An Empirical Analysis of Risk Premia in Futures Markets

Hendrik Bessembinder

INTRODUCTION

The classic economic rationale for the existence of futures markets is that they facilitate the transfer of risk to those most able or willing to bear it. The market price, or risk premium, for this transfer is the expected change over time in a contract's settlement price.[1] Several empirical studies document that futures price changes are not entirely random, a finding which is consistent with the existence of risk premia.[2] However, much of the prior work is descriptive and is limited to studies of closely related subgroups of futures markets such as currencies or commodities.

In a recent comprehensive study, Kolb (1992) evaluates whether Keynes' (1930) theory of "normal backwardation," which implies positive futures premia, is consistent with actual price behavior. His evidence indicates only weak support for normal backwardation: mean returns are predominately zero, and there is only weak evidence of nonzero time trends in settlement prices.

Keynes' theory did not incorporate diversification opportunities. More recent models that incorporate many risky assets and opportunities for diversification imply that futures risk premia depend only on the covariation of futures prices with relevant economic aggregates.[3] A remaining question is whether Kolb's finding of near-zero mean returns in most futures markets is consistent with models of equilibrium in the presence of diversification opportunities.

The author thanks Eric Amel, David Hirshleifer, S. P. Kothari, Paul J. Seguin, Jay Shanken, and seminar participants at Rochester, Tulane, Texas A&M, and Arizona State Universities, and the 1990 American Finance Association meetings for valuable comments, and Jay Coughenour for research assistance. Financial support by the Bradley Policy Research Center at the University of Rochester and provision of data by the Columbia University Futures Center are gratefully acknowledged.

[1]Note that, since the futures price converges to the spot price at delivery, extending the time horizon to the delivery date causes this statement to coincide with the traditional definition of the futures risk premium as the difference between the current futures price and the expected spot price at contract maturity.

[2]The evidence is perhaps most prevalent for currency futures [e.g., Hodrick and Srivastava (1984, 1987)], but similar evidence has been documented for commodity futures [e.g., Bodie and Rosansky (1980); Chang (1985); Fama and French (1987)] and financial futures [e.g., Raynauld and Tessier (1984); Rzepczynski (1987)].

[3]Examples include Dusak (1973); Grauer and Litzenberger (1979); Breeden (1980); Richard and Sundaresan (1981); and Hodrick and Srivastava (1987).

Hendrik Bessembinder is an Assistant Professor in the Department of Finance at Arizona State University.

The Journal of Futures Markets, Vol. 13, No. 6, 611–630 (1993)

CCC 0270-7314/93/060611-20

This study uses a comprehensive sample of futures price changes.[4] In addition to providing evidence that corroborates Kolb's with respect to mean returns, a simple futures pricing model is tested. The tests evaluate formally whether portfolios which include futures positions lie on the same estimated security market line (SML) as asset portfolios. This hypothesis has three testable implications: (1) that expected percentage changes in futures prices are a linear cross-sectional function of the futures' systematic risk; (2) that this linear function has a zero intercept; and (3) that the slope coefficient(s) of this linear function is equal to the systematic risk premia in asset markets.

These hypotheses are tested by adapting the multivariate Cross-Sectional Regression T^2 (CSR T^2) test presented by Shanken (1985). Results are mixed. The zero intercept condition, which implies that futures positions that contain no systematic risk earn a zero return on average, is not rejected. Point estimates of the zero beta rate for futures are remarkably close to zero. The joint hypothesis that futures returns are linear in systematic risk and that the linear function has a zero intercept is rejected when a single (equity) beta is used in the linear relation, but is not rejected when multiple betas are allowed for. However, the joint hypothesis of linearity, a zero intercept, and a risk premium on the equity beta that equals the risk premium observed in equity markets is rejected whether a single or multiple betas are included. This rejection implies that, in principle, it is possible to construct portfolios which include futures positions that dominate asset only portfolios in terms of mean return and observable beta risk. A possible interpretation for this result is provided by the model of Hirshleifer (1988), in which set-up costs cause futures returns to conditionally deviate from the SML implied by a frictionless equilibrium.

FUTURES PRICING HYPOTHESES

Futures positions can be held, in combination with financial assets, in purely financial portfolios that need not include a spot position in any deliverable good.[5] The focus of the empirical tests in this study is the proposition that expected returns to such financial portfolios are consistent with opportunities available in traditional asset markets such as equities. Modern asset pricing models predict a linear relation between expected asset returns and asset systematic risk, stated as[6]:

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

where $E(R_i^a)$ is the expected return on asset i, γ_0 is a constant (the zero-beta rate), β_i^a is a $1 \times p$ vector of sensitivities (betas) of asset i returns to each of p economic aggregates, and γ_1 is a $p \times 1$ vector of risk premia.

[4]The sample of 22 futures markets, which is also used in Bessembinder (1992), overlaps substantially with the sample of 29 markets employed by Kolb. However, five financial markets are included that Kolb does not: Treasury Notes, Eurodollars (the largest futures market by dollar volume), the S&P 500 Equity Index, the Value Line Index, and the Canadian Dollar. In contrast, Kolb employs more agricultural commodities than the six included in this study. Use of a broad cross-section of futures is particularly important for the formal hypothesis tests conducted because, as results in Stambaugh (1982) indicate, results of testing asset pricing models can be sensitive to the exclusion of securities.

[5]The purely financial model of futures premia examined here is linked to traditional storage models, in which the futures price differs from the contemporaneous spot price of the good as a function of interest rates and the net cash flow (dividend or convenience yield less storage costs) from holding a spot position, by the fact that the futures price will converge to the spot price at the delivery date. Fama and French (1987) provide an excellent discussion of the two complementary views of futures pricing.

[6]The multiple betas in (1) can result from use of a multivariate proxy for the return to the market portfolio in the Capital Asset Pricing Models of Sharpe (1964) and Black (1972), or as proxies for the multiple state variables in the models of Merton (1973), Ross (1976), and Connor (1984).

If investors can trade costlessly, principles of competitive equilibrium suggest that futures prices adjust until there is no longer a gain to entering futures contracts; that is, until relation (1) also characterizes expected returns on portfolios that include futures positions. The equilibrium condition that asset and futures markets are thus integrated provides the hypotheses tested in this study.

Proposition 1: If (1) applies to portfolios of assets and futures then the expected percentage change in futures price j, $E(R_j^f)$, is:

$$E(R_j^f) = \beta_j^f \gamma_1 \tag{2}$$

where β_j^f is a $1 \times p$ vector of sensitivities (multiple regression coefficients) of percentage changes in futures price j to the p economic aggregates, and γ_1 is defined above.

There are three testable hypotheses embedded in Proposition 1. These are: (1) expected percentage changes in futures prices are linear in the contracts' systematic risk; (2) premia for beta risk are equal across futures and asset markets; and (3) the linear risk/return function in futures markets has a zero intercept. The equal risk premia imply a uniform risk/return trade-off across asset and futures markets, while the zero intercept reflects the zero capital investment in a futures position.[7]

Propostion 1 also formalizes the equilibrium relation between the risk premium in a futures market, which takes the form of a forecastable time trend in the futures price, and the risk premium in the associated spot market, which takes the form of an expected return that exceeds the zero-beta rate. Specifically, the futures and spot premia differ only to the extent that the systematic risk of futures and spot positions differ. In markets where the spot–futures basis is relatively deterministic, systematic risk and hence risk premia should not differ significantly across the spot and the future. In contrast, in markets where the spot–futures basis displays more random variability (e.g., because of stochastic convenience yields or short sale constraints that hinder arbitrage strategies) differences in equilibrium risk premia for the spot and its associated future can exist.

To prove Proposition 1, consider a portfolio consisting of asset i and a position in one futures contract, j.[8] Let V_{it} denote the time t value of a unit of asset i, D_{it} denote per unit dividends paid by asset i at time t, F_{jt} denote futures price j at time t, N_i denote the number of units of asset i held, and N_j denote the number of long (negative if short) futures contracts held. The uncertain return on this portfolio from entry at time zero to the first settlement at time one, R_p, is given by:

$$1 + \tilde{R}_p = [N_i(\tilde{V}_{i1} + \tilde{D}_{i1}) + N_j(\tilde{F}_{j1} - F_{j0})]/(N_i V_{i0}) \tag{3}$$

where a tilde denotes a random variable. Defining $H \equiv N_j F_{j0}/N_i V_{i0}$ to denote the magnitude of the futures position relative to the asset position, the portfolio return can be stated as:

[7]In practice, futures traders must provide capital for initial margin, but this merely serves to bond performance under the contract and is not a purchase price. Further, margin can generally be provided in the form of interest-bearing instruments, with interest accruing to the futures trader.

[8]Note that it is not necessary to restrict the identity of the asset in the portfolio. In particular, margin requirements can be accommodated by allowing the asset to be a portfolio that includes the requisite margin. The asset portfolio need not include the underlying deliverable good. Therefore, considerations such as storage costs and convenience yields, which arise in the storage model of futures pricing, do not enter the pricing equations here.

$$\tilde{R}_p = \tilde{R}_i^a + \tilde{R}_j^f H \tag{4}$$

where $\tilde{R}_i^a$ is the return on asset i and $\tilde{R}_j^f$ is the percentage change in futures price j. This specification suggests that the percentage change in the futures price is analogous to a rate of return on assets. Although a slight misnomer, percentage changes in futures prices are hereafter referred to as futures returns. The expected return, $E(R_p)$, and systematic risk, β_p, of the portfolio are:

$$E(R_p) = E(R_i^a) + E(R_j^f)H \tag{5}$$

and,

$$\beta_p = \beta_i^a + H\beta_j^f \tag{6}$$

If (1) holds for this portfolio then:

$$E(R_p) = \gamma_0 + \beta_p\gamma_1 = \gamma_0 + [\beta_i^a + H\beta_j^f]\gamma_1 \tag{7}$$

Substituting into (5) from (1) and (7) yields (2).

The linear relation (2) is implied by the condition that asset/futures portfolios lie on the same SML as assets. The converse also holds; rejection of any of the three hypotheses embedded in (2) implies that there can exist asset/futures portfolios that dominate asset only portfolios in terms of expected return and beta risk. This is more readily formalized if expected futures returns are given, not by (2), but by an alternate linear function:

$$E(R_j^f) = \gamma_0^f + \beta_j^f\gamma_1^f \tag{8}$$

Then, substituting (8) and (1) into (5) yields:

$$E(R_p) = \gamma_0 + \beta_p\gamma_1 + H[\gamma_0^f + \beta_j^f(\gamma_1^f - \gamma_1)]. \tag{9}$$

The first two terms on the right side of (9) define the asset SML. If γ_0^f is not equal to zero and/or systematic risk premia in futures, γ_1^f, do not equal systematic risk premia in assets, γ_1, then the futures position, H can in principle be selected such that the third term is positive and the expected return on the portfolio exceeds the expected return on an asset portfolio with equal beta risk. If futures returns are given by a nonlinear function rather than (2), it remains true that it is possible to select futures positions such that expected returns exceed the level implied by (1), though the requisite positions will depend on the specific nonlinear function.

Issues in Interpreting Tests of Proposition 1

There is some ambiguity in interpreting the results of testing Proposition 1, since a rejection of its implications can derive from more than one source. Eq. (2) relies on the validity of the asset pricing relation (1) as well as on effective integration of asset and futures markets. Thus, even if betas are measured without error, Proposition 1 can be rejected either because of a failure of the underlying asset pricing theory or because markets are segmented. In addition, as Roll (1977) emphasizes for the case of the CAPM, inability to accurately measure the asset pricing theory's state variables introduces error into estimates of systematic risk. This measurement error can cause the rejection of Proposition 1 even if the asset pricing theory is correct and markets are integrated. As Roll observes, no empirical test of eq. (1) yet derived provides direct evidence on the validity of the asset pricing theory.

Similarly, the tests of equation (2) presented here do not provide definitive evidence on the validity of theories of capital market equilibrium. Nevertheless, results of testing Proposition 1 are of practical importance. Tests of Proposition 1 that are implemented using empirically observable proxies for the asset pricing theory's state variables are really tests of the question "do portfolios that include futures lie on the same estimated pricing function (SML) as do asset only portfolios?" While the answer to this question may not provide definitive evidence regarding asset pricing theory, it is of interest since investment portfolio performance must be evaluated relative to observable benchmarks. Regardless of the source, a rejection of Proposition 1 implies that it is possible to select future positions so that a portfolio of assets and futures earns mean returns greater than the mean returns earned by asset only portfolios with the same level of observed beta risk.

CHARACTERISTICS OF THE DATA

This section describes the data base used in this study. Descriptive statistics are presented and contrasted with findings in earlier studies.

Daily settlement prices for 22 futures markets, for the interval January, 1967 (or the inception of contract trading) to December, 1989, are obtained from the Center for the Study of Futures Markets at Columbia University and from Data Resources, Inc. These include six financial futures (Treasury Bonds, Treasury Notes, Treasury Bills, Eurodollars, the S&P 500 Index, and the Value Line Composite Index), six agricultural futures (Live Cattle, Soybeans, Sugar, Wheat, Cotton, and Corn), five foreign currency futures (the Canadian Dollar, the British Pound, the Japanese Yen, the German Mark, and the Swiss Franc), as well as five mineral futures (Gold, Silver, Platinum, Copper, and Crude Oil). These 22 markets are selected to provide a cross-section of different underlying goods, and on the basis of their relatively large trading volumes.

Individual futures contracts have a finite life defined by the contractual delivery date. Long-term futures return series are compiled as daily percentage changes in the settlement price of the contract with the nearest delivery date, except within the delivery month when daily percentage changes in the settlement price of the second nearest contract are used.[9] Monthly futures returns are obtained by linking these daily returns.

This study investigates the sensitivity of test results to the use of several economic aggregates for purposes of measuring systematic risk. These include the return on the Center of Research in Security Prices (CRSP) value-weighted equity portfolio, as well as six macroeconomic variables similar to those employed by Chen, Roll, and Ross (1986): unexpected inflation, the change in expected inflation, the change in short-term

[9]The percentage change is always computed using successive prices of the same contract, and never across contracts. The quoted price in some markets does not reflect the delivery price. For example, a quoted Treasury-Bill price of QP indicates that sellers are obligated to deliver (at contract maturity) 90-day Treasury Bills for a price such that the annualized return to the buyer is 100-QP%. In these cases, quoted prices are converted to implied delivery prices, and returns computed as the percentage change in these implied delivery prices.

Treasury-Bill yields, the change in the term structure (long- term government-bond yield less Treasury-Bill yield), the change in default premium (the yield on bonds rated BAA by Moody's less the long-term government bond yield), and the unexpected change in U.S. industrial production. Data sources and construction of these variables is described in the Appendix.

Mean Futures Returns

Table I displays mean daily percentage returns (annualized by multiplying by 250) for each of the 22 futures markets, and for four portfolios of futures positions. Since prior studies have indicated results can be sensitive to the method of return computation, Table I also presents means of returns computed as the natural logarithm of the current to prior day settlement price.

Although the samples do not fully overlap, results closely mirror those reported by Kolb (1992). For simple returns, the hypothesis that the mean futures return is zero can be rejected (p-value < 0.10) in only three of the 22 markets. None of the four portfolios of futures positions earn a significantly positive mean return. In addition to a lack of statistical significance, mean returns are fairly small in economic terms as well.[10] Annualized mean returns for portfolios of futures are 1.8% for financials, 0.7% for foreign currencies, 2.4% for commodities, and 4.5% for metals and oil. Results based on log returns are similar: the hypothesis that the mean return is zero is not rejected in 20 of the 22 individual markets or for any of the four portfolios.

Table I also reports standard deviations of futures returns, annualized by multiplying the standard deviation of daily percentage returns by the square root of 250. As a group, metals and oil futures are the most risky, with annualized standard deviations ranging from 23.7% in Gold to 33.1% in Crude Oil. Commodity futures are also relatively risky, with annualized standard deviations ranging from 20.0% in Corn futures to 47.9% in World Sugar futures. Despite widespread concern as to high levels of volatility, currency futures are the least risky as a group, with annualized standard deviations ranging from 3.9% for the Canadian Dollar to 12.0% for the Swiss Franc. Risk in financial futures is quite diverse, with annualized return standard deviations ranging from only 0.5% for Eurodollar futures to 22.2% for S&P 500 futures.

Comparison with Prior Results

In some respects, the results reported here contrast notably with earlier findings. Bodie and Rosansky (1980) examined prices in commodity futures and reported significantly positive mean returns. Their results have been widely cited as evidence of economically large mean returns in commodity futures, and as evidence that futures returns do not appear to conform to simple models such as the CAPM.[11] In contrast to their results, the more recent data examined here indicate that mean returns in commodity futures are near zero, suggesting that their evidence is not representative of more recent experience in commodity markets or for returns on more diverse sets of futures contracts. The next

[10]The two equity index futures provide noteworthy exceptions. In these markets mean rates of price change are large enough to be of economic interest, even though not statistically significant. As noted above, the risk premium in a given future approximates the excess of the expected return on the corresponding spot asset over the zero-beta rate. The observed mean rates of price change for equity index futures do approximate typical estimates of spot equity premia. The lack of statistical significance on these means reflects, in part, the relatively short sample since introduction of equity futures.

[11]For example, texts by Copeland and Weston (1988) and Sharpe and Alexander (1990) report the Bodi and Rosansky results prominently.

Table I
MEAN RETURNS (% PER DAY × 250) IN FUTURES
(t-STATISTICS IN PARENTHESES)

	Sample Start Date	Mean Simple	Mean Log	Standard Deviation		Sample Start Date	Mean Simple	Mean Log	Standard Deviation
Financial futures					Commodity futures				
T-Bills	1/76	0.12	0.12	0.62	Live Cattle	1/67	7.88	6.06	20.74
		(0.74)	(0.72)				(2.12)	(1.63)	
T-Bonds	9/77	1.99	1.05	13.66	Soy Beans	1/67	2.81	−0.04	23.91
		(0.51)	(0.27)				(0.57)	(−0.01)	
T-Notes	5/82	7.54	7.16	8.71	World Sugar	1/67	−2.76	−10.25	47.86
		(2.41)	(2.29)				(−0.28)	(−1.43)	
Eurodollars	5/82	0.53	0.53	0.48	Wheat	1/67	0.92	(−1.80)	23.37
		(3.11)	(3.11)				(0.19)	(−0.37)	
S&P 500 Index	5/82	12.84	10.27	22.19	Cotton	1/67	7.39	5.07	21.53
		(1.61)	(1.25)				(1.64)	(1.13)	
Value Line Index	5/82	9.09	7.03	20.06	Corn	1/67	−1.56	−3.55	19.96
		(1.26)	(0.96)				(−0.38)	(−0.86)	
Financial portfolio[a]		1.84	1.51	8.04	Commodity portfolio[a]		2.39	1.25	15.09
		(0.86)	(0.71)				(0.76)	(0.40)	
Metals & Oil									
Gold	1/75	−0.40	−3.21	23.72	Currencies	6/72	0.07	−0.01	3.92
		(−0.06)	(−0.53)		Canadian Dollar		(0.08)	(−0.01)	

(continued)

Table I (continued)

	Sample Start Date	Mean Simple	Mean Log	Standard Deviation		Sample Start Date	Mean Simple	Mean Log	Standard Deviation
Silver	1/67	2.91	−2.29	31.85	British Pound	6/72	0.20	−0.35	10.45
		(0.44)	(−0.34)				(0.08)	(−0.14)	
Platinum	1/67	1.49	−3.22	30.66	Japanese Yen	6/72	1.76	1.25	10.08
		(0.23)	(−0.49)				(0.74)	(0.53)	
Copper	1/67	6.42	(2.83)	26.76	German Marc	6/72	0.49	−0.05	10.38
		(1.15)	(0.51)				(0.20)	(−0.02)	
Crude Oil	4/83	16.28	10.79	33.10	Swiss Franc	6/72	1.01	0.30	11.97
		(1.28)	(0.85)				(0.36)	(0.10)	
Metals & Oil portfolio[a]		4.53	2.07	22.14	Currency portfolio[a]		0.71	0.42	7.52
		(0.98)	(0.45)				(0.40)	(0.24)	

Note: Simple returns are percentage changes in settlement prices while log returns are the natural logarithm of the ratio of current to prior settlement price. Each mean is annualized by multiplying the simple daily return mean by 250. The standard deviation is annualized by multiplying the simple daily return standard deviation by $\sqrt{250}$. The sample ends 12/89.

[a]Each portfolio return is a simple average of traded component returns.

section of this article provides formal evidence regarding whether returns in this more recent and broader sample are consistent with equilibrium predictions.

TESTS OF EQUILIBRIUM FUTURES PRICING HYPOTHESES

This section reports results of testing the null hypotheses against the general alternative that eq. (2) does not hold. Eq. (5) for the return on an asset/futures portfolio is precise for the settlement interval, i.e., the period between successive "markings to market." Since futures positions are settled daily, all hypothesis tests are conducted using daily futures returns.

Test Methodology

Proposition 1 is tested by adapting the Cross-Section Regression T^2 (CSR T^2) test developed by Shanken (1985).[12] The test is based on a determination of whether differences between realized mean futures returns and an estimated futures SML are large.

Let $\overline{F}$ denote a $k \times 1$ vector of sample mean returns in k futures markets, let p denote the number of sources of systematic risk (betas), and let $\hat{\beta}_{\mathbf{F}}$ denote the $k \times p$ matrix of estimated futures betas. The vector γ^f, obtained by a GLS regression of $\overline{\mathbf{F}}$ on $\hat{\beta}_{\mathbf{F}}$, defines the estimated futures SML. The zero intercept implication can be incorporated in the test by excluding the constant from the GLS regression. The $k \times 1$ vector $\hat{\mathbf{F}} = [\mathbf{1} : \hat{\mathbf{B}}_{\mathbf{F}}]\boldsymbol{\gamma}^{\mathrm{f}}$ contains fitted or predicted returns for each futures market, comprising a set of k points on the estimated futures SML. Letting $\mathbf{e} = \overline{\mathbf{F}} - \hat{\mathbf{F}}$ denote the vector of differences between realized and fitted mean futures returns, the test statistic for the CSR T^2 test is:

$$Q^* = Q/A$$

where

$$Q = Te'\boldsymbol{\Sigma}^{-1}e, \qquad A = (1 + \hat{\gamma}_1^{f'}\Psi\hat{\gamma}_1^f) \tag{11}$$

T is the sample size, $\boldsymbol{\Psi}$ is the estimated covariance matrix of the p economic aggregates, and Σ is the covariance matrix of the residuals from the regression of returns on economic aggregates.

Acceptance or rejection of the linear hypothesis is governed by the magnitude of the test statistic Q^*. If the sample mean futures returns do (do not) conform closely to the estimated futures SML the elements of the error vector $\mathbf{e}$ will be small (large) and Q, as a weighted sum of squared errors, will be small (large). Q^* is a downward adjusted statistic relative to Q, with the adjustment reflecting noise due to the use of estimated betas rather than true betas. Shanken (1985) shows that a transformation of Q^* is approximately distributed as an F variate.[13] Simulation evidence presented by Amsler and Schmidt (1985) supports the small sample reliability of the CSR T^2 test.

[12]The CSR T^2 test applied in asset markets is equivalently a test of whether a linear combination of the economic aggregates utilized is mean-variance efficient within the sample. Since futures contracts are (due to their zero net investment) not part of the theoretical market portfolio, results of this test implemented in futures do not have direct implications for mean variance efficiency of the market portfolio among sets of assets, and do not constitute a direct test of the Capital Asset Pricing Model.

[13]Specifically, letting T denote the number of observations, k denote the number of futures markets in the test, p denote the number of sources of systematic risk, and r denote the number of restrictions imposed on the γ vector, the variable $F = Q^*(T - k - r + 1)/(T - p - 1)(k - p - 1 + r)$ is (approximately) distributed as an F variate with $k - p - 1 + r$ and $T - k - r + 1$ degrees of freedom. See Shanken (1992) for a more detailed discussion.

The structure of the CSR T^2 test requires that the number of futures contracts employed be constant. Futures contracts used in this study commence trading at widely disparate dates throughout the sample period. To effectively use as many data as possible, the test is performed separately for five subperiods, within which the number of contracts is constant. All contracts trading at the beginning of a subperiod are included. Subperiod starting dates are chosen to match contract origination dates, while achieving roughly equal numbers of observations in each period. An overall probability value for the linear expected returns hypothesis is obtained by aggregating subperiod results in the manner described by Shanken (1985).[14]

Empirical Results

Beta Estimation Interval

For reasons discussed above, tests are conducted using daily futures returns. However, macroeconomic aggregates such as inflation and industrial production are reported on a monthly basis rather than daily, necessitating the use of betas estimated from monthly data. Beta estimates for this sample are reported in Bessembinder (1992).

To first investigate the sensitivity of results to the beta estimation interval, the single beta specification is estimated using both daily and monthly betas. Panel A of Table II reports results of a CSR T^2 test that incorporates the zero intercept restriction and is based on betas estimated from daily data. To correct for the well-known problems in estimating systematic risk from daily data, betas are estimated using the Cohen et al. (1983) generalization of the Scholes and Williams (1977) estimation technique, with ten leads and lags. The aggregate p-value from this specification is 0.039, suggesting rejection of the joint hypothesis of a linear expected return relation and the zero intercept constraint. The first (January, 1967 to May, 1972) and fourth (May, 1982 to December, 1985) subperiod F-statistics are largest; thus futures returns during these subperiods contribute most to this rejection.

Panel B of Table II reports results of a CSR T^2 test identical to that in panel A except that the beta estimates are obtained by OLS regression of monthly futures returns on monthly equity returns, rather than Scholes–Williams estimates from daily returns. Results, in terms of both the estimated risk price and the F-statistic are, in each subperiod and overall, virtually indistinguishable from those reported on panel A.

Results in Table II support two conclusions. First, the joint hypothesis that expected futures returns are linear in a single (equity) beta and that the linear function has a zero intercept is not consistent with the data. Second, inference is not affected by use of betas estimated from monthly returns as compared to use of betas estimated from daily returns. This second conclusion is important since the use of a single beta estimate may improperly measure systematic risk, while testing with multiple betas requires that systematic risk be estimated from monthly data.

[14]The F-statistic for each subperiod is converted to the standard normal variate with the same p-value. These standard normal variates are summed across subperiods and divided by the square root of the number of subperiods to obtain an overall standard normal variate, from which the p-value for the joint hypothesis that the mean is zero in all subperiods can be ascertained. The aggregate p-value is the probability, under the null hypothesis, of observing an aggregate standard normal statistic as large or larger than the one actually obtained. This aggregation procedure is evaluated further by Gibbons and Shanken (1987).

Table II
CSR T^2 TEST OF HYPOTHESIS THAT EXPECTED FUTURES RETURNS ARE LINEAR IN A SINGLE BETA

H_0: $E(R_j) = B_j \gamma_1$

	Period					
	1/67–5/72	**6/72–8/77**	**9/77–4/82**	**5/82–12/85**	**1/86–12/89**	**Aggregate**
	(T = 1304, K = 8)	(T = 1298, K = 14)	(T = 1161, K = 17)	(T = 918, K = 21)	(T = 1002, K = 22)	
Panel A: Daily returns and daily equity market beta[a]						
$\hat{\gamma}_1$ (EVW)	0.032	–0.049	–0.125	0.082	0.101	
F-statistic	2.67	0.66	0.65	1.82	1.41	
p-Value	0.010	0.800	0.836	0.015	0.104	0.039
Panel B: Daily returns and monthly equity market beta[b]						
$\hat{\gamma}_1$ (EVW)	0.019	–0.027	–0.083	0.082	0.101	
F-statistic	2.69	0.69	0.73	1.80	1.41	
p-Value	0.009	0.771	0.763	0.017	0.104	0.028

Note: T is the number of observations and K the number of futures markets employed in each subperiod. The F-statistic is based on a weighted average of squared deviations between realized mean returns and returns predicted by the estimated security market line. In each case the SML is estimated by a GLS regression of mean daily returns on estimated betas. γ coefficients are (250 times) the slope coefficients obtained in GLS regression.

[a]Betas estimated by Scholes–Williams regressions of daily futures returns on daily CRSP value-weighted index returns, using ten leads and ten lags.

[b]Betas estimated by simple regression of monthly returns on monthly CRSP value-weighted index returns.

Sensitivity of the CSR T^2 test to Pricing Constraints

Before investigating the effect of allowing for multiple sources of systematic risk, it may be useful to examine the sensitivity of the results reported in Table II to the constraints that Proposition 1 places on parameters of the estimated futures SML. Panel A of Table III reports results of a CSR T^2 test based on the single equity beta, but that does not impose the zero intercept constraint. Inference is largely unaffected by relaxing this constraint; the aggregate p-value for the linear expected returns hypothesis is 0.07. Inference is unaltered by relaxing the zero intercept constraint because the unconstrained intercept estimates are small. For each subperiod the absolute value of the intercept estimate reported in Table II is less than or equal to 0.6% per year. These point estimates imply that futures positions that contain no systematic risk earn mean returns very near zero.

Proposition 1 implies that risk premia in futures markets do not differ from premia in asset markets. Sensitivity to this implication can be ascertained by constraining the futures premia to equal the premia observed in an asset market. To implement this test, the slope coefficient on the equity beta, γ_1, is, in each subperiod, constrained to equal the ex post premium realized in equity markets.[15] Panel B of Table III reports the results of the CSR T^2 test subject to this constraint. Inference is again little affected. The aggregate p-value for the test of Proposition 1 is 0.051. Once again, the reason is that unconstrained risk premia estimates (reported in panel A) differ little from the constrained premia from equity markets (reported in panel B).

The relatively low p-values reported in Tables III and IV suggest that Proposition 1 is, subject to the maintained assumption that covariation with the CRSP value-weighted equity portfolio return adequately measures systematic risk, rejected. This rejection implies the existence of asset/futures portfolios with mean returns that exceed the level implied by the SML estimated from equity return data.

The relative insensitivity of these test results to restriction on the parameters of the futures security market line suggest that the rejection of Proposition 1 is caused by the linearity condition itself. Failure of this implication could reflect failure of the underlying asset pricing theory or effective market segmentation. However, the rejection could also result because the single beta estimated on the basis of covariation with a diversified equity index is not sufficient to capture the systematic risk which is actually priced in financial markets.

The CSR T^2 Test with Multiple Betas

Table IV reports results of testing a specification in which futures returns are fit to a multidimensional SML by a cross-sectional GLS regression of mean returns on multiple betas. The seven sources of systematic risk include, in addition to covariation with the CRSP value-weighted equity portfolio, covariation with six macroeconomic variables similar to those employed by Chen, Roll, and Ross (1985). This test of Proposition 1 provides evidence on whether mean returns to portfolios that include futures positions lie on the same multidimensional SML as do asset portfolios. Panel A of Table IV reports a version in which the parameters of the futures SML are unconstrained. The aggregate p-value from this specification is 0.553. Thus, in contrast to the single beta results, the

[15]Specifically, γ_1 is constrained to equal the ex post mean of the excess of the return to the CRSP value-weighted equity portfolio over the one-month T-bill yield.

Table III
SENSITIVITY OF SINGLE BETA CSR T^2 TESTS TO INTERCEPT AND RISK PREMIUM CONSTRAINTS

	H_0: $E(R_j) = \gamma_0 + B_j\gamma_1$ Period					
	1/67–5/72	6/72–8/77	9/77–4/82	5/82–12/85	1/86–12/89	Aggregate
	(T = 1304, K = 8)	(T = 1298, K = 14)	(T = 1161, K = 17)	(T = 918, K = 21)	(T = 1002, K = 22)	
Panel A: Without $\gamma_0 = 0$ intercept constraint						
γ_0 (Intercept)	0.006	–0.002	–0.001	0.005	0.001	
γ_1 (EVW)	0.016	–0.019	–0.086	0.100	0.104	
F-statistic	2.52	0.75	0.77	1.30	1.47	
p-Value	0.020	0.701	0.710	0.173	0.084	0.070
Panel B: With $\gamma_0 = 0$ and $\gamma_1 = \gamma_1^a$ constraints						
γ_1^a (EVW)	0.035	–0.003	0.007	0.102	0.105	
F-statistic	2.40	0.65	0.74	1.73	1.34	
p-Value	0.014	0.827	0.761	0.028	0.133	0.051

Note: T is the number of observations and K the number of futures markets employed in each subperiod. The F-statistic is based on a weighted average of squared deviations between realized mean returns and returns predicted by the estimated security market line. In each case the SML is estimated by a GLS regression of mean daily returns on estimated betas. γ coefficients are (250 times) the slope coefficients obtained in GLS regression.

Table IV

CSR T^2 TESTS OF HYPOTHESIS THAT EXPECTED RETURNS ARE LINEAR IN MULTIPLE BETAS

$$H_0: E(R_j) = \gamma_0 + \sum_{i=1}^{7} \gamma_1 B_{jf}$$

	Period					
	1/67–5/72	6/72–8/77	9/77–4/82	5/82–12/85	1/86–12/89	Aggregate
	(T = 1304, K = 8)	(T = 1298, K = 14)	(T = 1161, K = 17)	(T = 918, K = 21)	(T = 1002, K = 22)	
Panel A: Intercept and slopes unconstrained						
γ_0 (Intercept)	0.017	0.019	–0.011	0.007	0.019	
γ_1 (EVW)	0.081	–0.455	0.110	0.047	0.148	
γ_2 (CEI)	1.260	0.169	0.100	0.177	–0.129	
γ_3 (EUI)	–3.920	–1.425	2.056	–0.945	0.215	
γ_4 (CT3)	–4.309	5.805	–3.920	0.699	2.875	
γ_5 (CTS)	3.560	0.270	5.063	–2.046	–3.270	
γ_6 (CRP)	–1.498	–1.997	–0.016	1.087	–2.698	
γ_7 (CIP)	–16.15	–2.508	–1.275	7.245	–3.955	
F-statistic	0.95	0.51	0.73	0.57	1.76	
p-Value	0.331	0.805	0.682	0.817	0.039	0.553
Panel B: $\gamma_0 = 0$ constraint imposed						
γ_1 (EVW)	–0.816	–0.415	–0.082	0.063	0.160	
γ_2 (CEI)	–2.481	–0.033	–0.256	0.174	–0.375	
γ_3 (EUI)	4.712	–2.083	2.618	–1.802	0.175	

Table IV (continued)

	$H_0: E(R_j) = \gamma_0 + \sum_{i=1}^{7} \gamma_1 B_{jf}$					
	Period					
	1/67–5/72	6/72–8/77	9/77–4/82	5/82–12/85	1/86–12/89	Aggregate
γ_4 (CT3)	2.977	3.040	2.733	–3.210	0.860	
γ_5 (CTS)	–2.358	2.601	–1.310	1.688	–1.513	
γ_6 (CRP)	–1.827	–1.947	–0.023	1.519	–2.382	
γ_7 (CIP)	–20.931	–2.507	–4.461	7.727	–4.186	
F-statistic	0.26	0.49	0.81	0.75	1.95	
p-Value	0.611	0.851	0.620	0.727	0.016	0.516
Panel C: $\gamma_0 = 0$ and $\gamma_1^f = \gamma_1^a$ constraints imposed						
γ_1^a (EVW)	0.035	–0.003	0.007	0.102	0.105	
γ_2 (CEI)	–0.984	0.049	–0.202	–0.110	0.476	
γ_3 (EUI)	1.309	–2.078	2.655	–0.458	–1.095	
γ_4 (CT3)	7.508	1.207	2.648	–3.448	1.185	
γ_5 (CTS)	–5.450	1.412	–1.205	1.954	–1.708	
γ_6 (CRP)	–3.172	–0.964	–0.072	0.976	–0.736	
γ_7 (CIP)	–14.564	4.905	–3.733	4.085	2.303	
F-statistic	1.69	0.74	0.75	1.03	2.87	
p-Value	0.185	0.650	0.692	0.422	0.000	0.042

Note: T is the number of observations and K the number of futures markets employed in each subperiod. The F-statistic is based on a weighted average of squared deviations between realized mean returns and returns predicted by the estimated security market line. In each case the SML is estimated by a GLS regression of mean daily returns on estimated betas. Betas are estimated by multiple regression of monthly futures returns on CRSP value weighted index returns and six macroeconomic variables. Gammas reported are (250 times) coefficients obtained in GLS regression.

multi-beta test does not reject the existence of a linear relation between expected returns and systematic risk in futures markets.[16]

Proposition 1 implies not only a linear relation, but places restriction on the parameters of the linear function. Panel B of Table IV reports test results when the zero intercept implication is imposed on the multi-beta futures SML. Inference is again unaffected by this constraint; the aggregate *p*-value is reduced only to 0.516. As in the single beta case, inference is unaffected by the zero intercept constraint because the unconstrained estimates are near zero.

Panel C of Table IV reports test results when, in addition to the zero intercept constraint, the slope coefficient on equity beta risk is constrained to equal that estimated in equity markets.[17] Inference *is* affected by this constraint; the aggregate *p*-value for the joint hypothesis is reduced to 0.042, suggesting rejection of Proposition 1.

Inference is affected in this case because the estimate of the premium for covariation with equity returns obtained in the sample of futures (the unconstrained γ_1 estimate) differs significantly from the premium actually observed in the equity markets. Though the *p*-value on the joint hypothesis is reduced by this constraint in four of the five subperiods, the effect is most notable in the first subperiod (January, 1967 to May, 1972) where the unconstrained estimate of premia for equity risk is an implausible –81.6% per year, compared to the constrained value of 3.5% per year.

As discussed earlier, there are at least three reasons that could underlie the rejection of Proposition 1. First, asset and futures markets may be partially segmented. This could arise, for example, from the existence of informational costs to futures trading as suggested by Hirshleifer (1988). Second, the asset pricing relation (1) may not hold.[18] Third, the measures of systematic risk used here may not be adequate to proxy for the marginal utility of the representative investor.

Regardless of the underlying reason, rejection of (2) implies the existence of portfolios comprised of assets and futures positions that dominate (as well as portfolios that are dominated by) asset-only portfolios in terms of mean return and systematic risk, when systematic risk is estimated with respect to the macroeconomic variables used in this analysis. This point is relevant because portfolio performance evaluation must necessarily be based on observable benchmarks.

The existence of such superior performing portfolios is a logical implication of rejecting Proposition 1. However, this observation alone does not give rise to identifiable portfolio strategies. The tests based on multiple beta estimates do not reject the implication that the relation between futures returns and betas is linear. Assuming that the relation is in fact linear, eq. (9) describes returns to asset/futures portfolios. From (9) it is evident that the identification of futures positions that lead to superior portfolio performance

[16]Increasing the number of betas in the cross-sectional regression necessarily reduces the weighted average pricing error, Q. Note, though, that the conversion from the Q-statistic to the F-statistic and the degrees of freedom in the F-test are both altered by the number of betas used in the test. That is, the *p*-value on the test does not mechanically increase as the number of betas increases.

[17]To estimate the remaining risk premia subject to this constraint, the transformed dependent variable $\overline{F}^* = \overline{F} - \beta_1 \gamma_1^a$, where β_1 is the estimated equity beta and γ_1^a is the constrained equity premium, is regressed on the remaining betas.

[18]Several studies, including Shanken (1985), have tested and rejected the proposition that one of the economic aggregates used here, the return to the CRSP value-weighted equity index, is mean-variance efficient. It may be tempting to conclude that given this existing evidence, the rejection of Proposition 1 likely derives from failure of the asset pricing theory. However, as Roll (1977) emphasizes, such rejections may also reflect errors in measuring systematic risk.

requires knowledge of the signs of divergences between factor premia in futures and factor premia in assets $(\gamma_1^f - \gamma_1)$.

The rejection of Proposition 1 is against the unspecified general alternative that eq. (2) does not hold, and thus does not identify the economic determinants of divergences in factor premia across markets. Further, point estimates in Table IV indicate that the ex post difference in factor premia changes sign across subperiods, implying that the composition of superior performing portfolios changes across time. Elton, Gruber, and Rentzler (1987) report that positions in professionally managed commodity funds would have detracted from rather than enhanced the performance of asset portfolios during the 1980s. Thus, their evidence suggests that the technical trading rules followed by the managers of these funds do not succeed in identifying the superior performing portfolios that this analysis suggests exist.

SUMMARY AND CONCLUSIONS

This study presents evidence on mean futures returns and tests a simple model of futures risk premia using a comprehensive cross-section of prices in 22 markets, including financial, foreign currency, agricultural, and metals futures. Similar to the results reported by Kolb (1992), the hypothesis that mean returns in futures are zero is rejected in three of the 22 markets. For commodity futures in particular, the evidence indicates generally zero mean returns. These results contradict the widely cited findings of Bodie and Rosansky (1980).

The equilibrium hypothesis examined is that relations between capital investment, risk, and returns do not differ across futures markets and asset markets. More precisely, the hypotheses tested are: (1) that expected futures returns are a linear function of futures' systematic risk; (2) that the linear function has (to reflect the zero capital investment) a zero intercept; and (3) that risk premia (slope coefficients) equal to those in asset markets. Tests are conducted using both a single and multiple betas.

Results are mixed. The zero intercept condition, which implies that futures positions that contain no systematic risk earn a zero return on average, is not rejected. Point estimates of the zero beta rate for futures are remarkably close to zero. The joint hypothesis that futures returns are linear in systematic risk and that the linear function has a zero intercept is rejected when a single (equity) beta is used in the linear relation, but is not rejected when multiple betas are allowed. However, the joint hypothesis of linearity, a zero intercept, and a risk premium on the equity beta that equals the risk premium observed in equity markets is rejected whether a single or multiple betas are included.

While alternate explanations, including failure of the underlying asset pricing theory or improper measurement of systematic risk cannot be ruled out, the evidence is consistent with the reasoning that futures markets are not fully integrated with asset markets. This evidence suggests that further progress in understanding futures pricing may derive from models that explicitly incorporate market imperfections.

Roll (1977) argues that, due to difficulty in observing relevant economic aggregates, tests of asset pricing models are less than fully definitive. Such ambiguity also applies to results reported here. It cannot be stated with certainty that the rejection of Proposition 1 derives from failure of the asset pricing theory, improper measurement of systematic risk, or market segmentation.

Despite these ambiguities, these results are of practical importance. The empirical rejection of Proposition 1 is really a negative answer to the question: "Do all portfolios that include futures lie on the same *estimated* pricing function (SML) as do asset only portfolios?" This is important because portfolio performance must be evaluated relative to observable benchmarks.

Appendix

FUTURES PRICES

Daily settlement prices from 22 markets, over the interval January, 1967 (or the inception of trading) until December, 1989 are obtained from the Columbia University Futures Center and from Data Resources, Inc. The futures markets are associated with the following exchanges:

Chicago Board of Trade
- Treasury Bonds
- Treasury Notes
- Soybeans
- Corn
- Wheat

Chicago Mercantile Exchange
- Treasury Bills
- Eurodollars
- S&P 500 Index
- Janpanese Yen
- Deutsche Mark
- Canadian Dollar
- Swiss Franc
- British Pound
- Live Cattle

Kansas City Board of Trade
- Value Line Stock Index

Commodity Exchange, Inc.
- Gold
- Silver
- Copper

New York Mercantile Exchange
- Crude Oil
- Platinum

New York Cotton Exchange
- Cotton #2

Coffee, Sugar, and Cocoa Exchange
- Sugar #11

MACROECONOMIC VARIABLES

Monthly data on U.S. industrial production and U.S. inflation (based on the CPI) are obtained from the International Financial Statistics data tape. Interest rate data is obtained from Data Resources, Inc. (DRI). These include the yield on three-month Treasury Bills, the yield to maturity on 20-year Treasury Bonds, and the yield to maturity on long-term corporate bonds rated BAA by Moody's. The term structure variable is the (the change in) the Treasury-Bond yield less the Treasury-Bill yield. The default premium series is (the change in) the BAA bond yield less the Treasury-Bond yield.

The unexpected change in industrial production series is created as the residuals from an ARIMA(3, 1, 0) model fitted to the raw industrial production data. The inflation series is decomposed into expected and unexpected components using the "naive interest rate model" evaluated by Fama and Gibbons (1984), who find that the this simple method provides inflation forecasts as accurate as more complex methods. Expected inflation is determined as the difference between the one-month Treasury-Bill yield and the expected real interest rate, while the expected real interest rate is determined, in turn, as the simple average of ex post realized real T-Bill returns over the prior 12 months.

Bibliography

Amsler, C., and Schmidt, P. (1985): "A Monte Carlo Investigation of the Accuracy of Multivariate CAPM Tests," *Journal of Financial Economics,* 14:349–375.

Bessembinder, H. (1992): "Systematic Risk, Hedging Pressure, and Risk Premia in Futures Markets," *Review of Financial Studies,* 5:637–667.

Bodie, Z., and Rosansky, V. (1980): "Risk and Return in Commodity Futures," *Financial Analysts Journal,* 27–39.

Breeden, D. (1980): "Consumption Risk in Futures Markets," *Journal of Finance,* 35:520–530.

Black, F. (1972): "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business,* 45:444–455.

Chang, E. (1985): "Returns to Speculators and the Theory of Normal Backwardation," *Journal of Finance,* 40:193–208.

Chen, N., Roll, R., Ross, S. (1986): "Economic Forces and the Stock Market," *Journal of Business,* 383–403.

Cohen, K., Hawawini, G., Maier, S., Schwartz, R., and Whitcomb, D. (1983): "Friction in the Trading Process and Systematic Risk," *Journal of Financial Economics,* 12:263–278.

Connor, G. (1984): "A Unified Beta Pricing Theory," *Journal of Economic Theory,* 34:13–31.

Copeland, T., and Weston, J. (1988): *Financial Theory and Corporate Policy.* New York: Addison-Wesley.

Dusak, D. (1973): "Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premium," *Journal of Political Economy,* 81:1387–1406.

Elton, E., Gruber, J., and Rentzler, J. (1987): "Professionally Managed, Publicly Traded Commodity Funds," *Journal of Business,* 60:175–199.

Fama, E., and French, K. (1987): "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage," *Journal of Business,* 60:55–73.

Fama, E., and Gibbons, M. (1984): "A Comparison of Inflation Forecasts," *Journal of Monetary Economics,* 13:327–348.

Gibbons, M., and Shanken, J. (1987): "Subperiod Aggregation and the Power of Multivariate Tests of Portfolio Efficiency," *Journal of Financial Economics,* 19:389–394.

Grauer, F., and Litzenberger, R. (1979): "The Pricing of Commodity Futures Contracts, Nominal Bonds and other Risky Assets under Commodity Price Uncertainty," *Journal of Finance,* 33:69–83.

Hodrick, R., and Srivastava, S. (1987): "Foreign Currency Futures," *Journal of International Economics,* 22:1–24.

Hodrick, R., and Srivastava, S. (1984): "An Investigation of Risk and Return in Forward Foreign Exchange," *Journal of International Money and Finance,* 3:5–29.

Hirshleifer, D. (1988): "Residual Risk, Trading Costs, and Commodity Futures Risk Premia," *Review of Financial Studies,* 1:173–193.

Kolb, R., (1992): "Is Normal Backwardation Normal?," *The Journal of Futures Markets,* 12:75–91.

Keynes, J. (1930): *A Treatise on Money.* London: Macmillan: vol. 2.

Merton, R. (1973): "An Intertemporal Capital Asset Pricing Model," *Econometrica,* 41:867–887.

Raynauld, J., and Tessier, J. (1984): "Risk Premiums in Futures Markets: An Empirical Investigation," *The Journal of Futures Markets,* 4:189–211.

Richard, S., and Sundaresan, M. (1981): "A Continuous Time Equilibrium Model of Commodity Prices in a Multigood Economy," *Journal of Financial Economics,* 9:347–371.

Roll, R. (1977): "A Critique of the Asset Pricing Theory's Tests; Part 1: On Past and Potential Testability of the Theory," *Journal of Financial Economics,* 4:129–176.

Rzepczynski, M. (1987): "Risk Premiums in Financial Futures Markets: The Case of Treasury Bond Futures," *Journal of Futures Markets,* 7:653–662.

Ross, S. (1976): "The Arbitrage Theory of Capital Asset Pricing," *The Journal of Economic Theory,* 13:341–360.

Scholes, M., and Williams, J. (1977): "Estimating Beta from Non-synchronous Data," *Journal of Financial Economics,* 5:309–327.

Shanken, J. (1992): "On the Estimation of Beta Pricing Models," *Review of Financial Studies,* 5:1–33.

Shanken, J. (1985): "Multivariate Tests of the Zero-beta CAPM," *Journal of Financial Economics,* 14:327–348.

Sharpe, W., and Alexander, G. (1990): *Investments.* Englewood Cliffs, NJ: Prentice-Hall.

Sharpe, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance,* 19:425–442.

Stambaugh, R. (1982): "On the Exclusion of Assets from Tests of the Two-parameter Model: A Sensitivity Analysis," *Journal of Financial Economics,* 10:237–268.