided by the referee. Any remaining errors are my responsibility.

0304-405x/82/0000-0000/$02.75

little justification for interpreting any such inferences as evidence about the CAPM. However, if alternative market indexes produce similar inferences about mean-variance efficiency (and the 'zero-beta' intercept), the justification is stronger for proceeding to draw inferences about the asset pricing theory.

This study constructs a number of market indexes and finds that they produce identical inferences about the CAPM. In addition to common stocks, the indexes include bonds, residential real estate, and consumer durables — all are major components of total wealth. Even when common stocks represent only 10% of the index's market value, inferences about the model are the same as those obtained with a stocks-only index.

The basis for the tests conducted here is the familiar linear relation between expected return and systematic risk implied by mean-variance efficiency of the market portfolio. That is, for any asset i,

$$E(r_i)=\gamma_1+\gamma_2\beta_i, \tag{1}$$

where $\beta_i=\operatorname{cov}(r_i,r_M)/\sigma^2(r_M)$, r_i is the return on asset i, and r_M is the return on the market portfolio. The relation in (1) is initially tested with a set of assets (the i's) composed of common stock portfolios, preferred stocks, and bond portfolios. The tests do not reject the hypotheses that the relation is linear and that the slope or 'risk premium' (γ_2) is positive, but the tests reject the equality of the intercept or 'zero-beta return' (γ_1) to a risk-free rate of interest. These inferences are obtained using any of the market indexes constructed. Thus, the indexes are found to be efficient but not equal to the Sharpe–Lintner tangency portfolio.

The tests of (1) are then repeated for alternative sets of assets. For example, the relation is also tested with common stock portfolios alone. The results of these tests, particularly the estimates of the zero-beta return and the risk premium (the γ's), often differ from results obtained with a broader set of assets that includes bonds and preferred stocks. In general, it is found that a test of the CAPM is more sensitive to the selection of assets [the i's in (1)] than to the composition of the market index.

Section 2 explains the construction of the market indexes and describes their statistical properties. Section 3 reviews the CAPM's hypotheses and develops some methodology for testing (1). Section 4 tests (1) with different market indexes, but one set of assets is used for all of the tests in order to focus on index specification. Section 5 tests (1) with alternative sets of assets.

2. Market indexes

Construction of a market index requires (i) rates of return on a broad range of assets and (ii) market values by which to weight these returns. Estimates of market values and monthly returns for seven classes of assets are assembled for the period from February 1953 through December 1976.[1] Section 2.1 presents estimates of market values and index weights. Section 2.2 discusses the potential for double-counting assets. Estimates of asset returns are analyzed individually in section 2.3 and then combined to form several market indexes in section 2.4.

2.1. Estimates of asset market values and asset weights

Table 1 displays estimates of year-end index weights and market values for the seven classes of assets: (1) NYSE common stocks, (2) corporate bonds, (3) U.S. Government bonds, (4) Treasury bills, (5) residential real estate (structures), (6) housefurnishings, and (7) automobiles. (Brief descriptions of the market value estimates are provided in the appendix.) The weights shown are those that each asset receives in an index combining all seven assets. The weight for NYSE common stocks ranges from 15% to 34% and averages about 25%. Consumer durables and real estate together represent about half

Table 1

Market index weights for selected years (end of year).

Assets	1952	1955	1958	1961	1964
NYSE common stocks	0.155	0.222	0.255	0.300	0.317
Corporate bonds	0.078	0.068	0.060	0.055	0.055
U.S. Government bonds	0.174	0.153	0.129	0.115	0.099
Treasury bills	0.030	0.025	0.028	0.034	0.038
Residential real estate	0.384	0.364	0.360	0.345	0.341
Housefurnishings	0.121	0.109	0.106	0.097	0.096
Automobiles	0.058	0.059	0.062	0.054	0.054
Total value ($ billions)	721.4	903.6	1080.9	1289.9	1470.2

Assets	1967	1970	1973	1976
NYSE common stocks	0.339	0.294	0.263	0.240
Corporate bonds	0.047	0.050	0.057	0.052
U.S. Government bonds	0.083	0.074	0.061	0.073
Treasury bills	0.039	0.042	0.042	0.047
Residential real estate	0.331	0.361	0.394	0.399
Housefurnishings	0.104	0.116	0.123	0.130
Automobiles	0.057	0.063	0.060	0.059
Total value ($ billions)	1735.8	2068.3	2571.5	3408.1

[1]Generally speaking, the lack of return data, particularly monthly, prevents an asset's inclusion in a market index. The task of obtaining market values leads to the literature on national wealth estimation, which stems from the seminal work of Raymond Goldsmith (1952, 1955). Revised and expanded estimates by numerous researchers have since appeared. Kendrick and Lee (1976) provide a concise history of work in this area. Even estimates of human capital have been published [see Kendrick (1974)].

of the index. Thus it appears that, in terms of asset values, an index constructed from these seven assets is significantly broader than an index composed solely of stocks or of stocks and bonds.

2.2. Double-counting

An attempt to construct a portfolio of aggregate wealth must consider the possibility of 'double-counting', i.e., including multiple financial claims on the same underlying assets. For most developed countries, the value of outstanding financial claims, including those held by intermediaries, exceeds estimates of real physical wealth (excluding human capital). Goldsmith (1969) defines a 'financial inter-relation ratio' (FIR) as the ratio of all financial assets to all physical assets. His estimates of the FIR in 1963 are about 1.3 for the United States and 1.7 for Great Britain and Japan. One can imagine how a naive assortment of assets and financial claims could double-count certain forms of wealth, e.g., real estate and real estate mortgages.

The issue of double-counting is most likely to arise for government securities.[2] Barro (1974) shows that an increase in the amount of government debt will generally not result in an increase in perceived (aggregate) household wealth and, thereby, an increase in desired (aggregate) consumption.[3] An increase in government debt is simply offset by implied future tax liabilities, both at the personal and corporate levels. However, inclusion of government debt in a market portfolio is consistent with this argument. The problem is one of accounting. *Market values* of the assets (claims) in table 1 reflect expected future tax liabilities. For example, the corporate sector's assets are priced net of any expected tax payments attributable to government debt repayment. In other words, the fraction of corporate assets that produces the (expected) future tax payments is not counted by the 'common stock' column. Also, the portion of government debt that is (expected) to be repayed by personal taxes on labor income is not counted elsewhere in table 1, since no other claims on human capital are included.

The total value of outstanding NYSE common stocks also double-counts assets to the extent that firms on the Exchange hold each other's shares. Estimates of interfirm holdings for all domestic stocks are included in an S.E.C. study by Tri (1971) and in an appendix by Eilbott (1973) to the *Institutional Investors Study*. Both studies estimate that 15% to 20% of all domestic stock is held by corporations. However, the estimated market value of other domestic stock (other exchanges and OTC) is roughly equal to the value of interfirm holdings. As a result, the total market value of traded

[2]For example, see Sharpe (1978).

[3]The statement requires the existence of an interior solution for the amount of bequest or intergenerational transfer.

common stock, once double-counting is eliminated, only slightly exceeds the reported value of the NYSE. Therefore, the return on NYSE common stocks is weighted by the total reported value of the Exchange. It is recognized that this both double-counts some NYSE stocks and excludes other non-NYSE stocks.

2.3. Asset return series

Monthly rates of return for the seven assets are obtained as follows:

(1) *NYSE common stocks.* The value-weighted monthly return index is compiled by the Center for Research in Security Prices (CRSP) at the University of Chicago.

(2) *Corporate bonds.* This series is constructed by Ibbotson and Sinquefield (1976) from the Salomon Brothers' High-Grade Long-Term Index (beginning in 1969) and from other Salomon Brothers' yield data (from 1946 to 1968).

(3) and (4). *U.S. Government bonds and Treasury bills.* These series are constructed by Ibbotson and Sinquefield (1976) from the CRSP U.S. Government securities file. The bond index approximates the total return on a 20-year maturity portfolio. The Treasury bill index consists of the return on the shortest-term bill outstanding with at least a one-month maturity.[4]

(5) *Residential real estate.* The return is estimated as the percent change in the home purchase component of the U.S. Consumer Price Index (CPI).[5] The home purchase index is based on transactions during the most recent three months for homes with F.H.A.-insured mortgages. The index also contains various lags: the Bureau of Labor Statistics receives the transactions up to one month after they are recorded at the F.H.A., and the price is recorded when the home is insured.[6] Nonetheless, as Fama and Schwert (1977, p. 119) conclude, the series 'seems to be the best available quality adjusted index of transaction prices for real estate'.

(6) *Housefurnishings.* This series is the percent change in the CPI housefurnishings component, which includes prices of textile housefurnishings, furniture, floor coverings, appliances, and other miscellaneous housefurnishings.

(7) *Automobiles.* This series is the percent change in the CPI used-automobiles component, which reflects prices of Chevrolet and Ford cars two to five years old.

[4]Use of the single Ibbotson–Sinquefield bond index was initially motivated by its availability. Given the results in section 4, further refinement to account for various maturities would surely produce negligible changes in results.

[5]Details of the CPI's components appear in U.S. Bureau of Labor Statistics (1966).

[6]For additional discussion of the housing indexes, see Housing Costs in the Consumer Price Index (1956).

The last three CPI return series require a simplifying assumption about the true underlying returns. Changes in the price index, at best, reflect only the capital gain on the asset. Price changes exclude the portion of the total return due to the flow of rental services (net of depreciation and other costs) provided by the asset. It is assumed that these net rental services represent the same percentage of the asset's value each period, i.e., the (real) net rental portion of the return is constant. This is proposed as a plausible characteristic of most durables and real estate, at least in the aggregate. Fama and Schwert (1977) essentially make this assumption for real estate. Thus, if these price series are used to estimate returns, the return on a market index including these estimates differs from the total return on an actual portfolio by a constant. However, the index return's variance and covariance with other returns is unaffected by exclusion of the constant. The latter property allows a test of the CAPM (or mean-variance efficiency), but the test must preserve the zero-beta return and the risk premium [γ_1 and γ_2 in (1)] as separate parameters. (This point is discussed in section 3.)

The three CPI series are also subject to limitations common to many CPI components. A comprehensive revision of the CPI, completed in 1953, substantially increased the scope of the index. Prior to January 1953, the home purchase index is unavailable and the used automobiles index is available only quarterly. Data for the non-CPI series are available before 1953, but a sensitivity analysis that excludes consumer durables and, especially, real estate is less interesting. The CPI is also contaminated by the price controls imposed in August 1971 and phased out during 1973 and 1974.

The tests use returns from the period February 1953 to December 1976, and this total period is divided into four subperiods: (1) 2/53–3/59, (2) 4/59–5/65, (3) 6/65–7/71, and (4) 8/71–12/76. The subperiods are chosen to be roughly equal in length and to confine the period of price controls to the last subperiod. Real returns are used throughout the study.[7] Table 2 displays estimates of means, standard deviations, and autocorrelations of returns on the seven assets.[8] A striking (not surprising) feature of the statistics is the relatively high standard deviation of common stock returns — at least twice as high as those of the next highest series (bonds and automobiles) and up to twenty-five times the standard deviation of T-bill returns.

2.4. The market indexes

The weights in table 1 are used to combine the return series discussed

[7]The total CPI is used to construct real returns as defined by Fama (1976, pp. 172–173). Most of the tests were also conducted with nominal returns, and inferences were virtually unchanged.

[8]Similar statistics for nominal returns on many of these assets are also reported by Fama and Schwert (1977).

Table 2

Index component returns: Sample means, standard deviations, and first three autocorrelations.[a]

Asset	Mean[b]	Standard deviation[b]	Autocorrelations[c] ρ_1	ρ_2	ρ_3
2/53–3/59					
NYSE stocks	1.312	3.311	0.11	0.11	0.09
Corporate bonds	0.021	1.449	0.36	0.06	−0.02
Government bonds	−0.036	1.561	0.20	−0.01	0.00
Treasury bills	0.047	0.233	0.09	0.17	−0.02
Housefurnishings	−0.162	0.522	−0.21	0.01	−0.13
Automobiles	−0.232	1.655	0.25	0.32	0.28
Real estate	−0.037	0.366	0.16	−0.12	−0.24
4/59–5/65					
NYSE stocks	0.839	3.433	0.10	−0.10	0.07
Corporate bonds	0.269	0.823	0.19	−0.22	−0.19
Government bonds	0.217	1.021	−0.11	0.04	−0.29
Treasury bills	0.129	0.167	−0.04	0.01	−0.19
Housefurnishings	−0.131	0.274	−0.18	−0.22	−0.02
Automobiles	0.100	1.988	0.10	0.31	−0.11
Real estate	−0.032	0.303	0.09	−0.10	−0.21
6/65–7/71					
NYSE stocks	0.156	4.114	0.13	0.01	−0.09
Corporate bonds	−0.251	2.199	0.09	0.00	−0.14
Government bonds	−0.269	2.632	−0.16	0.11	−0.18
Treasury bills	0.074	0.167	0.12	0.01	−0.06
Housefurnishings	−0.127	0.224	0.14	0.06	−0.11
Automobiles	−0.150	2.070	0.35	−0.24	−0.32
Real estate	0.006	0.355	0.12	0.10	−0.06
8/71–12/76					
NYSE stocks	0.042	5.133	0.04	0.01	0.24
Corporate bonds	0.175	2.356	0.12	−0.01	0.15
Government bonds	0.104	2.050	0.17	−0.01	0.03
Treasury bills	−0.079	0.271	0.03	0.19	0.17
Housefurnishings	−0.105	0.400	0.31	0.17	0.01
Automobiles	0.173	2.543	0.77	0.37	−0.05
Real estate	−0.065	0.433	0.67	0.45	0.27

[a]Monthly real returns.
[b]Values multiplied by 100.
[c]Standard error is approximately 0.12 if true autocorrelations are zero at all lags.

above to form four market indexes, which are numbered 1 through 4. Index no. 1 consists of the value-weighted NYSE common stock series. Index no. 2 combines index no. 1 with corporate bonds, government bonds, and T-bills according to relative market values. (Monthly market values are obtained from year-end values by linear interpolation.) Index no. 3 adds housefurnishing, automobiles, and real estate to index no. 2. As noted above, common stocks average about 25% of index no. 3 when weights reflect estimated market values. (See table 1.) In order to extend the sensitivity analysis and partially allow for the exclusion of other assets, index no. 4 assigns the value-weighted common stock return a weight of 0.10 and the remaining six returns in index no. 3 a weight of 0.90 (with the relative weights among the latter six returns determined by their market values).

Table 3

Four market indexes: Sample means, standard deviations, autocorrelations, and correlations between indexes.[a]

			Autocorrelations[d]			Correlations between indexes		
Index[b]	Mean[c]	Standard deviation[c]	ρ_1	ρ_2	ρ_3	no. 2	no. 3	no. 4
2/53–3/59								
No. 1	1.312	3.311	0.11	0.11	0.09	0.901	0.867	0.514
No. 2	0.604	1.532	0.20	0.13	0.12		0.957	0.760
No. 3	0.229	0.695	0.19	0.12	0.15			0.856
No. 4	0.082	0.474	0.19	0.15	0.11			
4/59–5/65								
No. 1	0.829	3.433	0.10	−0.10	0.07	0.986	0.970	0.751
No. 2	0.571	1.981	0.11	−0.15	0.07		0.987	0.823
No. 3	0.260	0.969	0.11	−0.19	0.12			0.885
No. 4	0.116	0.433	0.06	−0.29	0.04			
6/65–7/71								
No. 1	0.156	4.114	0.13	0.01	−0.09	0.983	0.979	0.812
No. 2	0.032	3.010	0.12	0.02	−0.10		0.993	0.891
No. 3	−0.002	1.451	0.09	0.03	−0.07			0.913
No. 4	−0.053	0.692	0.01	0.06	−0.03			
8/71–12/76								
No. 1	0.042	5.133	0.04	0.01	0.24	0.986	0.966	0.878
No. 2	0.036	3.365	0.10	−0.02	0.25		0.980	0.914
No. 3	−0.009	1.396	0.14	0.04	0.29			0.950
No. 4	−0.012	0.798	0.26	0.11	0.36			

[a]Monthly real returns.
[b]Index no. 1: NYSE common stocks (value-weighted).
Index no. 2: no. 1 plus corporate and Government bonds and Treasury bills.
Index no. 3: no. 2 plus real estate, housefurnishings, and automobiles.
Index no. 4: same as no. 3 but with NYSE weighted by 0.10.
[c]Values multiplied by 100.
[d]Standard error is approximately 0.12 if true autocorrelations are zero at all lags.

Table 3 displays descriptive statistics for indexes no. 1 through no. 4. The high variance of stock returns is apparent in several of the statistics. First observe that the standard deviation of the return on the index falls dramatically as stocks receive less weight.[9] In addition, the correlations between indexes are generally quite high, although the correlations between indexes no. 1 and no. 4 range from 0.514 to 0.878. Such high correlations are also consistent with the dominant effect of common stocks. These numbers suggest that, as long as common stocks are included, one can construct an index that is highly correlated with the market portfolio, but it is difficult to capture the market's true mean and standard deviation.

As Roll (1977, p. 130) warns, though, *high correlations can be deceiving*:

> . . . most reasonable proxies will be very highly correlated with each other and with the true market whether or not they are mean-variance efficient. This high correlation will make it seem that the exact composition is unimportant, whereas it can cause quite different inferences.

Roll's caveat seems well-suited to the indexes constructed here. Section 4 investigates the extent to which these different indexes do, in fact, cause different inferences.

3. Testing the CAPM: Methodology

This section develops procedures that deal with some of the statistical problems in testing the CAPM. The major assumptions are that (1) monthly asset returns obey a multivariate normal distribution and (2) the parameters of this distribution are stationary over subperiods of T months, where T is typically 74 (about six years). These assumptions are also made in previous tests of the CAPM.[10]

[9]In three of the four periods, the mean also declines, but the means for indexes no. 3 and no. 4 must be viewed cautiously. Those indexes contain the three CPI components (Autos, Housefurnishings, and Real Estate) whose measured returns reflect only capital gains.

[10]Fama (1976) performs extensive diagnostics with common stock returns and concludes that 'the normal distribution is a good working approximation for monthly security and portfolio returns in the post-World War II period' (pp. 33–35). Blattberg and Gonedes (1974) find that daily returns are better represented as Student t than stable, which implies that monthly returns are approximately normal. Also see Blume (1968) and Officer (1971) for evidence supporting monthly return normality.

Black, Jensen and Scholes (1972) assume stationarity over five-year periods; Fama and MacBeth (1973) use five- to eight-year periods. Gonedes (1973) examines prediction errors from market model regressions and concludes that a seven-year estimation period produces the lowest mean-squared and absolute prediction errors. Recursive residuals tests for stationarity of the regressions used in this study are reported in Stambaugh (1981), and the results support the stationarity assumptions made here.

3.1. The model's testable hypotheses

This study tests three hypotheses implied by the CAPM:

H1: *Linearity.* Expected return is linearly related to beta. That is, for some γ_1 and γ_2,

$$\boldsymbol{\mu} = \gamma_1 \boldsymbol{\iota}_K + \gamma_2 \boldsymbol{\beta}, \tag{2}$$

where $\boldsymbol{\mu}$ is a K-vector of expected returns, $\boldsymbol{\beta}$ is a K-vector with typical element β_i, and $\boldsymbol{\iota}_K$ is a unit K-vector.

H2: $\gamma_1 = r_F$. The intercept, γ_1, is equal to a risk-free rate of interest.

H3: $\gamma_2 > 0$. The risk premium, γ_2, is positive.

Hypotheses H1 and H3 together are equivalent to the condition that the market portfolio lies on the positively-sloped portion of the minimum-variance boundary. Hypothesis H2 is implied by the original Sharpe–Lintner version of the model. If H2 is false, but H1 and H3 are true, then γ_1 is the expected return on an asset whose return is uncorrelated with the market return (the Black version of the model).[11]

Mean-variance efficiency also implies $\gamma_2 = \mu_M - \gamma_1$, where μ_M is the expected return on the market portfolio. However, as this section later demonstrates, that additional restriction is not testable with the broader market indexes (no. 3 and no. 4) due to the *a priori* error in their mean returns noted in section 2.3. Therefore, the tests are confined to the first three hypotheses.

3.2. The basic inference problem

Tests of the model are complicated by the fact that neither $\boldsymbol{\mu}$ nor $\boldsymbol{\beta}$ is directly observable. For example, tests of linearity essentially ask whether a set of estimated points, $\{(\hat{\beta}_i, \hat{\mu}_i)\}$, is 'likely' given the underlying restriction in (2). Previous studies, such as Fama and MacBeth (1973), condition their tests on prior estimates of $\boldsymbol{\beta}$. Fama and MacBeth's tests for nonlinearity also assume that violations of linearity are related to β_i^2 or $\sigma(\varepsilon_i)$ (standard deviation of the market-model disturbance). MacBeth (1975) employs a multivariate T^2 test of linearity that no longer assumes violations of linearity are related to β_i^2 or $\sigma(\varepsilon_i)$, but he still conditions on prior estimates of $\boldsymbol{\beta}$ as well as γ_1 and γ_2. Most previous tests of H2 and H3 have also conditioned on $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$ in obtaining estimates and standard errors of γ_1 and γ_2.[12]

[11]See Sharpe (1964), Lintner (1965), and Black (1972). Also see Fama (1976, ch. 8) for a discussion of these models.

[12]Litzenberger and Ramaswamy (1979) develop errors-in-the-variables estimators for γ_1 and γ_2 that do not condition on $\hat{\boldsymbol{\beta}}$ but condition instead on the variance of the estimation error in $\hat{\boldsymbol{\beta}}$.

The tests conducted here are not conditioned on the assumption that some of the model's parameters (e.g., $\boldsymbol{\beta}$) are equal to previously calculated estimates. Rather, the tests formally acknowledge that *none* of the model's parameters can be estimated without error.

3.3. Defining the framework: Parameter restrictions

Let $\boldsymbol{r}_i' = (r_{i1}, r_{i2}, \ldots, r_{iT})$ denote a T-vector of returns on the ith asset generated independently over T time periods. The tests include K such assets. Similarly, let $\boldsymbol{r}_M$ denote a T-vector of returns on the market index. Multivariate normality implies K 'market-model' regression equations,

$$\boldsymbol{r}_i = \boldsymbol{\iota}_T \alpha_i + \boldsymbol{r}_M \beta_i + \boldsymbol{\varepsilon}_i, \qquad i = 1, \ldots, K, \tag{3}$$

where $\boldsymbol{\iota}_T$ is a unit T-vector, and $\boldsymbol{\varepsilon}_i$ is a T-vector of regression disturbances. The K equations in (3) are written more compactly as[13]

$$\underset{TK\times 1}{\boldsymbol{r}} = \Big(\underset{TK\times K}{I \otimes \boldsymbol{\iota}_T}\Big)\boldsymbol{\alpha} + \Big(\underset{TK\times K}{I \otimes \boldsymbol{r}_M}\Big)\underset{K\times 1}{\boldsymbol{\beta}} + \underset{TK\times 1}{\boldsymbol{\varepsilon}}, \tag{4}$$

where $\boldsymbol{r}' = (\boldsymbol{r}_1', \ldots, \boldsymbol{r}_K')$, $\boldsymbol{\varepsilon}' = (\boldsymbol{\varepsilon}_1', \ldots, \boldsymbol{\varepsilon}_K')$, and $\boldsymbol{\alpha}' = (\alpha_1, \ldots, \alpha_K)$. Also,

$$E\{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\} = \underset{K\times K}{\Sigma} \otimes \underset{T\times T}{\boldsymbol{I}}, \tag{5}$$

where Σ is a full matrix with (i,j) element $\sigma_{ij} = \operatorname{cov}(\varepsilon_{it}, \varepsilon_{jt})$. Assume that the K individual assets are selected so that no linear combination of $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_K$ yields $\boldsymbol{r}_M$. (This prevents a singular Σ.)

The linear CAPM relation in (2) places restrictions on the market-model parameters. First note that the intercept vector, $\boldsymbol{\alpha}$, consists of

$$\boldsymbol{\alpha} = \boldsymbol{\mu} - \mu_M \boldsymbol{\beta}. \tag{6}$$

The linear pricing restriction in (2) is substituted for $\boldsymbol{\mu}$ in (6) to yield

$$\boldsymbol{\alpha} = \gamma_1 \boldsymbol{\iota}_K + (\gamma_2 - \mu_M)\boldsymbol{\beta}, \tag{7}$$

which is rewritten slightly as

$$\boldsymbol{\alpha} = \gamma_1 \boldsymbol{\iota}_K + \gamma_2^* \boldsymbol{\beta}, \tag{8}$$

where $\gamma_2^* = \gamma_2 - \mu_M$. Thus, the linear relation of the CAPM imposes $K-2$

[13]The symbol '$\otimes$' denotes a Kronecker product of two matrices. Theil (1971, pp. 303–306) provides a definition and some useful properties of the Kronecker product.

non-redundant restrictions on the parameters of the market-model equations. [Two of the K equations in (8) can be solved for γ_1 and γ_2^*, and the solution can be substituted in the remaining $K-2$ equations.]

Observe that the risk premium, γ_2, can be eliminated from the parameter set, because the pricing relation also requires $\gamma_2=\mu_M-\gamma_1$ (i.e., $\gamma_2^*=\gamma_1$). This additional restriction reduces (7) to

$$\boldsymbol{\alpha}=\gamma_1(\boldsymbol{\iota}_K-\boldsymbol{\beta}). \tag{9}$$

Eq. (9) underlies the estimates of a time series for the zero-beta factor by Black, Jensen and Scholes (1972). More recently, Gibbons (1982) performs a likelihood ratio test of (9) and computes maximum likelihood estimates of γ_1 using the multivariate regression framework in (4). The tests here preserve γ_2 (or γ_2^*) as a separate parameter due to the known error in the returns on the broader indexes. Assume as in section 2.3 that the reported index return used in the tests differs from the total return on an actual portfolio of the same assets by a constant. That is,

$$r_{Pt}=r_{It}+c, \qquad t=1,\ldots,T, \tag{10}$$

where r_{Pt} is the total return on the portfolio, with mean μ_P, and r_{It} is the index return, with mean μ_I.[14] Assume the *portfolio*, P, is mean-variance efficient, so that

$$\boldsymbol{\mu}=\gamma_1\boldsymbol{\iota}_K+(\mu_P-\gamma_1)\boldsymbol{\beta}_P, \tag{11}$$

where the ith element of $\boldsymbol{\beta}_P$ is $\beta_{iP}=\operatorname{cov}(r_i,r_P)/\sigma^2(r_P)$. The market-model equations in (4) are specified in terms of the *index*, I, so (6) becomes

$$\boldsymbol{\alpha}=\boldsymbol{\mu}-\mu_I\boldsymbol{\beta}_I=\boldsymbol{\mu}-\mu_I\boldsymbol{\beta}_P, \tag{12}$$

where the second equality follows from $\boldsymbol{\beta}_I=\boldsymbol{\beta}_P$, given the assumption in (10). Substitution of (11) for $\boldsymbol{\mu}$ in (12) yields

$$\boldsymbol{\alpha}=\gamma_1\boldsymbol{\iota}_K+(\mu_P-\gamma_1-\mu_I)\boldsymbol{\beta}_P=\gamma_1\boldsymbol{\iota}_K+(\gamma_2-\mu_I)\boldsymbol{\beta}_P=\gamma_1\boldsymbol{\iota}_K+\gamma_2^*\boldsymbol{\beta}_P, \tag{13}$$

which is identical to (8) where $\gamma_2=\mu_P-\gamma_1$ and $\gamma_2^*=\gamma_2-\mu_I$. However, (13) simplifies to (9) only if $\mu_P=\mu_I$, i.e., $c=0$. If it is known *a priori* that $c\neq 0$, then the restriction in (9) is inappropriate.[15]

[14]Recall that in the broader indexes, c arises from the left-out rental return on real estate and durables.

[15]This issue is investigated empirically in Stambaugh (1982), where tests of (8) and (9) are conducted with both stocks-only indexes and broader indexes.

3.4. Maximum likelihood estimators of γ_1 and γ_2

Estimates of the zero-beta return (γ_1) and the risk premium (γ_2) are meaningful only if linearity is satisfied. However, the tests of linearity are more easily discussed after defining estimators for the γ's.

Let L_R denote the likelihood function of the restricted parameters, which are collectively represented as the vector $\boldsymbol{\theta}$. In this case, $\boldsymbol{\theta}$ contains γ_1, γ_2^*, μ_M, σ_M^2 [$=\mathrm{var}(r_M)$], $\boldsymbol{\beta}$, and the $K(K+1)/2$ distinct elements of Σ. Maximum likelihood (ML) estimators, $\hat{\boldsymbol{\theta}}$, are obtained by maximizing L_R with respect to $\boldsymbol{\theta}$. Standard errors are obtained from the asymptotic covariance matrix, $R^{-1}(\boldsymbol{\theta})$, where

$$R(\boldsymbol{\theta}) = -E[\partial^2 \log L_R / \partial\boldsymbol{\theta}\partial\boldsymbol{\theta}']. \tag{14}$$

A consistent estimator of the asymptotic standard error is obtained by evaluating (14) at $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}$.[16] (The ML estimator of γ_2 is obtained by adding the estimator $\hat{\mu}_M$ to the estimator $\hat{\gamma}_2^*$.) The ML estimators of the γ's are computationally equivalent to the restricted nonlinear Aitken-type estimators derived by Gallant (1975).[17] He shows that the estimators possess strong consistency (almost sure convergence) rather than simple consistency (convergence in probability). Note that the standard errors of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ incorporate the uncertainty associated with estimates of *all* of the parameters, in particular $\boldsymbol{\beta}$. $R^{-1}(\boldsymbol{\theta})$ contains entries for variances and covariances among *all* of the parameter estimates.

Optimal asymptotic properties of ML estimators are well known.[18] However, finite-sample properties often depend on the specific application. The finite-sample properties of the estimators were investigated in Monte Carlo experiments with sample sizes and parameter values typical of those in tests of the CAPM. The results indicate that the estimators are approximately normally distributed with variances equal to the Cramer–Rao bound.[19]

3.5. Tests of linearity

A likelihood ratio (LR) test is a natural approach suggested by the above framework, but it is not the only way to use the likelihood function to test linearity and the resulting restriction in (8). Two alternatives are

[16] A technical appendix that details the likelihood function, first and second derivatives, and the asymptotic covariance matrix of $\hat{\gamma}_1$ and $\hat{\gamma}_2^*$ will be furnished upon request to the author.

[17] The quadratic form that is minimized here is the same as the expression minimized by Gallant (1975, p. 38, last equation).

[18] Shanken (1980) considers the asymptotic properties of several estimators for CAPM parameters, including the ML estimators developed here.

[19] See Stambaugh (1981, ch. 7) for details.

the Wald test and the Lagrangian multiplier (LM) test. Both tests are large-sample equivalents of the LR test in that their test statistics converge to the same asymptotic distribution.[20] However, the tests do not necessarily produce the same inferences in a finite sample.[21]

Monte Carlo tests that replicate sample characteristics encountered here reveal substantial differences in finite sample distributions of the LR, LM, and Wald tests.[22] The Wald test accepts the null hypothesis too often when critical values are based on the asymptotic distribution, and for this reason it is not discussed further. However, the LM test statistic conforms more closely to its asymptotic distribution than does the LR test. Experiments are conducted in which the number of market-model equations, K, takes values of five, ten, and fifteen. Both the LR and LM tests conform reasonably well to the asymptotic distribution when K equals five and ten. When K equals fifteen, the LM test continues to reject the null at the appropriate rate, but the rejection rate of the LR test is too high (e.g., 12% to 13% versus an asymptotic rate of 5%). This suggests a tendency for the LR test to reject the null hypothesis too often as the number of equations increases. Therefore, the linearity tests reported here are confined primarily to the LM test.

To construct the LM test, let $\boldsymbol{\theta}$ be the parameter vector containing $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, since they are the market-model parameters involved in the restriction in (8). Next evaluate the first derivatives of the *unrestricted* log-likelihood function, $\log L_u$, at the *restricted* maximum likelihood estimators, $\hat{\boldsymbol{\theta}}$. Define the vector $\boldsymbol{d}(\hat{\boldsymbol{\theta}})$ of length $2K$ as

$$\boldsymbol{d}(\hat{\boldsymbol{\theta}}) = [\partial \log L_u / \partial \theta_i]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \tag{15}$$

Define $R(\boldsymbol{\theta})$ as in (14) except that L_R is replaced by L_u. Then

$$\chi^2 = \boldsymbol{d}'(\hat{\boldsymbol{\theta}}) R^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{d}(\hat{\boldsymbol{\theta}}) \tag{16}$$

is asymptotically distributed $\chi^2(K-2)$ under the null hypothesis of linearity. Silvey (1975) calls this the 'χ^2 test', but it is originally called the 'Lagrangian multiplier test' in Aitchison and Silvey (1958) and Silvey (1959). If the null hypothesis is true, then the partial derivatives in (15) are 'close' to zero when evaluated at the restricted estimates. (They are exactly zero when evaluated at the unrestricted estimates.)

If linearity is violated, then for any γ_1 and γ_2,

$$\boldsymbol{\mu} = \gamma_1 \boldsymbol{\iota}_K + \gamma_2 \boldsymbol{\beta} + \boldsymbol{\xi}, \tag{17}$$

[20]See Silvey (1975, ch. 7) for a discussion of these tests.
[21]See, for example, Berndt and Savin (1977).
[22]Details appear in Stambaugh (1981, ch. 7).

where ξ is a K-vector with at least two different elements. The elements of ξ need not obey any relation. For example, rejection of linearity does not require that the elements of ξ are related to β_i^2 or $\sigma(\varepsilon_i)$.[23] Monte Carlo tests construct ξ as $\xi = hz$, where z contains values drawn independently from the standard normal distribution. This affords a convenient tabulation of the power function in terms of the scalar h. For example, with ten equations and $h=0.003$, the LM test rejects linearity at the 5% significance level in 83% of the samples. [A tabulation is reported in Stambaugh (1981).]

4. Tests using different market indexes

4.1. The initial set of market-model equations

This section tests the CAPM using the different market indexes described in section 2. The objective is to discover whether the different indexes yield different inferences about the model. In order to focus on this question, all of the tests use one set of assets. The tests developed in section 3 require a market-model equation [as in (3)] for each asset. Thus, this section changes the composition of the index (r_M) common to all market-model equations, but the equations contain the same asset returns (the r_i's) throughout. There are 28 assets selected: 19 common stock industry portfolios, 4 preferred stocks, and 5 bond portfolios. Preferred stocks and bonds are included to extend the range of asset types and parameter values beyond those typically encountered in tests of the CAPM. The objective is not to use a comprehensive sample of bonds and preferred stocks, but rather the aim is to begin investigating the sensitivity of tests to including such assets. As section 5 demonstrates, different inferences are often obtained with subsets of the 28 assets (such as the stock portfolios alone). Descriptions of the 28 assets are provided below.

Nineteen industry common stock portfolios are formed with the same method used by MacBeth (1975). The return on a portfolio is the arithmetic average of returns for firms on the NYSE with the appropriate S.E.C.-assigned two-digit industry code for the given month. The assignment of all NYSE stocks to portfolios each month is exhaustive (no unassigned stocks).[24] Returns are obtained from the CRSP Monthly Returns File, as are a majority of the two-digit industry codes.[25] Table 4 displays the number of securities in each portfolio at several dates, the S.E.C. codes assigned to each portfolio, and betas calculated with an equally-weighted NYSE index.

[23]Roll (1977) makes this criticism of Fama–MacBeth-type tests of linearity.

[24]MacBeth's portfolio no. 20 (miscellaneous) is omitted, and his original set of 20 portfolios is reduced to 19.

[25]Codes for approximately 450 firms are from the U.S. Securities and Exchange Commission, *Directory of Companies Required to File Annual Reports with the Securities and Exchange Commission*. Another 350 codes were obtained from a list compiled by CRSP.

Table 4

Industry portfolio S.E.C. codes, number of firms, and estimated betas.

Portfolio description	S.E.C. codes	Number of firms			Estimated betas[a]			
		12/53	12/65	12/76	2/53–3/59	4/59–5/65	6/65–7/71	8/71–12/76
1. Mining	10–14	43	42	52	1.19	0.87	0.97	0.71
2. Food and beverages	20	77	79	77	0.69	0.84	0.83	0.86
3. Textile and apparel	22, 23	41	47	54	1.12	1.01	1.13	1.15
4. Paper products	26	29	30	33	1.12	0.93	0.99	0.83
5. Chemical	28	69	85	89	0.99	1.13	0.88	0.79
6. Petroleum	29	40	34	30	1.05	0.76	0.74	0.59
7. Stone, clay, and glass	32	24	38	40	1.05	0.94	1.01	0.98
8. Primary metals	33	74	75	68	1.45	1.16	0.99	0.89
9. Fabricated metal	34	29	37	45	1.00	1.00	1.08	1.07
10. Machinery	35	74	101	103	1.16	1.12	1.13	1.06
11. Appliances, elec. equipment	36	44	80	83	1.17	1.33	1.30	1.22
12. Transportation equipment	37	83	63	57	1.15	1.04	1.14	1.09
13. Misc. manufacturing	38, 39	26	44	66	0.96	1.19	1.23	1.17
14. Railroads	40	54	37	16	1.41	1.17	1.12	0.77
15. Other transportation	41, 42, 44, 45, 47	32	33	40	1.20	1.08	1.34	1.14
16. Utilities	49	91	116	148	0.51	0.63	0.49	0.62
17. Department stores	53	31	29	29	0.80	0.88	1.05	1.18
18. Other retail trade	50–52, 54–59	52	76	117	0.71	0.99	0.99	1.16
19. Banking, finance, real estate	60–67	41	73	217	0.92	0.94	1.02	1.12

[a]Estimated using ordinary least squares with the equally-weighted NYSE index.

Portfolios are formed here primarily because they provide a convenient way to limit the computational dimensions of the methods described in section 3. (The limiting dimension is essentially that of a full contemporaneous covariance matrix of disturbances, Σ.) Industry portfolios also allow rejection of the CAPM due to the presence of additional industry-related variables in the pricing relation. That is, the elements of $\boldsymbol{\xi}$ in (17) may be related to industrial classification.[26]

Returns on non-callable preferred stocks for the 2/53–12/76 period are compiled for four firms: American Can, Ligget and Myers, Pacific Telephone and Telegraph, and Uniroyal.[27] The five bond portfolios consist of the Ibbotson–Sinquefield corporate and government long-term portfolios, described in section 2.3, and three portfolios with shorter maturities: (i) 1 to 2 years, (ii) 2 to 5 years, and (iii) 5 to 10 years. The monthly return on a portfolio is the arithmetic average of all returns on securities outstanding with the given remaining maturity. The return and maturity data are obtained from the CRSP U.S. Government Securities File. The methodology is the same as that of Bildersee (1975), except that his maturity classifications produce months in which there are no observations for some maturities. No 'empty cells' are encountered with the three classifications selected here. Bonds with special estate-tax features (flower bonds) are excluded.

4.2. Linearity tests

The results of a Lagrangian multiplier (LM) test of linearity are reported in table 5. None of the test statistics rejects linearity at conventional significance levels. The interesting aspect of the results is that the same inference occurs for all four market indexes.

The effects of changing the market index are seen by scanning down the columns of table 5. Test statistics and their corresponding p-values for each subperiod are displayed in the first four columns.[28] Each of the test statistics is asymptotically distributed as $\chi^2(26)$ under the null hypothesis, and the p-values are calculated according to that distribution. A test statistic for the overall period is constructed by observing that the subperiod statistics are

[26]There are also potential disadvantages of portfolio formation. As Roll (1979) notes, if non-zero ξ_i's in (17) exist for individual securities, and if these ξ_i's are distributed across securities in a manner independent of the portfolio classification scheme, then the ξ_i's cancel each other in the formation of portfolios. While this is true, it does not by itself imply that portfolio formation reduces the power of linearity tests. The reason is that $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$ in (17) are not directly observed, and these parameters, particularly $\boldsymbol{\beta}$, can be estimated more precisely for portfolios than for individual securities. There is a trade-off between parameter precision on one hand and preservation of individual deviations from linearity on the other. The net effect on the power of linearity tests has not been resolved.

[27]I am grateful to Pat Hess for providing a portion of these data.

[28]The p-value is the probability of the test statistic's exceeding the reported value if the null hypothesis is true. In general, 'small' p-values reject the null hypothesis.

Table 5

Lagrangian multiplier test of linearity (assets: 19 common stock industry portfolios, 4 preferred stocks, and 5 bond portfolios).

Market index	χ^2 statistic and p-value[a] 2/53–3/59	4/59–5/65	6/65–7/71	8/71–12/76	Overall χ^2 statistic and p-value[b]
1. NYSE (value-weighted)	26.60 (43.1%)	27.38 (39.0%)	17.41 (89.6%)	23.01 (63.3%)	94.39 (74.0%)
2. 1 plus Corporate bonds, Government bonds, and Treasury bills	27.74 (37.1%)	27.14 (40.2%)	17.67 (88.7%)	22.97 (63.5%)	95.53 (71.2%)
3. 2 plus Housefurnishings, Automobiles, and Real estate	25.86 (47.1%)	26.79 (42.0%)	17.67 (88.8%)	23.07 (62.9%)	93.39 (76.3%)
4. Same as 3 but with the NYSE weighted by 0.10	29.36 (29.5%)	26.35 (44.4%)	17.54 (89.2%)	23.30 (61.6%)	96.54 (68.6%)

[a]Test statistics are asymptotically distributed $\chi^2(26)$ under the null hypothesis of linearity; p-value in parentheses.

[b]The sum of the individual subperiod statistics, which is distributed $\chi^2(104)$ under the null hypothesis of linearity; p-value in parentheses.

independent χ^2 variates. Thus, the sum of the four subperiod statistics is asymptotically distributed $\chi^2(104)$ under the null hypothesis. This overall statistic and the associated p-value are shown in the fifth column. Values of the test statistics and p-values for both the overall period and the subperiods are virtually unchanged by altering the composition of the market index.

4.3. Estimates of γ_1 and γ_2

Maximum likelihood (ML) estimates of the zero-beta return and the risk premium are displayed in tables 6 and 7, respectively. The results find a positive risk premium but reject equality of the zero-beta return to the risk-free rate. These inferences, like those about linearity, are unaffected by choice of a market index.

The first four columns of table 6 contain ML estimates of the zero-beta return, γ_1. (Average real returns on one-month T-bills are reported in the last row for comparison.) A test of $\gamma_1 = r_F$ for the overall 2/53 to 12/76 period is constructed by observing that, under the null hypothesis, the coefficient estimate for each subperiod is approximately normally distributed with mean r_F and standard deviation equal to the estimated standard error. By first 'standardizing' each estimate by its mean and standard deviation, and then summing these standardized estimates over the subperiods, a single standard-

Table 6

Maximum likelihood estimates of the zero-beta return, γ_1 (assets: 19 common stock industry portfolios, 4 preferred stocks, and 5 bond portfolios).

Market index	Estimate of γ_1 and standard error[a]				Overall test statistic and p-value[b]
	2/53–3/59	4/59–5/65	6/65–7/71	8/71–12/76	
1. NYSE (value-weighted)	0.118	0.192	0.172	−0.007	4.82
	(0.026)	(0.022)	(0.036)	(0.054)	(0.0%)
2. 1 plus Corporate bonds, Government bonds, and Treasury bills	0.110	0.190	0.172	−0.010	4.59
	(0.025)	(0.022)	(0.036)	(0.055)	(0.0%)
3. 2 plus Housefurnishings, Automobiles and Real estate	0.099	0.187	0.172	−0.011	4.23
	(0.027)	(0.027)	(0.036)	(0.055)	(0.0%)
4. Same as 3 but with the NYSE weighted by 0.10	0.094	0.172	0.177	−0.010	3.83
	(0.026)	(0.024)	(0.036)	(0.055)	(0.2%)
	Average real T-bill return				
	0.047	0.129	0.074	−0.079	

[a]Values multiplied by 100; standard error in parentheses.

[b]A standard normal (0, 1) test statistic under the null hypothesis $\gamma_1 = r_F$, where r_F is the average real return on a one-month T-bill in the given subperiod. The p-value in parentheses reflects a two-tailed test.

normal test statistic is obtained. For example, if there are N subperiods ($N=4$ here), then

$$z = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \frac{\hat{\gamma}_{1n} - r_{Fn}}{s(\hat{\gamma}_{1n})} \tag{18}$$

has a standard normal distribution where, in the nth subperiod, r_{Fn} is the average T-bill return and $\hat{\gamma}_{1n}$ is the ML estimate. The last column of table 6 reports this test statistic and the p-value for a two-tailed test.[29] Estimates of γ_1 exceed the T-bill rate in each of the four subperiods, though not always by two standard errors. However, the probability of observing this overall result if $\gamma_1 = r_F$ is very low, as evidenced by the p-values in the last column. The

[29]Although a *constant* risk-free rate is not observed, the standard deviation of T-bill returns is so small compared to other asset returns that ignoring the variation and simply using the mean each subperiod is a suitable approximation. A test using the monthly differences between a time series of $\hat{\gamma}_{1t}$ [see Stambaugh (1981)] and the T-bill rate yields p-values very close to those of table 6. Most of the tests were also conducted with 'excess returns', and inferences were identical to those reported here.

estimates of γ_1 tend to decline slightly as the index is broadened, but the effect is too small to change inferences.

Table 7 displays ML estimates of the risk premium, γ_2. An overall test statistic is obtained in the same manner as described above for γ_1, except that (18) becomes

$$z = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \frac{\hat{\gamma}_{2n}}{s(\hat{\gamma}_{2n})}. \tag{19}$$

A test for a positive risk premium consists of a positive one-tailed test of $\gamma_2 = 0$. All but one of the sixteen estimates are positive, though coefficient estimates in a single subperiod are not always large relative to their standard errors. However, the results across subperiods, when taken as a whole, are unlikely if $\gamma_2 = 0$. The p-values in the last column are less than 1%, thereby rejecting $\gamma_2 = 0$ in favor of $\gamma_2 > 0$. Once again, the effect on this inference of changing the composition of the market index is negligible.

The estimates of γ_2 decline as stocks receive less weight in the market index, but standard errors decline in direct proportion, so there is no effect on inferences about $\gamma_2 = 0$. The decline in $\hat{\gamma}_2$ is matched by a commensurate increase in estimated β's. The effect is roughly approximated as a simple multiplicative scaling, and it reflects the high variance of common stock

Table 7

Maximum likelihood estimates of the risk premium, γ_2 (assets: 19 common stock industry portfolios, 4 preferred stocks, and 5 bond portfolios).

Market index	Estimate of γ_2 and standard error[a]				Overall test statistic and p-value[b]
	2/53–3/59	4/59–5/65	6/65–7/71	8/71–12/76	
1. NYSE (value-weighted)	1.384 (0.394)	0.586 (0.408)	0.323 (0.485)	0.533 (0.649)	3.22 (0.1%)
2. 1 plus Corporate bonds, Government bonds, and Treasury bills	0.593 (0.184)	0.368 (0.234)	0.105 (0.352)	0.360 (0.423)	2.97 (0.2%)
3. 2 plus Housefurnishings Automobiles, and Real estate	0.330 (0.087)	0.206 (0.117)	0.053 (0.170)	0.136 (0.178)	3.31 (0.1%)
4. Same as 3 but with the NYSE weighted by 0.10	0.181 (0.063)	0.123 (0.059)	−0.043 (0.083)	0.050 (0.106)	2.48 (0.7%)

[a]Values multiplied by 100; standard error in parentheses.
[b]A standard normal (0, 1) test statistic under the null hypothesis $\gamma_2 = 0$. The p-value in parentheses reflects a one-tailed test against the alternative $\gamma_2 > 0$.

returns relative to the variances of the returns on the other index components.[30]

As noted above, it is often the case that the estimates of γ_1 and γ_2 in single subperiods do not yield statistically 'significant' results (i.e., reject $\gamma_1 = r_F$ or $\gamma_2 = 0$). This same phenomenon is encountered in previous tests of the CAPM. For example, Fama and MacBeth (1973) also accept $\gamma_2 > 0$ and reject $\gamma_1 = r_F$, but in both cases the overall inferences are supported in only two of their six subperiods. It takes the combined effect across subperiods to produce the overall inference. Fama and MacBeth point out the difficulty in forming precise estimates of these parameters in a single subperiod, and the same lesson is true here.

4.4. Discussion of results

The impression created by this section is that inferences about the CAPM are not sensitive to altering the composition of the market index. The tests conducted here produce identical inferences across all market indexes. Specifically, the tests accept linearity and find a positive risk premium, but they reject equality of the zero-beta return to the T-bill rate. Together these inferences imply that the index portfolios lie on the positively-sloped portion of the minimum-variance boundary, but that none of them is the tangency portfolio associated with the risk-free rate. Thus, based on any of the market indexes constructed, the tests reject the traditional Sharpe–Lintner CAPM but do not reject the more general Black version.[31]

The pricing relation is tested for common stocks, preferred stocks, and bonds, but the inferences are the same as those reported for tests using only common stocks [e.g., Black, Jensen, and Scholes (1972), Fama and MacBeth (1973), and Blume and Friend (1973)]. However, there are differences in the nature of the results obtained here versus those of earlier studies. For

[30]If the other index components have literally constant returns, then the scaling is precisely $w\gamma_2$ and $(1/w)\beta$, where w is the weight received by common stocks in the index. The average $\hat{\beta}$ for stocks based on index no. 1 is approximately 1.0, whereas the average stock beta for the broader index no. 3 is roughly 4.0. Stocks receive an average weight of 25% in the latter index.

[31]The insensitivity to index composition extends beyond the results reported. Indexes no. 1 through no. 4 were also constructed with the NYSE equally-weighted index in place of the value-weighted index, and all test results were virtually identical to those shown here. The insensitivity was also found in tests employing more traditional techniques, including a replication of all of the tests performed by Fama and MacBeth (1973) using the various indexes. See Stambaugh (1981) for details.

These findings are also similar to those obtained by Miller and Scholes (1972) in a study that examines possible sources of bias in tests of the CAPM. One source of bias they examine is the choice of an improper index. A government bond return index is combined at weights of both 25% and 50% with a common stock index. Their least-squares estimate for γ_1 from the familiar 'second-pass' regression,

$$\bar{r}_i = \gamma_1 + \gamma_2 \hat{\beta}_i + u_i,$$

is virtually identical to that obtained using a stocks-only index.

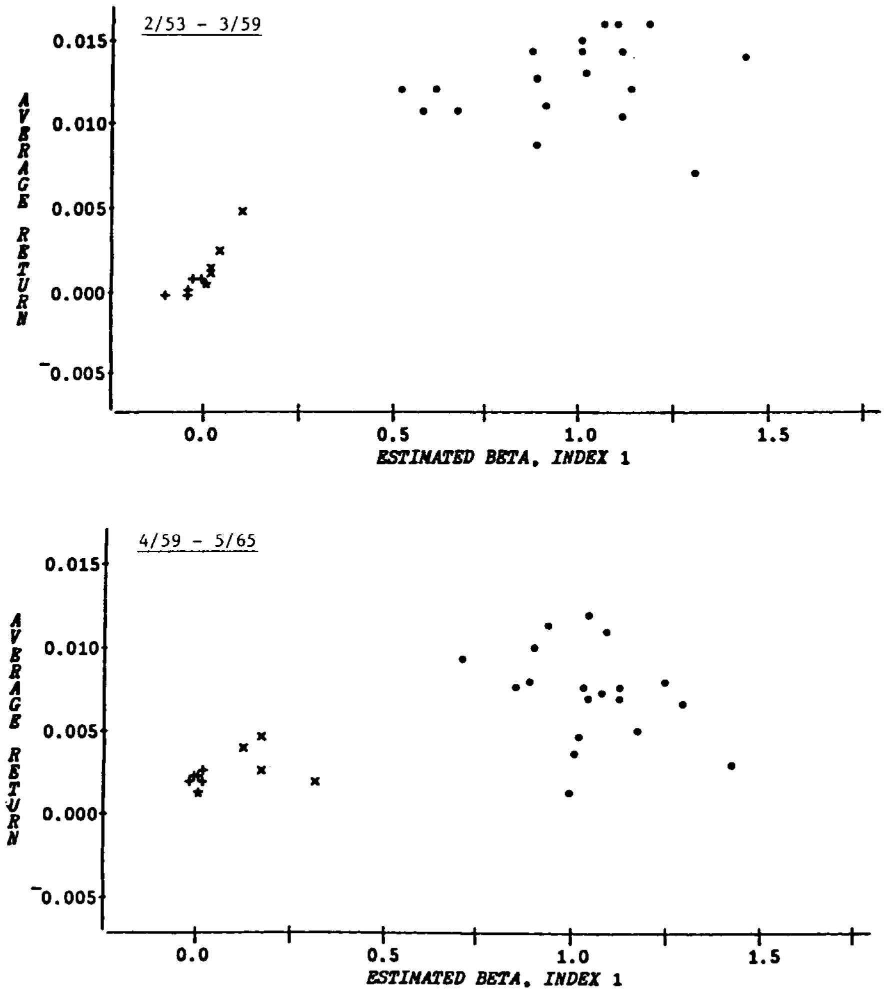

Fig. 1a. Scatter plot of average return versus estimated beta for index no. 1 (value-weighted NYSE). Key to symbols: ○: common stocks, +: bonds, ×: preferred stocks, *: one-month T-bill return.

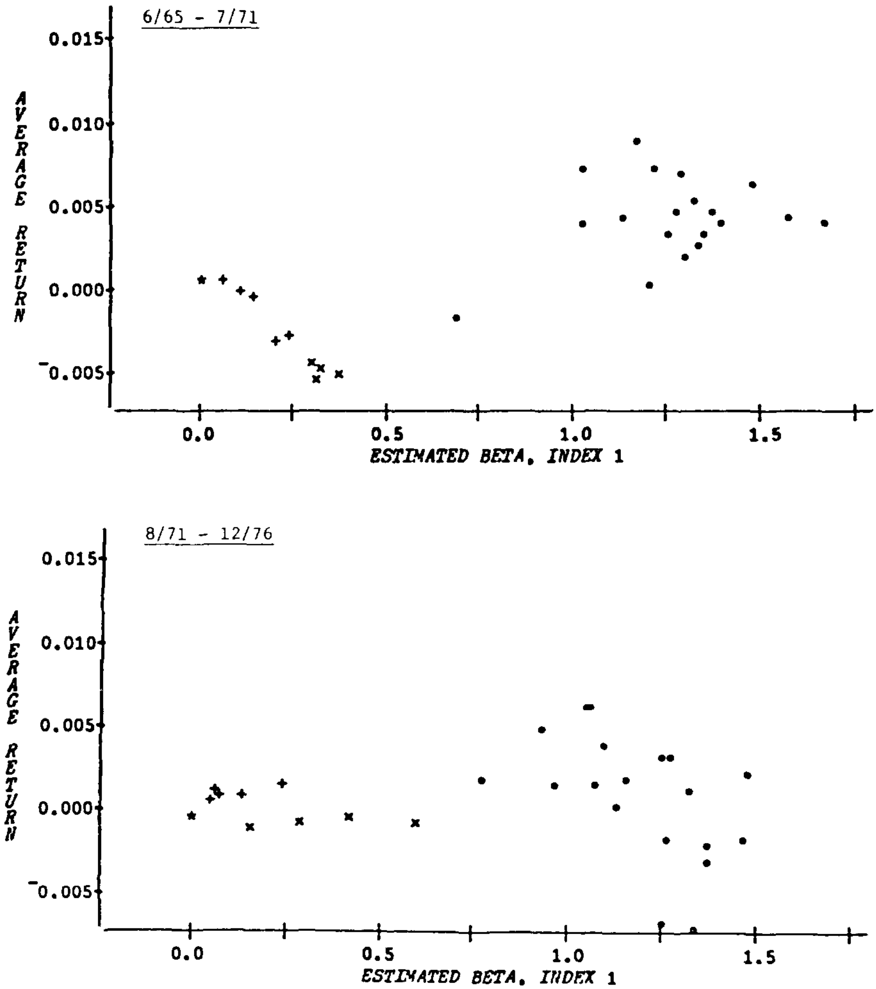

Fig. 1b. Scatter plot of average return versus estimated beta for index no. 1 (value-weighted NYSE). Key to symbols: ○: common stocks, +: bonds, ×: preferred stocks, *: one-month T-bill return.

example, the value of $\hat{\gamma}_1$ minus r_F in table 6 is 1.2% or less on an annualized basis, whereas the same difference is often 5% or more when γ_1 is estimated with only common stocks (as in the earlier studies as well as section 5 here). However, the overall p-values in table 6 are as unfavorable to the hypothesis $\gamma_1 = r_F$ as any p-values produced by using stocks alone. Thus, as the addition of bonds and preferreds produces a lower estimated zero-beta return, it also produces a lower standard error of the estimate. The reason is that the addition of these assets increases the range of parameter values, particularly beta. This is evident in the scatter plots of average return versus estimated beta in fig. 1.[32] One implication of the increased parameter range is that the zero-beta return can be estimated more precisely.

The scatter plots also appear to be different from the plots in Black, Jensen and Scholes (1972), hereafter BJS. At first glance, the BJS plots 'look' more linear. A closer inspection reveals that the main reason for the apparent difference is the scaling of the average return axis. BJS examine more subperiods over a longer total period. A greater range for average return is needed in order to use the same axes for all subperiods, particularly those pre-World War II. If the points in fig. 1 are plotted against the BJS axes, the scatter is 'compressed' from above and below and 'looks' more linear.[33]

5. Tests using different sets of market-model equations

The previous section tests the CAPM with various market indexes, but the same set of 28 assets is used throughout. This section tests the model with alternative sets of assets. In the context of the statistical methods, this means changing the vectors of asset returns (the r_i's) that appear on the left-hand side of the market-model equations in (3). Both the number of equations (K) and the identities of the assets vary. Results are reported for only one market index, the broader index no. 3, but the inferences obtained with the other market indexes are virtually identical to those reported.[34]

5.1. The alternative set of assets

Five alternative sets of assets are selected. The first three sets are subsets of the original 28 assets in section 4: (i) the 19 common stock industry

[32]The estimated betas in the scatter plots are based on the stocks-only market index (no. 1). This index is chosen simply because we are accustomed to seeing stock betas that average about 1.0. The broader indexes do not change the overall pattern — betas are simply scaled up. (Cf. footnote 30.)

[33]In particular, cf. BJS figures 10 and 11, which display plots for 105-month subperiods at least partially overlapping those used here. The points in those figures are compressed to appear as virtually flat lines.

[34]Most of these additional results are reported in Stambaugh (1981).

portfolios alone, (ii) the industry portfolios combined with the 4 preferred stocks, and (iii) the industry portfolios combined with the 5 bond portfolios. Two additional sets of assets consist of common stock portfolios formed by sorting on estimated betas: (i) 20 beta-sorted portfolios are formed in a manner first prescribed by BJS and (ii) 40 beta-sorted portfolios are constructed with the same method used by Gibbons (1982). Details of the beta-sorted portfolios are provided below.

In the BJS approach, the portfolios used in a given 'testing' subperiod are formed by sorting on betas estimated in a non-contemporaneous 'formation' subperiod. Twenty equally-weighted portfolios are formed here from the universe of all NYSE common stocks, and the CRSP equally-weighted NYSE index is used to calculate betas. The formation period for all testing subperiods but the first is the previous testing subperiod. (The second testing subperiod is used as the formation period for tests in the first subperiod as well as in the third.) In order to be included in a portfolio, a security must have data available for the entire testing subperiod and at least four years of the formation period. BJS propose this non-contemporaneous sorting in order to avoid a bias in market-model parameter estimates that arises with contemporaneous sorting. Gibbons (1982) points out that the joint ML estimation approach (used here) does not encounter this bias. He forms 40 portfolios by sorting on betas estimated in the *same* subperiod used in the tests. This contemporaneous sorting is used here to form 40 portfolios in each of the four subperiods, where the equally-weighted NYSE index is used to estimate betas.[35]

5.2. Test results

Table 8 displays results of Lagrangian multiplier tests of linearity for each of the five sets of assets described above. Overall p-values are at least 40%, so linearity is not rejected using any of the sets of assets. (Note that the *values* of the test statistics within a column are not directly comparable, because the degrees of freedom for the χ^2 distribution depends on the number of market-model equations.)

Estimates of the zero-beta return (γ_1) are shown in table 9. Inferences about the hypothesis $\gamma_1 = r_F$ are sensitive to the set of assets used. For example, equality is rejected using (i) the industry portfolios alone, (ii) the industry portfolios combined with bonds, and (iii) the 20 beta-sorted portfolios. The estimates of γ_1 in those cases usually exceed the real T-bill rate, and overall p-values are 0.2% or less. In contrast, equality is not rejected with (i) the industry portfolios combined with preferred stocks and (ii) the 40 beta-sorted portfolios.

[35]There is a slight difference in methodology because Gibbons excludes firms that change industry codes in subperiods prior to 1961.

Table 8

Lagrangian multiplier test of linearity (broad market index no. 3).[a]

No. of assets (K)	Description of assets	χ^2 statistic and p-value[b] 2/53–3/59	4/59–5/65	6/65–7/71	8/71–12/76	Overall χ^2 statistic and p-value[c]
19	Common stock industry portfolios	13.05 (73.3%)	24.78 (10.0%)	11.69 (81.8%)	14.64 (62.1%)	64.17 (60.9%)
23	Industry portfolios plus 4 preferred stocks	19.84 (53.2%)	25.93 (20.9%)	12.18 (93.5%)	18.96 (58.8%)	76.9 (69.6%)
24	Industry portfolios plus 5 bond portfolios	24.55 (31.9%)	26.44 (23.3%)	16.15 (80.8%)	21.34 (50.0%)	88.49 (46.5%)
20	Common stock beta-sorted portfolios (non-contemporaneous sorting)	12.16 (83.9%)	26.92 (8.1%)	11.76 (85.9%)	17.19 (51.0%)	68.03 (61.1%)
40	Common stock beta-sorted portfolios (contemporaneous sorting)	48.72 (11.4%)	47.62 (13.6%)	30.09 (81.6%)	28.56 (86.6%)	154.98 (41.7%)

[a]Combines the NYSE value-weighted index with six other assets using the weights in table 1.
[b]Test statistics are asymptotically distributed χ^2 with $K-2$ degrees of freedom under the null hypothesis of linearity; p-value in parentheses.
[c]The sum of the subperiod statistics, which is distributed χ^2 with $4K-8$ degrees of freedom under the null hypothesis; p-value in parentheses.

Table 10 contains estimates of the risk premium (γ_2). Inferences about γ_2 also vary across the sets of assets. The hypothesis $\gamma_2=0$ is not rejected in favor of $\gamma_2>0$ with (i) the 19 industry portfolios and (ii) the 20 beta-sorted portfolios. In those cases, estimates of γ_2 are negative as often as positive, and overall p-values exceed standard significance levels. However, the other three sets of assets support a positive risk premium with p-values of 0.1% or less.

Inferences about the CAPM are based on the combined inferences about linearity, the zero-beta return (γ_1), and the risk premium (γ_2). The sensitivity in inferences about the γ's, discovered above, creates sensitivity in the overall inference about the model. In fact, by appropriately selecting the set of assets, one can either reject or not reject the Sharpe–Lintner version or the Black version of the model. For example, both versions are rejected using (i) the 19 industry portfolios and (ii) the 20 beta-sorted portfolios. Rejection occurs because the risk premium is not significantly greater than zero. However, these two sets of assets also provide less dispersion in beta than the other sets of assets, and this makes precise estimation of γ_2 more difficult. The sets

Table 9

Maximum likelihood estimates of the zero-beta return, γ_1 (broad market index no. 3).[a]

No. of assets (K)	Description of assets	Estimate of γ_1 and standard error[b] 2/53–3/59	4/59–5/65	6/65–7/71	8/71–12/76	Overall test statistic and p-value[c]
19	Common stock industry portfolios	0.960 (0.237)	0.972 (0.362)	−0.566 (0.484)	1.336 (0.586)	3.64 (0.0%)
23	Industry portfolios plus 4 preferred stocks	0.379 (0.182)	0.338 (0.135)	−0.510 (0.281)	0.126 (0.227)	1.10 (27.3%)
24	Industry portfolios plus 5 bond portfolios	0.101 (0.029)	0.192 (0.024)	0.162 (0.037)	0.026 (0.059)	4.32 (0.0%)
20	Common stock beta-sorted portfolios (non-contemporaneous sorting)	0.248 (0.236)	1.529 (0.281)	0.288 (0.361)	−0.267 (0.486)	3.02 (0.2%)
40	Common stock beta-sorted portfolios (contemporaneous sorting)	0.053 (0.172)	0.708 (0.132)	−0.240 (0.163)	−0.724 (0.257)	−0.01 (99.7%)
		Average real T-bill return				
		0.047	0.129	0.074	−0.079	

[a]Combines the NYSE value-weighted index with six other assets using the weights in table 1.
[b]Values multiplied by 100; standard error in parentheses.
[c]A standard normal (0, 1) test statistic under the null hypothesis $\gamma_1 = r_F$, where r_F is the average real return on a one-month T-bill in the given subperiod. The p-value in parentheses reflects a two-tailed test.

of assets that include bond portfolios (here and in section 4) reject the Sharpe–Lintner version but do not reject the Black version: the tests do not reject linearity and they support a positive risk premium, but equality of the zero-beta return to the risk-free rate is rejected. Finally, the Sharpe–Lintner version is not rejected by (i) the industry portfolios plus preferred stocks and (ii) the 40 beta-sorted portfolios. A possible explanation of this last result is that these assets provide enough dispersion in beta to reject $\gamma_2 = 0$, but the exclusion of bonds prevents γ_1 from being estimated precisely enough to reject $\gamma_1 = r_F$. (Note in table 9 that the estimates of γ_1 for these sets of assets are often further from r_F than those obtained when bonds are included, but the standard errors are larger as well.)[36]

[36]The use of pre-tax returns on all assets ignores possible tax effects leading to 'after-tax' versions of the model [e.g., Litzenberger and Ramaswamy (1979)]. One could reasonably argue that the inclusion of bonds makes this simplification less comfortable and the inferences about the model more tenuous. However, the effect of differential taxes on asset returns, particularly bonds, remains an open empirical question.

Table 10

Maximum likelihood estimates of the risk premium, γ_2 (broad market index no. 3).[a]

No. of assets (K)	Description of assets	Estimate of γ_2 and standard error[b] 2/53–3/59	4/59–5/65	6/65–7/71	8/71–12/76	Overall test statistic and p-value[c]
19	Common stock industry portfolios	0.166 (0.100)	−0.054 (0.160)	0.346 (0.237)	−0.264 (0.232)	0.83 (20.4%)
23	Industry portfolios plus 4 preferred stocks	0.264 (0.097)	0.152 (0.124)	0.326 (0.198)	0.072 (0.184)	2.99 (0.1%)
24	Industry portfolios plus 5 bond portfolios	0.335 (0.088)	0.205 (0.117)	0.061 (0.171)	0.105 (0.179)	3.26 (0.1%)
20	Common stock beta-sorted portfolios (non-contemporaneous sorting)	0.433 (0.116)	−0.199 (0.145)	−0.031 (0.213)	0.197 (0.222)	1.55 (6.1%)
40	Common stock beta-sorted portfolios (contemporaneous sorting)	0.599 (0.108)	0.146 (0.121)	0.247 (0.179)	0.353 (0.190)	5.00 (0.0%)

[a]Combines the NYSE value-weighted index with six other assets using the weights in table 1.
[b]Values multiplied by 100; standard error in parentheses.
[c]A standard normal (0, 1) test statistic under the null hypothesis $\gamma_2=0$. The p-value in parentheses reflects a one-tailed test against the alternative $\gamma_2>0$.

5.3. Comparison with Gibbons' results

The results in table 8 are different from the results of Gibbons (1982, table 1). The Lagrangian multiplier (LM) test using the 40 beta-sorted portfolios does not reject linearity — the overall p-value in table 8 is 41.7%. In contrast, Gibbons performs a likelihood ratio (LR) test of (9) with the same 40 portfolios, and he rejects the restriction with a reported overall p-value of 0.0%. Recall that (9) is the linearity restriction in (8) with the additional restriction that $\gamma_2=\mu_M-\gamma_1$. As explained in section 3.3, (9) cannot be correctly tested on the broader indexes (no. 3 and no. 4). However, Gibbons uses the CRSP equally-weighted NYSE index, which captures the total return on a feasible portfolio, so the additional restriction contained in (9) is appropriate. The disparity in test results could stem from (i) differences in the restrictions tested [(8) versus (9)] and (ii) differences between the LR and LM tests. Both possibilities are investigated here.

Table 11 displays the results of LR *and* LM tests of restrictions (8) *and* (9). (The market index is the same equally-weighted NYSE index used by Gibbons.) The first two rows contain the overall test statistics and p-values

Table 11

Tests of market-model restrictions under CAPM (equally-weighted NYSE stock index).

No. of assets (K)	Description of assets	Overall χ^2 statistic and p-value[a]			
		$\alpha=\gamma_1\iota+\gamma_2^*\beta$ (8)		$\alpha=\gamma_1(\iota-\beta)$ (9)	
		Likelihood ratio[b]	Lagrangian multiplier[b]	Likelihood ratio[c]	Lagrangian multiplier[c]
40	Common stock beta-sorted portfolios (contemporaneous sorting)	247.63 (0.0%)	160.57 (30.1%)	267.18 (0.0%)	166.66 (26.5%)
20	Odd-numbered portfolios from above set of 40	88.92 (8.6%)	75.65 (36.1%)	91.86 (10.4%)	77.77 (42.2%)
20	Even-numbered portfolios from above set of 40	66.70 (65.4%)	59.02 (86.4%)	71.48 (62.6%)	63.00 (85.7%)

[a]The sum of four subperiod statistics for the period 2/53–12/76; p-value in parentheses.

[b]Asymptotically distributed χ^2 with $4K-8$ degrees of freedom under the null hypothesis in (8) (linearity).

[c]Asymptotically distributed χ^2 with $4K-4$ degrees of freedom under the null hypothesis in (9) (linearity and $\gamma_2=\mu_M-\gamma_1$, where γ_2 is the risk premium, μ_M is the mean return on the market, and γ_1 is the zero-beta return).

for the 40 beta-sorted portfolios. *Both* restrictions are rejected by the LR test with p-values of 0.0%, whereas *neither* restriction is rejected by the LM test [p-values are 30.1% for (8) and 26.5% for (9)].[37] Therefore, the disparity between the results of this study and the results of Gibbons is due to differences between the LM and LR tests.

Recall that the Monte Carlo evidence discussed in section 3.5 indicates that the LM test obeys its asymptotic distribution more closely than the LR test: the LR test appears to reject the null too often when the number of market-model equations is increased. This suggests that the low LR-test p-values are at least partially attributable to the *number* of assets. In order to pursue this question, two 20-asset subsets are selected from the 40 beta-sorted portfolios: (i) the odd-numbered portfolios and (ii) the even-numbered portfolios. If the restrictions in (8) or (9) are violated by some assets from the set of 40, then these assets are also contained in one of the 20-asset subsets. Therefore, if a test rejects the restrictions with 40 assets, one expects the same

[37]The time period used by Gibbons begins in 1926. However, the sum of his test statistics for the last four subperiods (1/56–12/75) equals 227.0, which also yields a p-value of 0.0%. Thus, his results are consistent with the LR test statistics of table 11.

result with at least one of the subsets.[38] The tests with the odd- and even-numbered subsets are reported in table 11, and, contrary to expectations, the results do not reject (8) or (9). Both subsets yield LR-test p-values above standard significance levels. However, this only weakly supports the conjecture that the *number* of assets produces the low LR-test p-values with 40 portfolios. More extensive Monte Carlo experiments would be needed to thoroughly investigate this question.

6. Conclusions

The various market index portfolios constructed here produce identical inferences about the CAPM. The portfolios include common stocks, corporate bonds, U.S. Government bonds, real estate, and consumer durables. In one portfolio, stocks represent only 10% of the total value. It remains *possible* that alternative market portfolios can reverse inferences about the model. But the results of this sensitivity analysis almost surely indicate that such an occurrence is less *likely* than Roll's (1977) arguments suggest.

Inferences prove to be sensitive to the set of assets used in the tests. Inferences based on the most inclusive set of assets — common stocks, bonds, and preferred stocks — reject the Sharpe–Lintner version of the CAPM but do not reject the more general Black version. Other sets of assets provide different inferences. However, it is important to view this sensitivity in the broader context of testing any asset pricing theory — not only the CAPM. A test of any pricing relation is based on a particular set of assets, and other sets of assets can, in principle, yield different inferences. Sensitivity to construction of the market index could imply that the CAPM is less testable than other models, but no such sensitivity is found in this study.

Appendix

The market values used to construct the portfolio weights (table 1) are obtained as follows:

(1) *NYSE common stocks.* This series is taken from a file maintained by CRSP at the University of Chicago.

(2) *Corporate bonds.* Roger Ibbotson at the University of Chicago has compiled a file containing information on most non-convertible corporate bond issues outstanding since 1926, over 14,500 issues in all. The file is used to construct year-end total par amounts outstanding and value-weighted average coupon rates and maturities. Moody's total corporate bond yield is

[38]This statement is intended to be suggestive, not rigorous. The power of the tests may also decline as the number of assets is decreased to 20. However, there is no reason to suspect this occurs here. In particular, most of the range in estimated betas found in the set of 40 portfolios is preserved in the subsets of 20.

combined with the average coupons and maturities to compute an 'average' price series. Total par values are multiplied by the 'average' prices to obtain estimated market values.

(3) and (4) *U.S. Government bonds* (including notes and certificates) *and Treasury bills.* These values are obtained from the CRSP U.S. Government securities file. The bond series is obtained by summing the month-end price times the par amount outstanding for each issue. Face values are used for Treasury bills due to their short maturities and relatively small discounts.

(5) *Residential real estate* (structures). These estimates appear in Musgrave (1976) and subsequent issues of the *Survey of Current Business.* Musgrave constructs the estimates with the 'perpetual inventory method', which is also discussed in Musgrave (1974) and in Young, Musgrave and Harkins (1971).

(6) and (7) *Housefurnishings and Automobiles.* Musgrave (1979) provides annual estimates of consumer durables, which he shows consist primarily of automobiles and housefurnishings (including appliances). The housefurnishings estimates are obtained by subtracting his estimates for automobiles from his estimates for total consumer durables. Musgrave also uses the perpetual inventory method with consumer durables, except for automobiles.

References

Barro, Robert J., 1974, Are government bonds net wealth?, Journal of Political Economy 82, 1095–1117.

Berndt, Ernst R. and N. Eugene Savin, 1977, Conflict among criteria for testing hypotheses in the multivariate linear regression model, Econometrica 45, 1263–1277.

Bildersee, John S., 1975, Some new bond indexes, Journal of Business 48, 506–525.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, Journal of Business 45, 444–454.

Black, Fischer, Michael C. Jensen and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in: Michael C. Jensen, ed., Studies in the theory of capital markets (Praeger, New York).

Blattberg, Robert C. and Nicholas J. Gonedes, 1974, A comparison of the stable and Student distributions as statistical models for stock prices, Journal of Business 47, 244–280.

Blume, Marshall, 1968, The assessment of portfolio performance, Ph.D. dissertation (University of Chicago, Chicago, IL).

Blume, Marshall and Irwin Friend, 1973, A new look at the capital asset pricing model, Journal of Finance 28, 19–34.

Eilbott, Peter, 1973, Estimates of the market value of the outstanding corporate stock of all domestic corporations, in: Raymond W. Goldsmith, ed., Institutional investors and corporate stock — A background study (National Bureau of Economic Research, New York).

Fama, Eugene F., 1976, Foundations of finance (Basic Books, New York).

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

Fama, Eugene F. and G. William Schwert, 1977, Asset returns and inflation, Journal of Financial Economics 5, 115–146.

Gallant, A. Ronald, 1975, Seemingly unrelated nonlinear regressions, Journal of Econometrics 3, 35–50.

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

Goldsmith, Raymond W., 1952, The growth of reproducible wealth of the United States of America from 1805 to 1950, in: Simon Kusnets, ed., Income and wealth of the United States: Trends and structures, Income and wealth series II (Johns Hopkins Press, Baltimore, MD).
Goldsmith, Raymond W., 1955, A study of saving in the United States (Princeton University Press, Princeton, NJ).
Goldsmith, Raymond W., 1969, Financial structure and development (Yale University Press, New Haven, CT).
Gonedes, Nicholas J., 1973, Evidence on the information content of accounting numbers: Accounting-based and market-based estimates of systematic risk, Journal of Financial and Quantitative Analysis 8, 407–444.
Housing costs in the consumer price index, 1956, Monthly Labor Review, 189–196 and 442–446.
Ibbotson, Roger G. and Rex A. Sinquefield, 1976, Stocks, bonds, bills, and inflation: Year by year historical returns, 1926–1974, Journal of Business 49, 11–47.
Kendrick, John W., 1974, The accounting treatment of human capital, Review of Income and Wealth 20, 439–468.
Kendrick, John W. and Kyu Sik Lee, 1976, Quarterly estimates of capital stocks in the U.S. private domestic economy by major industry groups, Review of Income and Wealth 22, 345–352.
Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economics and Statistics 47, 13–47.
Litzenberger, Robert H. and Krishna Ramaswamy, 1979, The effect of personal taxes and dividends on capital asset prices, Journal of Financial Economics 7, 163–195.
MacBeth, James D., 1975, Tests of the two parameter model of capital market equilibrium, Ph.D. dissertation (University of Chicago, Chicago, IL).
Miller, Merton H. and Myron Scholes, 1972, Rates of return in relation to risk: A re-examination of some recent findings, in: Michael C. Jensen, ed., Studies in the theory of capital markets (Praeger, New York).
Musgrave, John C., 1974, New estimates of residential capital in the United States, 1925–73, Survey of Current Business, Oct., 32–38.
Musgrave, John C., 1976, Fixed nonresidential and residential capital in the United States, 1925–75, Survey of Current Business, April, 46–52.
Musgrave, John C., 1979, Durable goods owned by consumers in the United States, 1925–77, Survey of Current Business, March, 17–25.
Officer, Robert R., 1971, A time series examination of the market factor of the New York Stock Exchange, Ph.D. dissertation (University of Chicago, Chicago, IL).
Roll, Richard, 1977, A critique of the asset pricing theory's tests; Part I: On past and potential testability of the theory, Journal of Financial Economics 4, 129–176.
Roll, Richard, 1979, A reply to Mayers and Rice (1979), Journal of Financial Economics 7, 391–399.
Ross, Stephen A., 1978, The current status of the capital asset pricing model, Journal of Finance 33, 885–901.
Shanken, Jay A., 1980, Some asymptotic properties of statistical methods in finance, Manuscript.
Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance 19, 425–442.
Sharpe, William F., 1978, Discussion, Journal of Finance 33, 917–920.
Silvey, S.D., 1959, The Lagrangian multiplier test, Annals of Mathematical Statistics 30, 389–407.
Silvey, S.D., 1975, Statistical inference (Chapman and Hall, London).
Stambaugh, Robert F., 1981, Missing assets, measuring the market, and testing the capital asset pricing model, Ph.D. dissertation (University of Chicago, Chicago, IL).
Stambaugh, Robert F., 1982, Testing the CAPM with broader market indexes: A problem of mean-deficiency, Working paper no. 1–82 (Rodney L. White Center for Financial Research, University of Pennsylvania, Philadelphia, PA).
Theil, Henri, 1971, Principles of econometrics (Wiley, New York).
Tri, Le Manh, 1971, The market value of outstanding corporate stock in the U.S., U.S. Securities and Exchange Commission, Statistical Bulletin, Sept., 42–47.
U.S. Bureau of Labor Statistics, 1966, The consumer price index: History and techniques, Bulletin no. 1517.
Young, Allan H., John C. Musgrave and Claudia Harkins, 1971, Residential capital in the United States, 1925–70, Survey of Current Business, Nov., 16–27.