The Arbitrage Pricing Theory and Supershares
Author(s): Mark Latham
Source: *The Journal of Finance*, Vol. 44, No. 2 (Jun., 1989), pp. 263-281
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328590
Accessed: 26/11/2009 07:25

THE JOURNAL OF FINANCE • VOL. XLIV, NO. 2 • JUNE 1989

The Arbitrage Pricing Theory and Supershares

MARK LATHAM*

ABSTRACT

In a single-period model with options on the market portfolio, linear factor pricing holds if and only if the variance of the market conditional on the factors is zero. There is no need for factors other than nonlinear functions of the market. For accurate linear pricing of all payoff patterns the factors must be rotationally equivalent to Hakansson's "supershares." In a multiperiod model, a similar set of results holds, but with consumption replacing the market payoff. The methodology of the empirical Arbitrage Pricing Theory literature is not consistent with either the single-period model or the multiperiod model.

LINEAR FACTOR MODELS HAVE been used extensively to analyze security returns, predominantly by researchers working with the "Arbitrage Pricing Theory" (APT) paradigm.[1] A coherent theoretical framework is, of course, crucial to give direction and interpretation to empirical work, as well as to guide applications. This paper presents some new basic theoretical results on factor models and security pricing. In a single-period model with market-index options, even though the Sharpe-Lintner CAPM may not hold, the market portfolio plays such a central role in pricing that the following condition is necessary and sufficient for a vector of factors to give linear pricing: the variance of the market return conditional on the factors must be zero; that is, the factors must determine the market.

Furthermore, there is no need for factors other than (potentially nonlinear) functions of the market return. Nonmarket risk in a factor can be eliminated with virtually no loss of linear pricing accuracy. In order to get linear pricing of all assets in a complete market, the factors must be rotationally equivalent to Hakansson's (1977, 1978) "supershares."

In an intertemporal model, a similar set of results holds, with consumption replacing the market return as the pivotal variable for pricing all securities.

These results are used to highlight some limitations of the existing theoretical APT literature as an interpretive foundation for APT empirical work. It is argued that the empirical work is consequently somewhat misdirected, so recommendations for redirection are made at the end of the paper.

Section I discusses single-period APT models and introduces a complete-markets model. Six theorems are derived which indicate that the potential role

* School of Business Administration, University of California, Berkeley. For helpful discussion I thank participants in the U. C. Berkeley finance seminar, especially Gregory Connor. Financial support from the Berkeley Program in Finance is gratefully acknowledged. An earlier version of this paper was circulated under the title of "The Arbitrage Pricing Theory: A State-Preference Analysis."

[1] See Huberman (1987) for a review of the empirical and theoretical literatures.

of factors is too limited to justify the usual interpretations of empirical APT studies. The theorems are extended to a multiperiod setting in Section II and related to the Connor and Korajczyk (1987) intertemporal APT model. These models share much of the logic of the consumption-based asset pricing analyses of Breeden and Litzenberger (1978) and Breeden (1979), as well as sharing their practical limitations. Section III discusses implications of these results for empirical work. Section IV gives a summary and conclusions.

I. Single-Period Models

There are different versions of the APT in the theoretical literature. The original version, in Ross (1976b), is inexact in the sense that it does not determine expected returns, but rather determines a bound on deviations of expected returns from a benchmark formula. With stronger assumptions, Connor (1984) derived an exact version, sometimes called an "Equilibrium-APT" because of the key role of aggregate supply (the market portfolio). Our model is closer to Connor's, so we will refer mainly to his results. Milne (1988) extended Connor's results but based his analysis on a nonfundamental assumption about induced preferences: his assumption A.8 is that investors will only want to hold portfolios spanned by the factors. This avoids the important issue of what conditions are needed for investors to want that—an issue that is central to both Connor's paper and the present paper.

The one piece of the APT that is truly based on arbitrage is that, if a security's return (or payoff) is an *exact* linear function (no error term) of the factors, then its expected-return premium (or current price) is a linear function of the factor premia (or factor-portfolio present values). This is just "value additivity." However, the important and interesting pieces of the APT are based on equilibrium rather than arbitrage analysis. They deal with the issues of deriving conditions under which the "epsilons" (error terms in the factor structure) are not priced (no return premium or present value) and conditions under which investors' portfolios have uniformly zero epsilons (typically the same conditions). Important questions not addressed by the APT are whether anything can be said about the factor premia and what factor weightings are appropriate for any given investor's portfolio.

In assessing the unique contributions of the APT, it is helpful to consider the theoretical reasons why the Sharpe-Lintner CAPM may not hold. These include insufficiency of mean and variance for portfolio choice, intertemporal hedging effects, missing assets, market segmentation, taxes, and transaction costs. This section will follow Connor (1984) in analyzing a model in which sufficiency of mean and variance is the only missing CAPM assumption. The limited role of factors in this setting is contrasted with the usual APT portrayal in which different factors represent different dimensions of risk. Note that, in the APT literature, factors are often orthogonalized and often identified with distinct macroeconomic variables as sources of risk.

We will make the additional simplifying assumption that enough securities exist to span the payoffs of Hakansson's (1977, 1978) supershares. For example, a complete set of options on the market portfolio would be sufficient. Then portfolios can be constructed to give any desired pattern of market-index-contingent payoffs. (The resulting equilibrium prices and portfolios are the same as if we assume complete markets, and assuming completeness simplifies many arguments in this paper, so we make this stronger assumption with no loss of generality.) The readiness of modern security markets to create new instruments (such as index options, portfolio insurance, and the SuperShares® introduced by Leland O'Brien Rubinstein Associates[2]) lends support to the position that the real-world equilibrium is not very different from a full supershare equilibrium. In any case, this supershare assumption, like those of no-transaction-costs and homogeneous beliefs, will help us focus on the more substantial issues of risk, return, and optimal portfolios.

Consider a one-period pure exchange economy of investors who trade their endowments of securities at time $t = 0$ and consume only at time $t = 1$. Consumption for any investor k equals her or his wealth at time 1, which is the sum of the time-1 payoffs on the securities in her or his optimally chosen equilibrium portfolio. Investor k maximizes her or his time-0 expectation of utility of time-1 wealth $E_0\{U_k(\tilde{W}_k^1)\}$. Utility is state independent. Assume $U_k' > 0$, $U_k'' < 0$, and $U_k'(0) = \infty$ for all k. Security payoffs are determined by the time-1 state of the economy. There are a finite number of possible states, indexed by s. All investors agree on the state probabilities $\{\pi_s\}$, which are all positive.[3] The security market is complete—there is a set of s securities with linearly independent payoffs across states.

It is well known that equilibrium prices in this setting are characterized by the existence of "pure security" prices $\{p_s\}$, where p_s is the time-0 price of a security paying \$1 in state s and \$0 in all other states.[4] These pure securities may or may not exist as traded securities, but any traded security's time-0 price y^0 will satisfy

$$y^0 = \textstyle\sum_s y_s p_s, \tag{1}$$

where y_s is the traded security's time-1 payoff in state s. Let M_s be the dollar value of the market portfolio (aggregate wealth) in state s at time 1. Then states with higher aggregate wealth will have lower pure-security prices per unit of probability (i.e., per unit of expected payoff): for any two states s and s^*,

$$M_s < M_{s^*} \qquad \text{if and only if} \qquad \frac{p_s}{\pi_s} > \frac{p_{s^*}}{\pi_{s^*}}. \tag{2}$$

Also, every investor will have the same ranking of states in order of his or her time-1 wealth:

$$M_s < M_{s^*} \qquad \text{implies} \qquad W_{sk}^1 < W_{s^*k}^1 \qquad \text{for all } k. \tag{3}$$

[2] SuperShares® is a registered trademark of Leland O'Brien Rubinstein Associates, Inc. They are similar to Hakansson's (1977, 1978) supershares in the way they make the security market more complete, but their exact payoff patterns are different.

[3] The assumption of homogeneous beliefs is stronger than required for this paper's results, but it allows a simpler exposition. A weaker sufficient condition is homogeneity of *conditional* beliefs (conditional on M, the market payoff).

[4] This model determines relative prices but not the overall price level, and thus not the interest rate, which can be set arbitrarily.

A consequence of (2) is that payoffs occurring in states with the same aggregate wealth have the same present value per unit of expected payoff and so can be grouped together in calculation (1) as follows. Letting m vary across all possible realizations of $\tilde{M}$,

$$\begin{aligned} y^0 &= \textstyle\sum_s y_s p_s \\ &= \textstyle\sum_m \sum_{\{s:M_s=m\}} y_s p_s \\ &= \textstyle\sum_m \sum_{\{s:M_s=m\}} (p_s/\pi_s)(y_s\pi_s) \\ &= \textstyle\sum_m g(m) \sum_{s:M_s=m\}} (y_s\pi_s/\text{prob}\{\tilde{M} = m\}) \\ &= \textstyle\sum_m g(m)E\{\tilde{y} \mid \tilde{M} = m\}, \end{aligned} \tag{4}$$

where $g(m) = \sum_{\{s:M_s=m\}} p_s = (p_s/\pi_s)\text{prob}\{\tilde{M} = m\}$ for any s such that $M_s = m$. In the case of a security with expected payoffs that are linear conditional on the market's payoffs, if

$$E\{\tilde{y} \mid \tilde{M} = m\} = a + bm, \tag{5}$$

then

$$\begin{aligned} y^0 &= \textstyle\sum_m g(m)(a + bm) \\ &= a \textstyle\sum_m g(m) + b \sum_m mg(m) \\ &= a \textstyle\sum_s p_s + b \sum_s M_s p_s \\ &= \frac{a}{1 + r} + bM^0 \end{aligned} \tag{6}$$

where r is the riskless rate and M^0 is the time-0 value of the market portfolio.

To translate the above result into rates of return, denote the return on the complex security by $\tilde{R}_i = (\tilde{y}/y^0) - 1$ and the market return by $\tilde{R}_m = (\tilde{M}/M^0) - 1$. Then, if

$$E\{\tilde{R}_i - r \mid \tilde{R}_m = R_m\} = \alpha_i + \beta_i(R_m - r), \tag{7}$$

then

$$\alpha_i = 0$$

so that

$$\mu_i - r = \beta_i(\mu_m - r). \tag{8}$$

This shows that equilibrium pricing will be the same as in the Sharpe-Lintner CAPM for any security with a linear conditional expectation relationship with the market, *even if other securities are not conditionally linear* and *even if the optimal allocation of wealth is not linear in the market payoff.* Dybvig and Ingersoll (1982) prove a similar result in a mean-variance framework, but mean-variance analysis is not required for the above. (Also, notice that joint normality of security returns is a special case of equation (7).)

Let $\tilde{f}$ denote a random *vector* $(\tilde{f}_1, \tilde{f}_2, \cdots, \tilde{f}_\kappa)$, called the factors. Let $f_{\lambda s}$ be the realization of the random-variable factor $\tilde{f}_\lambda$ in state s, and let f_s be the vector realization of $\tilde{f}$ in state s, that is $f_s = (f_{1s}, f_{2s}, \cdots, f_{\kappa s})$. Let f be a κ-vector variable, which we will use as an index, running across all possible realizations $\tilde{f}$.

Connor's main assumption is a linear factor structure for security payoffs. Let us adopt the simplified terminology "(time-1 payoff) $\tilde{y}$ is linear in $\tilde{f}$" to mean that there exist nonrandom scalar a and vector b such that

$$E\{\tilde{y} \mid \tilde{f} = f\} = a + b'f \qquad \text{for all possible realizations } f \text{ of } \tilde{f}. \tag{9}$$

Then Connor's assumption is that all time-1 security payoffs are linear in $\tilde{f}$. Connor's main result is that time-0 security prices are linear with the same coefficients:

$$y^0 = \frac{a}{1 + r} + b'f^0, \tag{10}$$

where $f^0 = \sum_s p_s f_s$ is the vector of time-0 prices of portfolios with payoffs equalling the factors $\tilde{f}$ in every state.[5] We will describe this as "$\tilde{f}$ gives accurate linear pricing." Important among Connor's other assumptions are the conditions that the market portfolio's payoff $\tilde{M}$ is a linear function of $\tilde{f}$ ($\tilde{M} = a_m + b_m'\tilde{f}$) and that for each factor $\tilde{f}_\lambda$ there exists a mimicking portfolio with payoffs $\tilde{y}$ such that $y_s = f_{\lambda s}$ for all states s.

In the complete-markets context of the present paper, we can weaken the condition of $\tilde{M}$ being a linear function of $\tilde{f}$ and find a necessary and sufficient condition for accurate linear pricing by $\tilde{f}$. The condition is that the market portfolio payoff $\tilde{M}$ can be *any* function of the factors; that is, knowing the realization of $\tilde{f}$ must tell you the realization of $\tilde{M}$. This relation between two random variables will recur often in this paper, so let us describe it by the simplified terminology "$\tilde{f}$ refines $\tilde{M}$." This would mean that, for any two states s and s^*, if $f_s = f_{s^*}$ then $M_s = M_{s^*}$. It also implies that

$$\operatorname{var}\{\tilde{M} \mid \tilde{f} = f\} = 0 \qquad \text{for all } f. \tag{11}$$

(Some readers may prefer to mentally substitute this last condition in place of "$\tilde{f}$ refines $\tilde{M}$" in the rest of this paper.) It makes intuitive sense that $\tilde{f}$ must determine $\tilde{M}$ in order for conditional pricing to work, because in this economy the present value of expected payoffs is linked to the market payoff by (2).

Theorem 1, below, shows that $\tilde{f}$ gives accurate *nonlinear* pricing of *any* payoff pattern if and only if $\tilde{f}$ refines $\tilde{M}$. Theorem 2 shows that $\tilde{f}$ gives accurate linear pricing of payoff patterns that are linear in $\tilde{f}$, if and only if $\tilde{f}$ refines $\tilde{M}$, even if not all securities' payoffs are linear in $\tilde{f}$. See Appendix for all proofs.

THEOREM 1 (Nonlinear Conditional Pricing): *For any vector of factors $\tilde{f}$, there is a present value function $h(f)$ such that, for any payoff pattern $\tilde{y}$, its present value y^0 satisfies*

$$y^0 = \textstyle\sum_f h(f) E\{\tilde{y} \mid \tilde{f} = f\} \tag{12}$$

if and only if $\tilde{f}$ refines $\tilde{M}$.

[5] Connor sets the interest rate to zero (see footnote 4).

THEOREM 2 (Linear Conditional Pricing): *$\tilde{f}$ gives accurate linear pricing for all securities with payoffs linear in $\tilde{f}$ if and only if $\tilde{f}$ refines $\tilde{M}$.*[6]

We have seen that, in a complete-markets single-period model, equilibrium pricing is closely linked to the market portfolio. So what role is played by "factors" other than the market? The next section addresses this question by projecting each factor onto the market-payoff random variable. For securities that were linear conditional on the original factors, the projected factors work just as well, but the projected factors have the potential weakness of mispricing some securities that were not linear with the original factors but are linear with the projected factors.

For any vector of factors $\tilde{f}$, define a new vector of factors $\tilde{\phi}$ as follows:

$$\phi_s \equiv E\{\tilde{f} \mid \tilde{M} = M_s\}. \tag{13}$$

Thus, $\tilde{\phi}$ is the pointwise projection of $\tilde{f}$ onto $\tilde{M}$. Note that this definition implies that $\tilde{M}$ refines $\tilde{\phi}$ and that $E\{\tilde{f} \mid \tilde{\phi} = \phi\} = \phi$.

LEMMA 1: *$\tilde{\phi}$ has the same vector of present values (and thus the same vector of factor premia) as $\tilde{f}$:*

$$\phi^0 = f^0. \tag{14}$$

LEMMA 2: *If and only if $\tilde{f}$ refines $\tilde{\phi}$, any security that is linear in $\tilde{f}$ is also linear in $\tilde{\phi}$, and with the same coefficients; that is, if*

$$E\{\tilde{y} \mid \tilde{f} = f\} = a + b'f, \tag{15}$$

then

$$E\{\tilde{y} \mid \tilde{\phi} = \phi\} = a + b'\phi. \tag{16}$$

THEOREM 3: *If $\tilde{f}$ refines $\tilde{M}$, any security that satisfies linear pricing on a vector of factors $\tilde{f}$ will satisfy linear pricing on $\tilde{\phi}$ (the projection of $\tilde{f}$ onto the market), with*

[6] Since Theorem 2 is a foundation for much of what follows, it is worth reiterating the dependence of these results on the assumptions of homogeneous beliefs (or just homogeneous conditional on the market), supershare spanning, and state-independent utility. To illustrate, the following is a counterexample to Theorem 2 for state-dependent utility. All investors are identical in utility functions and endowments, so there is no trading. There are four equiprobable states. The equilibrium pure security prices p_s calculated below have been normalized to sum to 1.

	State 1	State 2	State 3	State 4
W^1 (Wealth at $t = 1$)	1	1	2	2
Utility function	$\log(W^1)$	$\log(W^1)$	$\log(W^1)$	$-e^{-W^1}$
$U'(W^1)$	1	1	½	$1/e^2$
p_s	0.37946	0.37946	0.18973	0.05135
f_s	1	1	2	2
y_s	0	2	1	3

The single factor $\tilde{f}$ posited above refines the market, since it is the market. However, even though payoff $\tilde{y}$ is linear in $\tilde{f}$ (since $E\{\tilde{y} \mid \tilde{f} = f\} = f$), it is not accurately priced by $\tilde{f}$:

$$f^0 = \textstyle\sum_s p_s f_s = 1.24108,$$

$$y^0 = \textstyle\sum_s p_s y_s = 1.1027 \neq f^0.$$

the same coefficients and factor premia; that is, if

$$E\{\tilde{y} \mid \tilde{f} = f\} = a + b'f \text{ for all } f \tag{17}$$

and

$$y^0 = \frac{a}{1 + r} + b'f^0, \tag{18}$$

then

$$E\{\tilde{y} \mid \tilde{\phi} = \phi\} = a + b'\phi \text{ for all } \phi \tag{19}$$

and

$$y^0 = \frac{a}{1 + r} + b'\phi^0. \tag{20}$$

To summarize, if $\tilde{f}$ refines $\tilde{M}$ then its projection $\tilde{\phi}$ gives the same conditional-linear fit for time-1 payoffs of all securities linear in $\tilde{f}$, and both $\tilde{f}$ and $\tilde{\phi}$ price these securities accurately. So if all existing securities have expected payoffs linear in $\tilde{f}$, as in Connor (1984), then the projected factors $\tilde{\phi}$ work just as well. Therefore, for the purposes of pricing, there is no role for factors with dimensions of risk other than that of the market portfolio. The only contribution multiple factors can make is to linearize conditional payoffs that are nonlinear in the market; this is done by having factors that are simply nonlinear functions of the market payoff.

All the factor-pricing results in the present paper can be applied selectively to those securities with the required conditional payoff characteristics, while allowing that other securities may exist. This raises the question of whether the projected factors $\tilde{\phi}$ are useful in pricing securities that are nonlinear conditional on the original factors $\tilde{f}$. First we will show that there will always be payoff patterns linear in $\tilde{\phi}$ but not in $\tilde{f}$, unless $\tilde{\phi} = \tilde{f}$. Then we will present conditions under which $\tilde{\phi}$ gives accurate linear pricing of those patterns.

LEMMA 3:

$$\tilde{\phi} = \tilde{f} \tag{21}$$

if and only if $\tilde{M}$ refines $\tilde{f}$.

THEOREM 4: *In the case where $\tilde{f}$ refines $\tilde{M}$, there exist payoff patterns $\tilde{y}$ linear in $\tilde{\phi}$ but not linear in $\tilde{f}$, if and only if*

$$\tilde{f} \neq \tilde{\phi}. \tag{22}$$

Therefore, there can be many securities nonlinear in $\tilde{f}$ that become linear in its projection $\tilde{\phi}$. However, while most of them will be accurately priced by $\tilde{\phi}$, some may not be. Applying Theorem 3 to $\tilde{\phi}$, $\tilde{\phi}$ will accurately price all conditionally linear securities if *and only if* $\tilde{\phi}$ refines $\tilde{M}$, a condition which is not guaranteed by the fact that $\tilde{f}$ refines $\tilde{M}$. The following two examples illustrate this. (States are equiprobable in both.)

In Table I, $\tilde{\phi}$ refines $\tilde{M}$—there are no two states with different market levels

Table I
Example of Projected Factor $\tilde{\phi}$ Refining the Market Payoff $\tilde{M}$

	States→	$s1$	$s2$	$s3$	$s4$
Market payoff	$\tilde{M}$:	1	1	2	2
Original factor	$\tilde{f}$:	3	5	4	6
Projected factor	$\tilde{\phi}$:	4	4	5	5

and the same ϕ, so knowing ϕ tells you M. So ϕ correctly prices not only securities linear in $\tilde{f}$, but all others that are linear in $\tilde{\phi}$, such as

$$\tilde{y}: \quad 4 \quad 4 \quad 4 \quad 6.$$

Recognizing that state prices are the same when market levels are the same, a simpler and more powerful factor has been obtained by projecting on $\tilde{M}$.

In Table II, $\tilde{\phi}$ does not refine $\tilde{M}$, even though $\tilde{f}$ did. So, as shown in the "only if" proof of Theorem 2, there are payoff patterns linear in $\tilde{\phi}$ that do not fit the corresponding linear present-value relation, such as

$$\tilde{y}: \quad 4 \quad 3 \quad 5 \quad 4.$$

$E\{\tilde{y} \mid \phi\} = \phi$, but the present value $y^0 < \phi^0$ because $\tilde{y}$ is tilted more toward higher M values than $\tilde{\phi}$ is.[7] Notice that $\tilde{y}$ is not conditionally linear in $\tilde{f}$, although nonlinear pricing on $\tilde{f}$ would still be accurate by Theorem 1, since $\tilde{f}$ refines $\tilde{M}$.

However, in the case of Connor's model this problem does not arise. If the market portfolio is a linear function of $\tilde{f}$,

$$\tilde{M} = a + b'\tilde{f}, \tag{23}$$

then it will be the same linear function of $\tilde{\phi}$: For any state s,

$$a + b'\phi_s = a + b'(E\{\tilde{f} \mid \tilde{M} = M_s\}) \tag{24}$$

$$= E\{a + b'\tilde{f} \mid \tilde{M} = M_s\} \tag{25}$$

$$= E\{\tilde{M} \mid \tilde{M} = M_s\} \tag{26}$$

$$= M_s. \tag{27}$$

Thus, $\tilde{\phi}$ refines $\tilde{M}$ and, by Theorem 2, gives accurate linear pricing for all conditionally linear payoffs. Therefore, if we project Connor's factors onto $\tilde{M}$ so that each factor is a nonlinear function of the market, we lose nothing while gaining the potential of linearizing more payoff patterns, plus the simplicity of having no "diversifiable" (nonmarket) risk in the factors, which rational investors would not want to hold anyway.

It is interesting to ask what kind of factors it would take to make all securities conditionally linear in a complete market. The answer is that the factors must be rotationally equivalent to a set of "quasi-pure" securities paying off in "quasi-states," where a quasi-state is a group of states with the groups being mutually exclusive and collectively exhausting the state space. Thus, in each state, one

[7] This kind of counterexample would be impossible if the support of $\tilde{f}$ were a convex set.

Table II

Example of Projected Factor $\tilde{\phi}$ *Not* Refining the Market Payoff $\tilde{M}$

	States→	$s1$	$s2$	$s3$	$s4$
Market payoff	$\tilde{M}$:	1	1	2	2
Original factor	$\tilde{f}$:	3	5	2	6
Projected factor	$\tilde{\phi}$:	4	4	4	4

rotated factor has the value one and all the others are zero. Table III shows an example of this.

This also implies that the κ-dimensional random vector $\tilde{f}$ can only have κ possible distinct realizations—an extremely restrictive condition.[8]

For simplicity let us assume at this point that there are no redundant factors—no factor can be replicated by a linear combination of the others. Thus, each security's factor loading is unique. Also, let factor number one be a constant, $f_{1s} = 1$ for all s; there are still κ factors, so now only $\kappa - 1$ of them are risky. The constant term "a" in a linear combination of factors becomes the first element of the vector "b".

THEOREM 5: *For any payoff pattern $\tilde{y}$, there exists a vector b such that*

$$E\{\tilde{y} \mid \tilde{f} = f\} = b'f \tag{28}$$

for all possible realizations f of $\tilde{f}$, if and only if $\tilde{f}$ spans the same space of payoff patterns as a set of κ dummy variables $\tilde{\psi}$; that is, if and only if there is a nonsingular κ-by-κ matrix A such that $\tilde{\psi} \equiv A\tilde{f}$ has its realization ψ_s in any state s being a vector of all zeroes and a single one.

Theorem 5 says that the only way to get conditional expected payoffs of all securities to be linear in κ factors is to partition the state space into κ subsets and define each factor to be a "dummy variable" for a subset. Theorem 2 said that the only way to get accurate pricing of conditionally linear payoffs is to have $\tilde{f}$ refine $\tilde{M}$, that is, to ensure that no two levels of M are associated with the same factor-vector realization f. If we put these two results together, we see that the only way to get both linearity and accurate pricing is to use Hakansson's supershares as factors. That is, partition the state space by levels of M, so that each subset contains all states with a given M value (these subsets are called "superstates"); define κ "quasi-pure" securities (supershares), each paying off \$1 in its superstate and \$0 otherwise. Any (nonsingular) rotation of these payoffs will give accurate linear pricing. Factors based on finer partitions of the state space would also give accurate linear pricing, but the refinement would increase complexity with no improvement in linearity or pricing accuracy.

THEOREM 6: *For every payoff pattern $\tilde{y}$ there exists a vector b such that*

$$E\{\tilde{y} \mid \tilde{f} = f\} = b'f \tag{29}$$

[8] One way to satisfy this restrictive condition would be to augment the set of factors to include new factors that are "general or multiple options" on the existing factors, as defined (and denoted "M") in Ross (1976a). Such a general or multiple option would be any mapping from the vector of existing-factor outcomes to a scalar.

Table III

Rotated Factors as "Quasi-Pure" Securities

	States→	$s1$	$s2$	$s3$	$s4$	$s5$	$s6$
Rotated factors:	$rf1$→	1	1	0	0	0	0
	$rf2$→	0	0	1	1	1	0
	$rf3$→	0	0	0	0	0	1
	"quasi-states"→	—$qs1$—		——$qs2$——			$qs3$

for all f, and

$$y^0 = b'f^0, \tag{30}$$

if and only if $\tilde{f}$ is rotationally equivalent to Hakansson's supershares or to a refinement of supershares.

The results of this section show that, in the single-period complete-markets model, the only useful role for factors is to provide different nonlinear functions of the market portfolio, so that securities that are nonlinear conditional on the market may be linear conditional on the factors. This contrasts sharply with most discussions of APT empirical work, which emphasize different sources and dimensions of risk. What pricing effects do we really think factors are going to capture, if not just nonlinearity conditional on the market? Remaining possibilities include intertemporal effects, missing assets, market segmentation, taxes, and transaction costs. Each of these has different implications about how empirical work should be conducted and about how it should be applied to choosing optimal portfolios. The next section analyzes the case of intertemporal effects caused by stochastic investment opportunities.

II. Intertemporal Models

Models with more than one future consumption date allow the possibility that equilibrium pricing may not be so simply linked to the return on the market portfolio. This may make for a richer role for factors other than functions of $\tilde{M}$. Connor and Korajczyk (1987) (CK) develop an "Intertemporal Equilibrium APT" based on assuming a linear factor structure for future dividends. While it is more appealing to make the dividend process exogenous than to make the security price (or return) process exogenous, this has the effect of making aggregate consumption exogenous since it equals aggregate dividends in the CK model. If we are willing to make assumptions directly about consumption and how it correlates with individual securities' dividends, we can apply some known results from the consumption-based asset pricing literature.

Following CK, consider an economy with a representative investor who at time t chooses a consumption and portfolio strategy to

$$\text{maximize } u(C_t) + E_t\left\{\textstyle\sum_{\tau=1}^{\infty} \rho^\tau u(\tilde{C}_{t+\tau})\right\},$$

where $u(C_t)$ is utility of time-t consumption and ρ is an impatience parameter. As in the single-period model, $u'(\cdot) > 0$, $u''(\cdot) < 0$, and $u'(0) = \infty$. We assume spanning of Hakansson-type supershares on aggregate consumption at each date but do the analysis as if markets were fully complete since that equilibrium is the same.[9] Then equilibrium prices are characterized by pure security prices $p_{t+\tau,s}$, being the time-t price, denominated in the consumption good, for receiving one unit of the good at time $t + \tau$ in state s. The consumer's first-order condition determines that

$$p_{t+\tau,s} = \pi_{t+\tau,s}\rho^{\tau}u'(C_{t+\tau,s})/u'(C_t). \tag{31}$$

In this multiperiod setting, first consider the subset of securities that happen to pay dividends only at date $t + \tau$. Then all the single-period results from Section I above have counterparts here, with $\tilde{C}_{t+\tau}$ now playing the role of the market portfolio's payoff ($\tilde{M}$), which was equal to consumption in the single-period model. The counterparts are listed below, with "i" (for intertemporal) appended to the theorem numbers.

The time-t value v_t of a security paying dividend $\tilde{x}_{t+\tau}$ is

$$v_t = \textstyle\sum_s x_{t+\tau,s}p_{t+\tau,s} \tag{32}$$

$$= \textstyle\sum_c g(c,\ t + \tau)E_t\{\tilde{x}_{t+\tau} \mid \tilde{C}_{t+\tau} = c\}, \tag{33}$$

where

$$g(c,\ t + \tau) = \textstyle\sum_{\{s:C_{t+\tau,s}=c\}} p_{t+\tau,s}. \tag{34}$$

(This was shown by Breeden and Litzenberger (1978).) In the case of a security with time-$(t + \tau)$ dividends linear in consumption, if

$$E_t\{\tilde{x}_{t+\tau} \mid \tilde{C}_{t+\tau} = c\} = a + bc, \tag{35}$$

then

$$v_t = aB(t,\ t + \tau) + b\gamma(t,\ t + \tau), \tag{36}$$

where

$$B(t,\ t + \tau) = \textstyle\sum_s p_{t+\tau,s} \tag{37}$$

is a time-t unit bond price and

$$\gamma(t,\ t + \tau) = \text{time-}t \text{ value of } \tilde{C}_{t+\tau} \tag{38}$$

$$= \textstyle\sum_s C_{t+\tau,s}\, p_{t+\tau,s}. \tag{39}$$

THEOREM 1i (Nonlinear Conditional Pricing): *For any vector of factors $\tilde{f}$ with realizations at time $t + \tau$, there is a present value function $h(f)$ such that, for any time-$(t + \tau)$ dividend payoff pattern $\tilde{x}_{t+\tau}$, its present value v_t satisfies*

$$v_t = \textstyle\sum_f h(f)E_t\{\tilde{x}_{t+\tau} \mid \tilde{f} = f\} \tag{40}$$

if and only if $\tilde{f}$ refines $\tilde{C}_{t+\tau}$.

[9] Notice that the utility function is time additive and state independent; otherwise, the results might not hold. The assumption of homogeneous beliefs is maintained here, though again it could be weakened to homogeneity conditional on aggregate consumption.

THEOREM 2i (Linear Conditional Pricing): *$\tilde{f}$ gives accurate linear pricing for all securities paying dividends only at time $t + \tau$, with $\tilde{x}_{t+\tau}$ linear in $\tilde{f}$, if and only if $\tilde{f}$ refines $\tilde{C}_{t+\tau}$.*

Here, "$\tilde{x}_{t+\tau}$ linear in $\tilde{f}$" means that there exist scalar a and vector b such that

$$E_t\{\tilde{x}_{t+\tau} \mid \tilde{f} = f\} = a + b'f \tag{41}$$

for all realizations f of $\tilde{f}$. Also, "accurate linear pricing" means

$$v_t = aB(t, t + \tau) + b'F(t, t + \tau), \tag{42}$$

where

$$F(t, t + \tau) = \textstyle\sum_s p_{t+\tau,s} f_s \tag{43}$$

is the vector of time-t values of portfolios, with time-$(t + \tau)$ payoffs equaling the factors $\tilde{f}$ in every state.

Parallel to the single-period case, CK assume that $\tilde{C}_{t+\tau}$ is a linear function of the factors, which guarantees that $\tilde{f}$ refines $\tilde{C}_{t+\tau}$.

Now define the projected factors $\tilde{\phi}$ as

$$\phi_s = E_t\{\tilde{f} \mid \tilde{C}_{t+\tau,s}\}, \tag{44}$$

so $\tilde{C}_{t+\tau}$ refines $\tilde{\phi}$, and $E_t\}\tilde{f} \mid \tilde{\phi} = \phi\} = \phi$.

LEMMA 1i: *$\tilde{\phi}$ has the same vector of time-t values as $\tilde{f}$:*

$$\Phi(t, t + \tau) = F(t, t + \tau), \tag{45}$$

where

$$\Phi(t, t + \tau) = \textstyle\sum_s p_{t+\tau,s}\phi_s. \tag{46}$$

LEMMA 2i: *If and only if $\tilde{f}$ refines $\tilde{\phi}$, any time-$(t + \tau)$ dividend payoff pattern linear in $\tilde{f}$ is linear in $\tilde{\phi}$, and with the same coefficients.*

To summarize the remaining intertemporal counterpart theorems: if $\tilde{f}$ refines $\tilde{C}_{t+\tau}$, then the projected factors $\tilde{\phi}$ work just as well as $\tilde{f}$ for linear pricing of payoffs linear in $\tilde{f}$. More payoff patterns will be linear in $\tilde{\phi}$ than in $\tilde{f}$, except in the degenerate case where $\tilde{\phi} = \tilde{f}$, but some of these may not be accurately priced by $\tilde{\phi}$; however, that problem does not arise if $\tilde{C}_{t+\tau}$ is a linear function of $\tilde{f}$, as in CK. The only way for all payoff patterns to be linear in $\tilde{f}$ is for $\tilde{f}$ to be rotationally equivalent to "quasi-pure" securities on some partition of the state space, and the only way to have this linearity plus accurate pricing is for $\tilde{f}$ to be rotationally equivalent to "supershares" on consumption, or a refinement thereof.

Turning now to securities that may pay dividends in all future periods, we add up the present values of dividends from each future date. In the special case of a security with dividends linear in consumption with the same coefficients at each future date, if

$$E_t\{\tilde{x}_{t+\tau} \mid \tilde{C}_{t+\tau} = c\} = a + bc \qquad \text{for all} \quad \tau, c, \tag{47}$$

then

$$v_t = aB(t, *) + b\gamma(t, *), \tag{48}$$

where

$$B(t, *) = \textstyle\sum_{\tau=1}^{\infty} B(t, t + \tau) = \text{time-}t \text{ price of a consol} \tag{49}$$

and

$$\gamma(t, *) = \textstyle\sum_{\tau=1}^{\infty} \gamma(t, t + \tau) = \text{time-}t \text{ value of all future consumption.} \tag{50}$$

In the case of securities with dividends linear in a time series of factors $\tilde{f}_t$, where $\tilde{f}_t$ refines $\tilde{C}_t$ for each t, if

$$E_t\{\tilde{x}_{t+\tau} \mid \tilde{f}_{t+\tau} = f\} = a + b'f \qquad \text{for all} \quad \tau, f, \tag{51}$$

then

$$v_t = aB(t, *) + b'F(t, *), \tag{52}$$

where

$$F(t, *) = \textstyle\sum_{\tau=1}^{\infty} F(t, t + \tau) = \text{time-}t \text{ factor values.} \tag{53}$$

In the case (considered by CK) of a "zeroth" factor f_0 that is stochastic, affects dividends at time $t = \tau$, but becomes known one period earlier at time $t + \tau - 1$, it can be treated just like any other factor in the above valuation equations.

III. Implications for Empirical Work

The appropriate strategy for estimating and testing the APT (versus the CAPM) depends on which CAPM assumption is being violated. The violation considered in Connor (1984) is that mean and variance are not sufficient for portfolio choice or, alternatively, that security payoffs are not linear conditional on the market. The analysis in Section I above implies that, in such a case, the only kind of factors needed is nonlinear functions of the market portfolio. Existing APT empirical work is nothing like that. It uses multidimensional linear strategies such as factor analysis or principal components analysis.

Suppose instead that the CAPM fails to hold because of a stochastic investment opportunity set (intertemporal effects). Merton's (1973) ICAPM shows that securities with returns correlated with investment opportunity set shifts will tend to have hedging premia. However, standard methods focus on contemporaneous correlations, not serial correlations. There is no assurance that the serial correlations will be adequately captured in the factors rather than left in the epsilons, although one can try to make the epsilons as small as possible. This problem applies whether the theoretical model is stated in terms of Consumption-CAPM or ICAPM since the two are equivalent. Thus, besides ensuring zero correlation of epsilons with contemporaneous factors, we should try for zero correlation of epsilons with subsequent returns.

If the problem is that we lack data on some existing capital assets, then (as in the case of intertemporal effects) the arbitrage part of the APT still works: exact replication implies exact relative pricing. However, epsilons may be priced if they are correlated with returns on missing assets. Once again, finding factors that keep the epsilons small may reduce the model's inaccuracies. However, notice that standard empirical methodologies concentrate on correlations among assets with data and cannot try to ensure no correlation of epsilons with the missing data.

If incomplete integration of international capital markets is causing the CAPM not to hold, then, for example, the correlation of a U.S. stock with the Nikkei index may affect its pricing; however, factor analysis on U.S. equities alone may leave much of that correlation in the epsilons.

If there are tax effects, then, of course, the usual factor approaches are inappropriate. Differential taxation for different asset classes should be measured and allowed for directly.

Hand in hand with the above problems of pricing the epsilons is the question of whether an optimal portfolio may contain epsilon risk. Furthermore, addressing the issue of why the CAPM is not holding will lead to facing the problems of determining the equilibrium factor premia and the optimal portfolio factor loadings for specific types of investors.

IV. Summary and Conclusions

The possible roles for factors in a single-period model are not rich enough to provide a theoretical framework for the mainstream empirical APT research. This is clearest if we assume that securities span a complete set of options on the market portfolio or, equivalently, a complete set of Hakansson's "supershares." The market portfolio is so central to such a model that the following results are shown to hold. Linear factor pricing is exact if and only if the factors "refine" the market (i.e., knowing the factor outcomes tells you the market outcome; the variance of the market conditional on the factors is zero). If the factors refine the market, then they can be projected onto the market with no loss of asset-pricing power. Thus, the only "factors" needed are nonlinear functions of the market. Finally, to give accurate linear pricing of all payoff patterns, factors must be rotationally equivalent to Hakansson's supershares. A similar set of results holds in an intertemporal model but with consumption replacing the market payoff.

Under each of the many possible reasons why the Sharpe-Lintner CAPM may not hold, the current empirical APT focus on contemporaneous correlations of asset returns is too narrow. For example, if Merton's ICAPM holds instead, then in principle we should be looking for serial correlations in returns.

Appendix

Proof of Theorem 1:
"if":

Suppose $\tilde{f}$ refines $\tilde{M}$. Then any two states with $f_s = f_{s^*}$ will have $M_s = M_{s^*}$ and $p_s/\pi_s = p_{s^*}/\pi_{s^*}$. Let

$$h(f) = \textstyle\sum_{\{s:f_s=f\}} p_s = (p_s/\pi_s)\text{prob}\{\tilde{f} = f\} \tag{A1}$$

for any s such that $f_s = f$. Then any payoff pattern will satisfy

$$y^0 = \textstyle\sum_f h(f)E\{\tilde{y} \mid \tilde{f} = f\} \tag{A2}$$

by the same reasoning as for equation (4).

"only if":

Suppose there is such a function h. Consider any two states s and s^* such that $f_s = f_{s^*}$. Applying the present value function h to the pure securities for states s and s^* gives

$$p_s = h(f_s)\pi_s/\text{prob}\{\tilde{f} = f\} \tag{A3}$$

and

$$p_{s^*} = h(f_{s^*})\pi_{s^*}/\text{prob}\{\tilde{f} = f\}, \tag{A4}$$

implying $p_s/\pi_s = p_{s^*}/\pi_{s^*}$. Therefore, by (2), $M_s = M_{s^*}$, so $\tilde{f}$ refines $\tilde{M}$. Q.E.D.

Proof of Theorem 2:

"if":

By Theorem 1 any $\tilde{y}$ has present value

$$y^0 = \textstyle\sum_f h(f)E\{\tilde{y} \mid \tilde{f} = f\}, \tag{A5}$$

so any $\tilde{y}$ that is linear conditional on $\tilde{f}$ will have

$$y^0 = \textstyle\sum_f h(f)(a + b'f) \tag{A6}$$

$$= a \textstyle\sum_f h(f) + b' \sum_f fh(f) \tag{A7}$$

$$= a \textstyle\sum_s p_s + b' \sum_s f_s^1 p_s \tag{A8}$$

$$= a/(1 + r) + b'f^0. \tag{A9}$$

"only if":

Supposing $\tilde{f}$ gives accurate linear pricing of linear payoffs, but $\tilde{f}$ does not refine $\tilde{M}$, then we will derive a contradiction. Nonrefinement implies that there are two states $s1$ and $s2$ such that $f_{s1} = f_{s2}$ and $M_{s1} < M_{s2}$, so by (2) we have

$$p_{s1}/\pi_{s1} > p_{s2}/\pi_{s2}. \tag{A10}$$

Choose any one of the factors, $\tilde{f}_\lambda$. Consider the payoff pattern $\tilde{y}$ that has realizations identical to those of $\tilde{f}_\lambda$ except in states $s1$ and $s2$, in which

$$y_{s1} = f_{\lambda s1} + \pi_{s2} \tag{A11}$$

and

$$y_{s2} = f_{\lambda s2} - \pi_{s1}. \tag{A12}$$

Then

$$E\{\tilde{y} \mid \text{state } s1 \text{ or } s2\} = \frac{\pi_{s1}(f_{\lambda s1} + \pi_{s2}) + \pi_{s2}(f_{\lambda s2} - \pi_{s1})}{\pi_{s1} + \pi_{s2}} = f_{\lambda s1} = f_{\lambda s2}, \tag{A13}$$

so y is linear conditional on $\tilde{f}$:

$$E\{\tilde{y} \mid \tilde{f} = f\} = f_\lambda, \tag{A14}$$

since $\tilde{f}$ does not distinguish state $s1$ from state $s2$. Thus, $\tilde{f}$ should give the present

1970. Since we need one year's lag in order to construct a 1-year moving average dividend-price ratio, our sample period starts in January 1971.

Japanese bond markets did not develop until the 1970's, and data are therefore not available before 1970. There is no equivalent of Treasury bills in Japan; thus the short-term interest rate used here is a combined series of the call money rate (1971:1–1977:11) and the Gensaki rate (1977:12–1990:3). The Gensaki rate, an interest rate applied to bond repurchase agreements, is less subject to regulation than the call money rate, but it became available only after 1977. The call money rate is the "unconditional" rate, which is applied to transactions maturing in less than one month, and we use a Gensaki rate with one month maturity. For the long-term Japanese government yield, we use a value-weighted index of yields on bonds with 9 to 10 years to maturity.[10]

We also use one piece of data from outside the national financial markets of the U.S. and Japan. This is the monthly return on the Morgan Stanley Capital International (MSCI) World Index, a market-value-weighted index covering just under 1500 companies listed on the stock exchanges of 20 countries. Together, these companies account for about 60% of the total market capitalization of the countries included in the index. At the end of September 1990 the U.S. market had a weight of 35%, the Japanese market had a weight of 30%, and the European stock markets had a combined weight of 28% in the index.[11] The MSCI world index is measured in dollar terms, inclusive of dividends.

Finally, we note that in forming excess return series, we measure both U.S. and Japanese stock returns in dollars, relative to the U.S. Treasury bill rate. In earlier versions of this paper, we measured the Japanese stock return in yen, relative to the Japanese short-term interest rate.[12] Excess yen returns on Japanese stocks are slightly more predictable than excess dollar returns on Japanese stocks, but the difference is small and does not affect the qualitative results of the paper. We also measure all returns in continuously compounded (log) form. This is common practice in empirical work on asset pricing, and it has the advantage that it enables us to use excess returns without measuring a dollar or yen price deflator. However it may introduce some approximation error in that asset pricing models generally apply to simple rather than log returns.[13]

[10]This is the longest consistently available maturity. The yield index normally includes a benchmark issue, which has a large capitalization and therefore a high weight in the index.

[11]Earlier in our sample period the U.S. had a higher weight and Japan had a lower weight. The prevalence of cross-holding in the Japanese market tends to give it too high a weight in international value-weighted indices.

[12]The excess Japanese stock return, measured in yen, approximates the return that U.S. investors will receive if they finance their investment by borrowing yen. The approximation becomes accurate in the limit of continuous time if the yen return on Japanese stocks is uncorrelated with the dollar price of the yen (Stulz (1981)).

[13]As Bekaert and Hodrick (1990) point out, there can be systematic differences in expected log returns on different assets even if investors are risk-neutral; these result from Jensen's Inequality. Thus it is strictly speaking incorrect to refer to expected excess log returns as "risk premia."

since $\tilde{f}$ does not distinguish state $s1$ from state $s2$. However,

$$E\{\tilde{y} \mid \tilde{\phi} = \phi_{s1}\} = E\{\tilde{f} \mid \tilde{\phi} = \phi_{s1}\} + \pi_{s1}\pi_{s2}/\text{prob}\{\tilde{\phi} = \phi_{s1}\} \tag{A32}$$

$$= \phi_{s1} + \pi_{s1}\pi_{s2}/\text{prob}\{\tilde{\phi} = \phi_{s1}\} \tag{A33}$$

$$> \phi_{s1}. \tag{A34}$$

Q.E.D.

Proof of Theorem 3: If $\tilde{f}$ refines $\tilde{M}$, then $\tilde{f}$ refines $\tilde{\phi}$ (because $\tilde{M}$ refines $\tilde{\phi}$). Then apply Lemma 2 and Lemma 1.

Proof of Lemma 3:
"if":
$\tilde{M}$ refines $\tilde{f}$ implies that the distribution of $\tilde{f}$ conditional on any value of $\tilde{M}$ is a singleton—zero variance. So for any state s,

$$\phi_s \equiv E\{\tilde{f} \mid \tilde{M} = M_s\} = f_s. \tag{A35}$$

"only if":
$\tilde{\phi} = \tilde{f}$ means that, for any state s,

$$f_s = \phi_s \equiv E\{\tilde{f} \mid \tilde{M} = M_s\}. \tag{A36}$$

Consider any two states s and s^* with $M_s = M_{s^*}$. Then

$$f_s = E\{\tilde{f} \mid \tilde{M} = M_s\} = E\{\tilde{f} \mid \tilde{M} = M_{s^*}\} = f_{s^*}. \tag{A37}$$

Thus, $\tilde{M}$ refines $\tilde{f}$. Q.E.D.

Proof of Theorem 4:
"if":
$\tilde{f} \neq \tilde{\phi}$ implies that $\tilde{M}$ does not refine $\tilde{f}$, so there are states s and s^* such that $M_s = M_{s^*}$ and $f_{\lambda s} \neq f_{\lambda s^*}$ for some factor number λ. Consider payoff pattern $\tilde{y} = \tilde{\phi}_\lambda$. Then

$$E\{\tilde{y} \mid \tilde{\phi} = \phi\} = \phi_\lambda, \tag{A38}$$

so $\tilde{y}$ is linear in $\tilde{\phi}$. $\tilde{f}$ refines $\tilde{M}$ and $\tilde{M}$ refines $\tilde{\phi}$, so $\tilde{f}$ refines $\tilde{\phi}$; thus, by Lemma 2, *if* $\tilde{y}$ were linear in $\tilde{f}$ then

$$E\{\tilde{y} \mid \tilde{f} = f\} = f_\lambda. \tag{A39}$$

We will derive a contradiction from this. If $E\{\tilde{\phi}_\lambda \mid \tilde{f} = f\} = f_\lambda$ for all possible outcomes f, then

$$E\{\tilde{\phi}_\lambda \mid \tilde{f} = f_s\} = f_{\lambda s} \tag{A40}$$

and

$$E\{\tilde{\phi}_\lambda \mid \tilde{f} = f_{s^*}\} = f_{\lambda s^*}. \tag{A41}$$

$\tilde{M}$ refines $\tilde{\phi}$, so $\phi_{\lambda s} = \phi_{\lambda s^*}$. $\tilde{f}$ refines $\tilde{\phi}$, so the distribution of $\tilde{\phi}$ conditional on any value of $\tilde{f}$ is a singleton, so

$$E\{\tilde{\phi}_\lambda \mid \tilde{f} = f_s\} = \phi_{\lambda s} \tag{A42}$$

and

$$E\{\tilde{\phi}_\lambda \mid \tilde{f} = f_{s^*}\} = \phi_{\lambda s^*}. \tag{A43}$$

Linking up the last five equations gives $f_{\lambda s} = f_{\lambda s^*}$, a contradiction. [There must be an easier proof!]

"only if":

$\tilde{f} = \tilde{\phi}$ implies that any $\tilde{y}$ linear in $\tilde{\phi}$ must be linear in $\tilde{f}$. Q.E.D.

Proof of Theorem 5:

"if":

Assuming such an A exists, let $\tilde{y}$ be the payoff on pure security s^*, so $y_{s^*} = 1$, and $y_s = 0$ for $s \neq s^*$;

$$E\{\tilde{y} \mid \tilde{f} = f_{s^*}\} = \pi_{s^*}/\text{prob}\{\tilde{f} = f_{s^*}\}; \tag{A44}$$

$$E\{\tilde{y} \mid \tilde{f} = f\} = 0 \qquad \text{for all} \quad f \neq f_{s^*}. \tag{A45}$$

Let λ^* be the index corresponding to the unique rotated factor $\tilde{\psi}_{\lambda^*}$ that takes the value 1 in state s^*. Denote the λ^*th row of A as $(A_{\lambda^*})'$, a row-vector the transpose of which is A_{λ^*}. Then,

$$E\{\tilde{y} \mid \tilde{f} = f\} = b'f, \tag{A46}$$

where

$$b = [\pi_{s^*}/\text{prob}\{\tilde{f} = f_{s^*}\}] \times A_{\lambda^*}. \tag{A47}$$

To see this, note that, since $\tilde{f} = A^{-1}\tilde{\psi}$, the only states where $\tilde{\psi}_{\lambda^*}$ takes the value 1 are those with $f_s = f_{s^*}$, so $(A_{\lambda^*})'f = 1$ only when $f = f_{s^*}$, and zero otherwise. Therefore, all pure securities are conditionally linear in $\tilde{f}$. Any other payoff pattern is a linear combination of pure security payoffs, so it will also be conditionally linear in $\tilde{f}$. [Another, perhaps more enlightening way to prove this is to show that each pure security's $\tilde{y}$ is linear in $\tilde{\psi}$ and, hence, linear in $\tilde{f} = A^{-1}\tilde{\psi}$.]

"only if":

There may be some groups of the S states where the random vector $\tilde{f}$ has the same realization f_s for every state in the group. Let μ ($\leq S$) be the number of distinct possible realizations of $\tilde{f}$. Let Q be any set of μ states, each having a different realization of $\tilde{f}$. An important part of the proof will be to show that, if all patterns $\tilde{y}$ are linear in $\tilde{f}$, then $\mu = \kappa$. μ cannot be less than κ because, thinking of $\tilde{f}$ as a κ-by-S matrix,

$$\kappa = \text{dimension of } \tilde{f}\text{'s rowspace} = \text{dimension of } \tilde{f}\text{'s columnspace} \leq \mu. \tag{A48}$$

Construct a μ-by-κ matrix A by letting each row be the transpose of $[\text{prob}\{\tilde{f} = f_s\}/\pi_s] \times b$, where b is such that, for the payoff $\tilde{y}$ on pure security number s,

$$E\{\tilde{y} \mid \tilde{f} = f\} = b'f, \tag{A49}$$

and s ranges over all states in set Q. Let $\tilde{\psi} \equiv A\tilde{f}$. Imagine shrinking the matrices $\tilde{f}$ and $\tilde{\psi}$ from having S columns to having μ columns, by crossing out all columns except for those corresponding to states in Q. (Note that, since only duplicate columns are being crossed out, both for $\tilde{f}$ and for $\tilde{\psi}$, the columnspaces of each are unaffected.) Then we will show that the shrunken $\tilde{\psi}$ (which is now μ-by-μ) is an identity matrix.

For any two states s1 and s2 in Q, the (s1, s2) element of the shrunken matrix $\tilde{\psi}$ is

$$[\text{prob}\{\tilde{f} = f_{s1}\}/\pi_{s1}] \times E\{\tilde{y} \mid \tilde{f} = f_{s2}\}, \tag{A50}$$

where $\tilde{y}$ is the payoff on pure security #s1. If s1 = s2, then

$$E\{\tilde{y} \mid \tilde{f} = f_{s2}\} = \pi_{s1}/\text{prob}\{f = f_{s1}\}, \tag{A51}$$

so the diagonal elements all equal 1. If s1 $\neq$ s2, then $f_{s1} \neq f_{s2}$ by the construction of Q, so

$$E\{\tilde{y} \mid \tilde{f} = f_{s2}\} = 0. \tag{A52}$$

Therefore, the shrunken $\tilde{\psi}$ is the μ-by-μ identity matrix. Thus,

$$\mu = \text{dimension of columnspace of } \tilde{\psi} \tag{A53}$$

$$= \text{dimension of rowspace of } \tilde{\psi} \tag{A54}$$

$$\leq \text{dimension of rowspace of } \tilde{f} \text{ [since } \tilde{\psi} = A\tilde{f}\text{]} \tag{A55}$$

$$= \kappa, \tag{A56}$$

so we must have $\mu = \kappa$. Thus, the shrunken $\tilde{f}$ has dimensions κ by κ and is the inverse of A, and each column of $\tilde{\psi}$ has a one and the rest zeroes. Q.E.D.

Proof of Theorem 6: This is immediate from Theorems 2 and 5.

REFERENCES

Breeden, Douglas, 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.

——— and Robert Litzenberger, 1978, Prices of state-contingent claims implicit in option prices, *Journal of Business* 51, 621–651.

Connor, Gregory, 1984, A unified beta pricing theory, *Journal of Economic Theory* 34, 13–31.

——— and Robert Korajczyk, 1987, An intertemporal equilibrium beta pricing model, Finance Working Paper No. 176, Institute of Business and Economic Research, University of California, Berkeley.

Dybvig, Philip and Jonathan Ingersoll, Jr., 1982, Mean-variance theory in complete markets, *Journal of Business* 55, 233–251.

Hakansson, Nils, 1977, The superfund: Efficient paths toward efficient capital markets in large and small countries, in Haim Levy and Marshall Sarnat, eds.: *Financial Decision Making under Uncertainty* (Academic Press, New York).

———, 1978, Welfare aspects of options and supershares, *Journal of Finance* 33, 759–776.

Huberman, Gur, 1987, Arbitrage pricing theory, in John Eatwell, Murray Milgate, and Peter Newman, eds.: *The New Palgrave: A Dictionary of Economics* (Stockton Press, New York).

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Milne, Frank, 1988, Arbitrage and diversification in a general equilibrium asset economy, *Econometrica* 56, 815–840.

Ross, Stephen, 1976a, Options and efficiency, *Quarterly Journal of Economics* 90, 75–89.

———, 1976b, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.