Analytic Tests of Factor Pricing Models

Campbell R. Harvey
Duke University, Durham, NC 27708-0120
National Bureau of Economic Research, Cambridge, MA 02138

Christopher M. Kirby
The University of Maryland, College Park, MD 20742

Abstract

This paper outlines a new procedure for constructing analytic tests of linear and nonlinear factor pricing models. First we derive a general representation of factor pricing models that is fully consistent with intertemporal utility maximization and that nests a number of well-known linear pricing results as special cases. Then we exploit this general representation to develop a straightforward approach for constructing analytic versions of the efficient generalized method of moments (GMM) estimator and the test for overidentifying restrictions for such models. No restrictions on serial correlation or conditional heteroskedasticity are required. All of our results are derived in a dynamic setting so the test assets can include dynamic trading strategies. Thus our procedure for constructing analytic GMM tests is applicable to almost any factor pricing model that is likely to be encountered in practice.

Preliminary and incomplete. Comments welcome.
1. Introduction

Asset pricing models often imply that the expected return on an asset can be written as a linear combination of factor risk premia. Indeed, a linear relation between risk and expected return is implied by some of the most well-known models in financial economics. The Sharpe (1964) – Lintner (1965) capital asset pricing model (CAPM), the Black (1972) CAPM, the Merton (1973) intertemporal CAPM, the arbitrage pricing theory (APT) of Ross (1976), and the Breeden (1979) consumption CAPM can all be classified under the general heading of linear factor models. Although these models differ in terms of their underlying structural assumptions, each is associated with a pricing relation that is a linear function of one or more market-wide risk premiums. It is this linear tradeoff between systematic risk and expected return that makes such models both intuitively appealing and analytically tractable.

As is often the case, however, intuition and analytic tractability are not without their price. The linearity of the risk-return tradeoff in the aforementioned models stems from either an implicit or explicit restriction on the joint distribution of factor realizations and asset returns. If we relax the distributional restrictions, then factor-based pricing theories generally lead to the type of nonlinear APT studied by Bansal and Viswanathan (1993). In the nonlinear version of arbitrage pricing theory, the pricing kernel is an unknown nonlinear function of the factor realizations. Thus, to test the theory, we need a way to approximate the unknown pricing kernel consistently. Although a number of different approximation techniques are available, seminonparametric (SNP) regression using a polynomial series expansion is particularly convenient. The basic idea behind this form of SNP approximation is to replace the unknown pricing kernel with a polynomial series expansion in the factors. This yields an econometric specification that is linear in the parameters of interest.

A number of different methods have been employed to estimate and test factor pricing models. Gibbons (1982) and Stambaugh (1982), for example, suggest a maximum likelihood approach based on a multivariate normal data generating process. Although the maximum likelihood approach is still in use, it has been largely supplanted by more robust methodologies. Hansen’s (1982) generalized method of
moments (GMM) is the procedure of choice in most applications. Unlike maximum likelihood estimation, the GMM approach remains consistent under a wide range of distributional assumptions, and is robust to the presence of conditional heteroskedasticity and serial correlation in the data. The one potential drawback of using GMM is that it requires a two-stage, nonlinear numerical minimization. Thus, for large systems, the computational burden of GMM estimation can be nontrivial, and problems with the minimization algorithm, such as a failure to converge, are not uncommon.

There are some cases, however, where it is possible to compute GMM estimators analytically. MacKinlay and Richardson (1991) consider one such scenario. They construct GMM tests of the Sharpe–Lintner CAPM that explicitly allow for conditional heteroskedasticity in the model disturbances. By treating the normal equations for the market model regression as a vector of moment conditions, they derive an analytic version of the GMM test statistic. Although their results serve to illustrate the effects of conditionally heteroskedasticity in tests of the CAPM, MacKinlay and Richardson (1991) deal only with single factor models where the disturbance term is serially uncorrelated. More importantly, though, they do not make any distinction between the testable implications of the single-period CAPM and intertemporal versions of the same model. An intertemporal CAPM should be able to price not only the returns from buy-and-hold investment rules but also the payoffs from dynamic trading strategies.

An alternative treatment of analytic GMM tests is provided by Zhou (1994). He considers a K-factor latent variables model and develops analytic solutions to the GMM minimization problem for a certain class of weighting matrices. Although his tests are developed in intertemporal setting, they impose a number of assumptions that limit the overall applicability of the methodology. First, the tests assume that conditional expectations are linear in a set of information variables. The general validity of this assumption is open to question. Second, the analytic solutions are for a certain class of weighting matrices. In most applications the optimal weighting matrix does not fall within this class, and the resulting analytic estimator is less efficient than the optimal GMM estimator. Finally, Zhou (1994) restricts his analysis
of analytic tests to latent variables models. As Wheatley (1989) points out, tests of latent variables models can be viewed only as tests of distributional hypotheses about mean-variance efficient portfolios of unknown composition.

In this paper we outline a simple procedure for constructing analytic GMM tests of both linear and nonlinear factor pricing models. We begin by developing a general representation of factor pricing models that is fully consistent with intertemporal utility maximization. The pricing relation implied by this general representation nests a wide range of well-known pricing results as special cases. Then we use this general representation to show that it straightforward to construct analytic expressions for the optimal GMM estimator and test statistic for any factor pricing model. No restrictions on conditionally heteroskedasticity or serial correlation are required. Moreover, because we derive the factor pricing specification in an intertemporal setting, we can easily incorporate the payoffs to dynamic trading strategies into the analysis. As a result, we are able to construct analytic tests of factor pricing models that cover almost any scenario that is likely to be encountered in practice.

2. A General Representation of Factor Pricing Models

The conventional method for deriving factor pricing models is to impose an explicit factor structure directly on the return generating process. In our view, however, it is more informative to approach this subject from a different perspective. Our basic strategy is to start with the standard stochastic discount factor representation of dynamic asset pricing models and then nest the factor pricing specification within this more general pricing framework. Much of the analysis in this section of the paper builds on a simple insight of Chen and Ingersoll (1983) regarding exact pricing in linear factor models. Similar theoretical treatments can be found in recent work by Bansal and Viswanathan (1993), Carhart and Krail (1993), and Cochrane (1994).

2.1. The Economic Environment

We begin by considering a simple endowment economy in which an unspecified number of rational investors trade assets in securities markets. Both asset prices and
their payoffs are denominated in terms of a single consumption good and there are no transactions costs, short sale constraints, or other barriers to trade. All of the market participants in this economy solve dynamic portfolio problems in order to arrive at an optimal holding of securities. It is well known that under these circumstances the first order conditions for each agent’s intertemporal optimization problem take the form:

\[ E[\tilde{m}_{t+1} \tilde{R}_{t+1}^i \mid \mathcal{F}_t] = 1 \quad \forall i, \]  

(1)

where \( \tilde{R}_{t+1}^i \) is the total return on asset \( i \) at date \( t + 1 \), \( \tilde{m}_{t+1} \) is the agent’s intertemporal marginal rate of substitution between dates \( t \) and \( t + 1 \), and \( \mathcal{F}_t \) is the agent’s information set at date \( t \).

Equation (1) is consistent with a broad class of intertemporal asset pricing models. The basic implication of (1) is that if the total return on asset \( i \) is discounted by any agent’s intertemporal marginal rate of substitution it forms a stochastic process that has a constant conditional expectation. If markets are complete then marginal rates of substitution are equated across all investors and \( \tilde{m}_{t+1} \) is uniquely determined. Otherwise, marginal rates of substitution differ across individuals and the set of random variables that satisfies the pricing restriction in (1) contains multiple elements. Every element of this set is, however, a valid pricing kernel in the sense that it assigns the correct price to all traded assets.

2.2. Linear Factor Models

The pricing relation shown in (1) places few restrictions on the functional form of the agent’s intertemporal marginal rate of substitution. Note, however, that by applying the law of iterated expectations this equation can be written as:

\[ E \left[ E[\tilde{m}_{t+1} \mid \mathcal{F}_{t+1}] \tilde{R}_{t+1}^i \mid \mathcal{F}_t \right] = 1 \quad \forall i. \]  

(2)

When expressed in this form equation (1) suggests a straightforward way to model the behavior of \( \tilde{m}_{t+1} \) that is consistent with the standard linear factor specification. First let \( \tilde{f}_{t+1} \) denote the date \( t + 1 \) realization of an \( K \times 1 \) vector of common factors
that represent sources of undiversifiable risk in our endowment economy. Then partition the agent’s information set $\mathcal{F}_{t+1}$ as $[\mathcal{I}_{t+1}, \mathcal{F}_t]$ so that $\mathcal{I}_{t+1}$ represents the new information that becomes available at date $t + 1$. If we assume that conditional on $\mathcal{F}_t$ the joint distribution of $\tilde{m}_{t+1}$ and $\mathcal{I}_{t+1}$ is multivariate elliptical then it follows that $E[\tilde{m}_{t+1} \mid \mathcal{I}_{t+1}, \mathcal{F}_t]$ can be written a linear combination of the elements of $\mathcal{I}_{t+1}$.\textsuperscript{1} Thus by imposing a sufficient statistic restriction of the form:

$$E[\tilde{m}_{t+1} \mid \mathcal{I}_{t+1}, \mathcal{F}_t] = a_t + b'_t \tilde{f}_{t+1}$$

(3)

where $a_t$ and the elements of $b_t$ depend on the information set $\mathcal{F}_t$ we obtain a pricing kernel that is conditionally linear in the factor realizations.\textsuperscript{2}

Note that equation (3) does not constitute a restriction on risk preferences in the conventional sense. It permits a great deal of heterogeneity across investors while at the same time eliminating the need to explicitly model utility. We only require that a linear function of the factors be sufficient with respect to $\mathcal{I}_{t+1}$ for predicting at least one investor’s intertemporal marginal rate of substitution. If the agent in question has preferences that are both time and state separable then equation (3) reduces to an assumption that $\tilde{m}_{t+1}$ is linearly spanned by the factors.\textsuperscript{3} If nonseparabilities are present, however, the value of $\tilde{m}_{t+1}$ will typically be unknown to the agent at date $t + 1$. By imposing the sufficient statistic restriction we can cover this more general scenario, and it has a certain degree of intuitive appeal as well.

Combining the sufficient statistic restriction in (3) with the general pricing result in (2) yields:

$$E[(a_t + b'_t \tilde{f}_{t+1}) \tilde{F}_{t+1} \mid \mathcal{F}_t] = 1 \quad \forall i.$$  

(4)

\textsuperscript{1}It is clear that conditional multivariate normality would be sufficient for this result to obtain. But the multivariate normal distribution is just one member of the class of distributions that is characterized by linear conditional expectations. Cambanis, Huang, and Simons (1981) show that any multivariate elliptical distribution has this property.

\textsuperscript{2}Bansal and Viswanathan (1993) and Bansal, Hsieh and Viswanathan (1993) exploit a similar restriction to derive a nonlinear version of the APT, and Carhart and Kralj (1993) and Cochrane (1994) develop analogous models for $\tilde{m}_{t+1}$ in their analysis of the conditional CAPM.

\textsuperscript{3}With the additional assumption that the factors are traded assets we obtain the intertemporal analogue of the single-period result given by Chen and Ingersoll (1983). In this case, imposing the sufficient statistic restriction is equivalent to assuming that some agent’s optimal portfolio is a time-varying linear combination of the factors.
At first the pricing relation shown in equation (4) may not appear to be of the general form typically associated with linear factor models. Nevertheless, the pricing implications are identical. To see why this is true note that equation (4) can also be expressed as:

\[ E[\tilde{R}_{t+1} | F_t] = \gamma_0 + \lambda^i \text{cov}_t(\tilde{R}_{t+1}, \tilde{f}_{t+1}) \quad \forall i, \tag{5} \]

where \( \gamma_0 \equiv (a_t + b^t E[\tilde{f}_{t+1} | F_t])^{-1} \) and \( \lambda_t \equiv -\gamma_0 b_t \). Equation (5) implies that the conditional expected return on each asset is a linear function of the conditional covariance of its return with the vector of factor realizations. This is, of course, the same pricing result that emerges from the conventional approach to linear factor pricing. In the factor pricing literature, however, equation (5) is usually expressed as:

\[ E[\tilde{R}_{t+1} | F_t] = \gamma_0 + \beta^t \gamma_t \quad \forall i, \tag{6} \]

where \( \gamma_t \) denotes a \( K \times 1 \) vector of conditional factor risk premia and \( \beta_t \) denotes a \( K \times 1 \) vector of conditional factor loadings.\(^4\)

Now suppose we impose the additional assumption that \( a_t \) and \( b_t \) are time-invariant. Applying the law of iterated expectations to equation (4) yields:

\[ E[(a + b^t \tilde{f}_{t+1}) \tilde{R}_{t+1}] = 1 \quad \forall i. \tag{7} \]

Thus if we proceed along the same lines as before we obtain a linear pricing relation of the form:

\[ E[\tilde{R}_{t+1}] = \gamma_0 + \lambda \text{cov}(\tilde{R}_{t+1}, \tilde{f}_{t+1}) \quad \forall i, \tag{8} \]

where \( \gamma_0 \equiv (a + b^t E[\tilde{f}_{t+1}])^{-1} \) and \( \lambda \equiv -\gamma_0 b_t \). Equation (8) is the unconditional version of linear factor pricing. It implies that the unconditional expected return on each asset is a linear function of the unconditional covariance of its return with the vector of factor realizations.

The pricing results shown in (5) and (8) are consistent with a number of well-known asset pricing models. Risk-return tradeoffs of this form are implied by the

\(^4\)The conditional factor loadings for asset \( i \) are simply the slope coefficients in the conditional least-squares projection of \( \tilde{R}_{t+1} \) onto the space spanned by \( 1 \) and \( \tilde{f}_{t+1} \). Thus if we let \( V_{ft} \) denote the conditional variance-covariance matrix of the factors then \( \lambda^i \text{cov}_t(\tilde{R}_{t+1}, \tilde{f}_{t+1}) \) can be written as \( \beta^t \gamma_t \) where \( \gamma_t \equiv V_{ft} \lambda_t \).
Sharpe (1964) - Lintner (1965) capital asset pricing model (CAPM), the Black (1972) CAPM, the Merton (1973) intertemporal CAPM, the arbitrage pricing theory (APT) of Ross (1976), and the Breeden (1979) consumption CAPM. It is important to remember, however, that the linearity of the relation between systematic risk and expected return can be traced directly to our explicit restriction on the joint distribution of the factor realizations and asset returns. In the absence of such distributional restrictions the pricing relation will typically be nonlinear. As a consequence, a general treatment of factor pricing must cover nonlinear models as well.

2.3. Nonlinear Factor Models

The key assumption underlying the foregoing analysis is that the conditional expectation of \( \tilde{m}_{t+1} \) given \( \mathcal{F}_{t+1} \) can be written a linear function of a \( K \times 1 \) vector of factor realizations. It is well known, of course, that in general conditional expectations are not linear. If we relax the assumption of linear conditional expectations then we obtain a sufficient statistic restriction of the form

\[
E[\tilde{m}_{t+1} | \mathcal{F}_{t+1}] = \phi_t(\tilde{f}_{t+1})
\]

where \( \phi_t(\cdot) \) denotes some well-behaved nonlinear function. Equation (9) is similar to the dimensionality restriction that is the basis for the nonlinear arbitrage pricing theory studied by Bansal and Viswanathan (1993). The basic premise underlying the nonlinear APT is that a small number of factors is sufficient for predicting at least one agent’s intertemporal marginal rate of substitution.

Substituting the sufficient statistic restriction in (9) into the left-hand side of equation (2) yields the pricing relation:

\[
E[\phi_t(\tilde{f}_{t+1}) \tilde{R}_{t+1}^i | \mathcal{F}_t] = 1 \quad \forall i
\]

Although the nonlinear pricing kernel \( \phi_t(\cdot) \) is unknown, it can be approximated using a seminonparametric (SNP) approach. The basic idea behind the SNP approach is to apply classical methods of estimation and inference to models derived from truncated series expansions. A number of recent studies including those by Gallant
and Tauchen (1989) and Bansal and Viswanathan (1993) employ an SNP approach to estimate and test asset pricing models. There are several different variants of the SNP methodology. Given the context of our estimation problem, however, it makes sense to adopt a regression-based approach. The regression-based version of SNP analysis uses an expanding sequence of finite-dimensional parametric models to approximate an unknown regression function.

In order to implement the SNP procedure it is necessary to choose a class of approximating functions that form the basis for the series expansion. We focus on the polynomial class of functions because it leads to simple pricing relations that can easily be analyzed within the standard linear framework. Suppose, for example, that we model the unknown pricing kernel as a second-order polynomial series expansion in the factors:

\[
\phi(\tilde{f}) = a + \sum_{j=1}^{K} b^j \tilde{f}^j + \sum_{j=1}^{K} \sum_{k=j}^{K} b^{jk} \tilde{f}^j \tilde{f}^k.
\]  

Given this specification for \(\phi(\cdot)\) equation (10) becomes:

\[
E \left[ \left( a_t + \sum_{j=1}^{K} b^j_{t} \tilde{f}_{t+1}^j + \sum_{j=1}^{K} \sum_{k=j}^{K} b^{jk} \tilde{f}_{t+1}^j \tilde{f}_{t+1}^k \right) \tilde{R}_{t+1}^i \left| \mathcal{F}_t \right. \right] = 1 \quad \forall i.
\]  

Thus if we treat each of the monomials in the series expansion as a separate factor then equation (12) can be manipulated to obtain a linear pricing relation of the same form as that discussed earlier.

Using a polynomial series expansion to approximate an unknown function is a long-established tradition in statistics. One of the main advantages of this strategy is that the leading term of the expansion is linear in the factors. This leading term captures all of the linear effects while the higher-order terms accommodate any deviations from linearity. The truncation point of the series expansion is typically based on either a deterministic or adaptive rule. If an adaptive rule is used then the truncation point is determined by moving along an upward expansion path until some criterion, often a large-sample chi-square test, indicates that the specification is adequate. The preferred specification is then subjected to a battery of standard diagnostic procedures.
Gallant (1987) and Gallant and Nychka (1987) show how the SNP approach leads to consistent approximation of unknown functions, and recent work by Andrews (1991) establishes the consistency and asymptotic normality of SNP series estimators in the i.i.d. regression setting. Whether or not their results carry over to general time-series environments is an open question. As Gallant and Tauchen (1989) point out, however, nothing is lost if we choose to view an SNP specification as a conventional parametric model that has been subjected to sensitivity analysis. Then we are free to apply classical methods of estimation and inference in the regular manner. With respect to the current analysis this implies that our analytic estimator of the vector of unknown parameters in the SNP specification is consistent and asymptotically normal under standard regularity conditions.

The upshot of the foregoing analysis is that we can express the pricing relation for both linear and nonlinear factor pricing models as:

\[
E[\tilde{R}_{t+1}^i | \mathcal{F}_t] = \gamma_{0t} + \lambda_t \text{cov}_t(\tilde{R}_{t+1}^i, \tilde{s}_{t+1}) \quad \forall i,
\]

(13)

where the first \(K\) elements of \(\tilde{s}_{t+1}\) are the factors and the remaining elements, if any, represent higher-order terms associated with a polynomial series expansion. The unconditional version of the above pricing relation can be tested without any difficulty. Testing the conditional specification in (13) is a bit more complicated. Note that (13) says nothing about how \(\gamma_{0t}, \lambda_t\) and the conditional expectations operator \(E[\cdot | \mathcal{F}_t]\) depend on the elements of the information set \(\mathcal{F}_t\). It turns out, however, that there is a simple way to deal with this problem that meshes nicely with the SNP approach discussed above. Suppose we let \(\tilde{Z}_t\) denote the date \(t\) realization of an \(L \times 1\) vector of instrumental variables that are in \(\mathcal{F}_t\).\(^5\) If we model \(\gamma_{0t}, \lambda_t\) and conditional expectations \(E[\cdot | \mathcal{F}_t]\) as polynomial series expansions in \(\tilde{Z}_t\) then the conditional version of (13) becomes testable.

Modelling the aforementioned quantities as series expansions in \(\tilde{Z}_t\) can be motivated by recent work on SNP estimation of conditionally constrained heterogeneous processes. Gallant and Tauchen (1989), for example, use an SNP approach to approx-

\(^5\)We assume without loss of generality that the first element of the vector \(\tilde{Z}_t\) is the trial random variable 1.
imate the error density associated with a vector autoregression (VAR). They model the error density as a polynomial series expansion in the standardized residual of the VAR multiplied by a standard Gaussian density. This yields a nonnormal density that can take on a wide variety of different shapes. Then to allow for conditional heterogeneity in the process they set each of the coefficients in the series expansion involving the standardized residuals equal to a polynomial expansion in the regressors. The resulting density estimator provides a consistent approximation of the true unknown error density under fairly general conditions.

Our strategy for conducting SNP tests of the pricing relation shown in (13) can be viewed a natural variant of the Gallant and Tauchen (1989) approach. It is typically much easier to implement, however, because we do not attempt to approximate the entire law of motion for the process under consideration. Thus unlike Gallant and Tauchen (1989) we can usually get by with a relatively sparse parameterization. In many cases, for example, it may be reasonable to model $E[(\cdot) | \mathcal{F}_t]$, $\gamma_{0t}$ and the elements of $\lambda_t$ as linear functions of the set of instrumental variables. The result is a parsimonious specification that can be tested in a straightforward manner.

2.4. Pricing the Payoffs to Dynamic Trading Strategies

Although the focus of the analysis thus far has been on pricing the returns from buy-and-hold investment strategies, the general approach to factor pricing developed above has implications concerning the payoffs to dynamic trading strategies as well. Suppose, for instance, that we multiply both sides of the conditional linear pricing relation shown in (5) by $\tilde{Z}_t^\ell$ where $\ell \in \{1, 2, \ldots, L\}$. Because $\tilde{Z}_t^\ell$ is in the date $t$ information set the resulting restriction can be written as:

$$E[\tilde{R}_{t+1}^i \tilde{Z}_t^\ell | \mathcal{F}_t] = \gamma_{0t} \tilde{Z}_t^\ell + \lambda_t \text{cov}_t(\tilde{R}_{t+1}^i \tilde{Z}_t^\ell, f_{t+1}) \quad \forall i.$$  

(14)

The interpretation of (14) is straightforward. Think in terms of our endowment economy. The quantity $\tilde{R}_{t+1}^i \tilde{Z}_t^\ell$ can be viewed as the payoff from a dynamic trading strategy. This trading strategy is quite simple: invest $\tilde{Z}_t^\ell$ units of the consumption good in asset $i$ at date $t$ and receive a payoff of $\tilde{R}_{t+1}^i \tilde{Z}_t^\ell$ units of the consumption good.
at date $t + 1$. Equation (14) shows how to price the payoff from such an investment strategy. It says, not surprisingly, that the price of the payoff is a linear function of its conditional covariance with factor realizations. Thus payoffs to dynamic trading strategies are priced in the same manner as every other payoff in the economy.

Extending the analysis of dynamic trading strategies to models with nonlinear factor structures is straightforward. We just impose the nonlinear dimensionality restriction of equation (9), replace the nonlinear pricing kernel with a polynomial series expansion of the form shown in (11), and proceed in exactly the same manner as before. Of course in practice we usually want to know whether a model can simultaneously price the returns to both passive and dynamic strategies. Recall that $\tilde{Z}_t$ denotes the date $t$ realization of the $L \times 1$ vector $[1, \tilde{Z}^1, \tilde{Z}^2, \ldots, \tilde{Z}^{L-1}]'$. Thus a more general form of the pricing restriction shown in (14) is:

$$
E[\tilde{R}_{t+1}^i \tilde{Z}_t | \mathcal{F}_t] = \gamma_0 \tilde{Z}_t + \lambda_t \text{cov}_t(\tilde{R}_{t+1}^i \tilde{Z}_t, \tilde{s}_{t+1}) \quad \forall i.
$$

(15)

where $\tilde{s}_{t+1}$ is defined in the manner indicated earlier.

The pricing result shown in (15) is the basis for the class of asset pricing tests considered in this paper. Our approach to testing these pricing restrictions represents a significant departure from what is commonly seen in the empirical literature. First of all, we use GMM to avoid making strong assumptions concerning the distribution of the asset returns and instrumental variables. Thus unlike the maximum likelihood approach which assumes normality our methodology is robust to the presence of conditional heteroskedasticity and serial correlation in the data. More importantly, though, we show how to bypass the numerical minimization procedure that is typically required for GMM estimation of factor pricing models. As a consequence we are able to provide analytic results for a wide variety of different scenarios.

3. Analytic Methods of Estimation and Testing

In this section we outline a general strategy for conducting analytic tests of dynamic factor pricing models. Although our approach is similar in spirit to that of MacKinlay and Richardson (1991) and Zhou (1994), our analysis differs from theirs in a number
of important respects. Like us, MacKinlay and Richardson (1991) study analytic GMM tests of linear factor models. But they focus exclusively on tests of the Sharpe–Lintner CAPM under conditions where the disturbance term for the model is serially uncorrelated. Zhou (1994), on the other hand, considers a K-factor latent variables model and develops analytic solutions to the GMM minimization problem for a certain class of weighting matrices. But for the efficient weighting matrix to fall within this class, the model disturbances must be independently and identically distributed.

Unlike these previous studies, we develop simple procedures for conducting analytic GMM tests that are applicable to a broad range of dynamic factor pricing models. No restrictions on conditional heteroskedasticity or serial correlation are required. Our analytic estimators are efficient, and our analytic test criteria can easily be constructed using standard regression packages. The arguments used to develop the analytic test criteria are based on asymptotic results associated with Hansen’s (1982) generalized method of moments (GMM). We begin, therefore, with a brief discussion of the relevant distributional theory. This will serve to establish the notation that will be used throughout the remainder of the paper.

3.1. The GMM Framework

Suppose a model implies a set of moment conditions of the form:

$$E[h(\tilde{x}_t, \theta_0)] = 0,$$  \hspace{1cm} (16)

where $\tilde{x}_t$ is an $l \times 1$ vector of observable random variables, $\theta_0$ is an $m \times 1$ vector of unknown parameters and $h(\cdot)$ denotes an $n \times 1$ vector of equations. By imposing suitable restrictions, such as stationarity and ergodicity, on the data generating process we can insure that a version of central limit theorem holds. Thus if the model is valid the sample counterpart of (16):

$$g_T(\theta_0) \equiv \frac{1}{T} \sum_{t=1}^{T} h(\tilde{x}_t, \theta_0),$$  \hspace{1cm} (17)
should be close to zero when evaluated for a large number of observations. The basic strategy behind the GMM procedure is to estimate the unknown vector $\theta_0$ by choosing the value of $\theta$ that minimizes some measure of the overall deviation of the sample moments from zero.

If the system of moment conditions is exactly identified ($m = n$), the GMM estimator is the value of $\theta$ that sets the sample moments equal to zero. For the more commonly encountered scenario where the system is overidentified ($m < n$), finding a vector of parameters that sets all of the sample moments equal to zero is not feasible. It is possible, however, to find a value of $\theta$ that sets $m$ linear combinations of the $n$ sample moment conditions equal to zero; simply let $A_T$ be an $m \times n$ matrix such that $A_Tg_T(\theta) = 0$ has a well-defined solution. The value of $\theta$ that solves this system of equations is the GMM estimator.

Although we have considerable latitude in choosing the matrix $A_T$, Hansen (1982) shows that the variance-covariance matrix of the method of moments estimator is minimized by setting $A_T$ equal to $\hat{D}'\hat{S}^{-1}$ where $\hat{D}$ and $\hat{S}$ are consistent estimates of:

$$D = E\left[\left.\frac{\partial h(\bar{x}_t, \theta)}{\partial \theta'}\right|_{\theta_0}\right] \quad \text{and} \quad S = \sum_{j=-\infty}^{\infty} E[h(\bar{x}_t, \theta_0)h(\bar{x}_{t-j}, \theta_0)'].$$  \hspace{1cm} (18)

For this optimal choice of $A_T$ it follows that:

$$\sqrt{T}(\hat{\theta} - \theta_0) \overset{d}{\rightarrow} N(0, (D'S^{-1}D)^{-1}),$$  \hspace{1cm} (19)

where $\hat{\theta}$ denotes the efficient estimator of $\theta_0.$ In cases where the model is overidentified we can construct a test based on the overidentifying restrictions. Under the null hypothesis that the restrictions imposed by the model are true the quadratic form:

$$J_T \equiv T g_T(\hat{\theta})'S^{-1}g_T(\hat{\theta})$$  \hspace{1cm} (20)

It is assumed in the discussion that follows that $\bar{x}_t$ is generated by a stationary and ergodic process, that the second moment matrix of $h(\cdot)$ exists and is finite, and that the regularity conditions given by Hansen (1982) are satisfied. These assumptions are sufficient for the asymptotic distribution theory of GMM to hold.

\footnote{The limiting distribution for suboptimal choices of $A_T$ is given by:

$$\sqrt{T}(\hat{\theta} - \theta_0) \overset{d}{\rightarrow} N(0, (AD)^{-1}ASA'(AD)^{-1}',)$$  \hspace{1cm} (20A)

where $A_T \overset{p}{\rightarrow} A.$ This result provides a way to determine the relative efficiency of different estimators.}
converges to a central chi-square random variable with degrees of freedom equal to the
number of overidentifying restrictions. Although in general the $S$ matrix is unknown,
it can be replaced by a consistent estimate without affecting the limiting distribution
of the test statistic.

In principle we should be able estimate the unknown vector $\theta_0$ analytically. All
that we need are analytic expressions for the $D$ and $S$ matrices. The only problem
is that in most cases it is difficult to derive such expressions, and even when we can
derive them there is no guarantee that we can solve the resulting system of equations
without resorting to numerical methods. But estimation of factor pricing models
represents an exception to the general rule. Because these models are associated
with simple linear moment restrictions, it is easy to develop the required analytic
expressions and solve for the efficient GMM estimator. The remainder of this section
outlines the key features of the analytic approach to estimation and testing. First we
work through a straightforward example that illustrates the basic concepts. Then we
show how to extend the methodology to estimate and test more complicated models.

3.2. Analytic Tests of the Black CAPM

Let $\tilde{R}_{t+1}^i$ and $\tilde{R}_{t+1}^m$ denote the total return on asset $i$ and the total return on the
value-weighted market portfolio. The conventional specification of the Black (1972)
CAPM takes the form:

$$ E[\tilde{R}_{t+1}^i] = \gamma_0(1 - \beta_i) + \beta_i E[\tilde{R}_{t+1}^m] \quad \forall i, $$

(21)

where:

$$ \beta_i \equiv \frac{cov(\tilde{R}_{t+1}^i, \tilde{R}_{t+1}^m)}{var(\tilde{R}_{t+1}^m)} $$

denotes the beta of asset $i$ with respect to the return on the market portfolio and
$\gamma_0$ is the return on the zero-beta portfolio. In keeping with our earlier discussion,
however, we can express this pricing restriction as:

$$ E[\tilde{R}_{t+1}] = \gamma_0 1_N + \lambda cov(\tilde{R}_{t+1}, \tilde{R}_{t+1}^m), $$

(22)

where $\lambda \equiv E[\tilde{R}_{t+1}^m - \gamma_0]/var(\tilde{R}_{t+1}^m)$ denotes the price of covariance risk.
The Black CAPM, like all linear factor models, lends itself easily to generalized method of moments estimation. Assume, for example, that the data set contains $N$ assets. Let $\varphi$ and $\tilde{x}_{t+1}$ denote the $2 \times 1$ and $(N + 1) \times 1$ vectors $[\gamma_0, \lambda]'$ and $[\tilde{R}_{t+1}^m, \tilde{R}_{t+1}']$. We can estimate $\varphi$ using the sample moment conditions:

$$g_T(\hat{\mu}_m, \hat{\mu}_r, \tilde{x}_{t+1}, \varphi) = \frac{1}{T} \sum_{t=1}^{T} (\tilde{R}_{t+1} - \gamma_0 1_N - \lambda (\tilde{R}_{t+1}^m - \hat{\mu}_m)(\tilde{R}_{t+1} - \hat{\mu}_r)) \quad (23)$$

where $\hat{\mu}_m$ and $\hat{\mu}_r$ denote the sample analogues of $E[\tilde{R}_{t+1}^m]$ and $E[\tilde{R}_{t+1}]$. The above specification for estimating the Black CAPM differs from the traditional GMM specification in two respects. First, we replace $\mu_m$ and $\mu_r$ with their sample analogues prior to estimating $\gamma_0$ and $\lambda$ rather than estimating all of the parameters simultaneously. This form of sequential estimation simplifies the analysis and has no effect on the asymptotic distribution of the GMM estimator of $\varphi$. Second, we estimate the risk premium per unit of market risk instead of the $N \times 1$ vector of asset betas. This results in system of equations that is linear in the unknown parameters.

The proof that sequential estimation has no effect on the efficiency of the GMM estimator of $\varphi$ is quite simple. Suppose that instead of (23) we consider the full set of sample moment conditions:

$$g_T(\tilde{x}_{t+1}, \phi) = \frac{1}{T} \sum_{t=1}^{T} \left( \begin{array}{c} \tilde{R}_{t+1} - \mu_r \\ \tilde{R}_{t+1}^m - \mu_m \\ \tilde{R}_{t+1} - \gamma_0 1_N - \lambda (\tilde{R}_{t+1}^m - \mu_m)(\tilde{R}_{t+1} - \mu_r) \end{array} \right) \quad (24)$$

where $\phi \equiv [\mu_r', \mu_m, \gamma_0, \lambda]'$. First we partition the above vector as:

$$g_T(\tilde{x}_{t+1}, \phi) = \frac{1}{T} \sum_{t=1}^{T} \left( \begin{array}{c} h_1(\tilde{x}_{t+1}, \phi_1) \\ h_2(\tilde{x}_{t+1}, [\phi_1', \phi_2']) \end{array} \right) \quad (25)$$

where $\phi_1 \equiv [\mu_r', \mu_m]'$ and $\phi_2 \equiv [\gamma_0, \lambda]'$. Then we partition $D$ and $S$ accordingly. Because the expected value of the Jacobian of $h_2(\cdot)$ with respect to $\phi_1$ is zero, the partitioned version of $D$ is a diagonal matrix. Thus it follows that the GMM estimator of $\phi_2$ in (24) and the GMM estimator of $\varphi$ in (23) are equally efficient. This is easily verified by noting that: (i) when $A_{\phi T}$, the weighting matrix for the specification shown in (24), is set equal to the block diagonal matrix $diag\{\hat{D}_{11}^{'}, \hat{S}_{11}^{-1}, \hat{D}_{22}^{'}, \hat{S}_{22}^{-1}\}$
the resulting estimator for $\phi_2$ and the sequential estimator for $\phi$ coincide and; (ii) the variance-covariance matrix of the estimator for $\phi_2$ is given by $(D'_{22}S_{22}^{-1}D_{22})^{-1}$ for both the optimal choice of the weighting matrix and for the suboptimal choice $A_{\phi_T} = \text{diag}\{\hat{D}'_{11}\hat{S}_{11}^{-1}, \hat{D}'_{22}\hat{S}_{22}^{-1}\}$.

The analytic estimator of $\phi$ is the solution to the set of linear equations:

$$\hat{D}'\hat{S}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\hat{R}_{t+1} - \gamma_0I_N - \lambda(\hat{R}_{t+1}^m - \hat{\mu}_m)(\hat{R}_{t+1} - \hat{\mu}_r)\right) = 0,$$  \hspace{1cm} (26)

where $\hat{D}$ and $\hat{S}$ denote consistent estimators of the $D$ and $S$ matrices. It is readily apparent from (23) that $D$ is given by $-\{1_N, \sigma_{rm}\}$ where $\sigma_{rm}$ is the $N \times 1$ vector of covariances between asset returns and the return on the market portfolio. Using this result we can write equation (26) as

$$\hat{D}'\hat{S}^{-1}\hat{\mu}_r + \hat{D}'\hat{S}^{-1}\hat{D}\phi = 0.$$  \hspace{1cm} (27)

The solution to (27) is:

$$\hat{\phi} = -(\hat{D}'\hat{S}^{-1}\hat{D})^{-1}\hat{D}'\hat{S}^{-1}\hat{\mu}_r.$$  \hspace{1cm} (28)

Thus given a consistent estimate of $S$ we can compute the optimal GMM estimator for $\phi$ quite easily.\(^8\)

### 3.3. The Variance-Covariance Matrix

In constructing our estimator of $S$ we want to allow for the possibility that the disturbances for the model are serially correlated and conditionally heteroskedastic. As a consequence we have to employ some type of weighting scheme to insure that the estimator is always positive definite. Although a number of different weighting schemes have been proposed in the econometric literature, Andrews (1991) shows that the weights based on the quadratic spectral kernel are optimal in the sense that they minimize the asymptotic truncated mean squared error of the estimator. We follow the suggestion of Andrews (1991) and use quadratic spectral weights.

\(^8\)Cochrane (1994) discusses a similar approach for obtaining analytic estimators in the CAPM setting. His analysis, however, is framed in terms of noncentral second-moments.
The bandwidth needed to calculate the weights can be obtained using one of two methods. The simplest approach is to rely on the results of Anderson (1971) and set the bandwidth equal to \(0.5T^{1/5}\). This plug-in method of bandwidth selection is attractive because it yields an estimator of \(S\) that depends only on the size of the sample in question, which is in keeping with the spirit of the foregoing analytic analysis. The other alternative is to use the data to estimate the optimal bandwidth. This data-dependent approach can be viewed as a semi-analytic method. It is more complicated than the plug-in approach, but the Monte Carlo results of Andrews (1991) suggest that it may produce superior performance in small samples.

The procedure for constructing an estimate of \(S\) is as follows: first we compute a consistent estimate of \(\varphi\) using the formula:

\[
\hat{\varphi}^* = -(\hat{D}'\hat{D})^{-1}\hat{D}'\hat{\mu}_r.
\]  

Equation (29) is the analytic solution to the first-stage minimization problem in the two-stage Hansen and Singleton (1982) procedure. The first-stage estimate of \(\varphi\) is used to form the vector of sample disturbances:

\[
\hat{h}_{t+1} = (R_{t+1} - \hat{\gamma}_0^*1_N - \hat{\lambda}^*(R_{t+1}^m - \hat{\mu}_m)(R_{t+1} - \hat{\mu}_r)).
\]  

Then we specify the bandwidth either by the plug-in method or, as suggested by Andrews (1991), by fitting a first-order autoregressive model to each element of \(\hat{h}_{t+1}\) and using the resulting autocorrelation coefficients and residual variances to estimate the optimal bandwidth. The estimator for \(S\) is given by the formula:

\[
\hat{S} = \frac{1}{T} \sum_{t=1}^{T} \hat{h}_{t+1} \hat{h}_{t+1}' + \sum_{j=1}^{T-1} \left( \frac{1}{T} \sum_{t=j+1}^{T} \hat{w}(j)(\hat{h}_{t+1} \hat{h}_{t-j+1}' + \hat{h}_{t-j+1} \hat{h}_{t+1}') \right),
\]  

where \(\hat{w}(j)\) is the weight at lag \(j\) given by the quadratic spectral kernel using the bandwidth determined earlier.

Once we have a consistent estimate of \(S\) in hand we are in a position to construct an analytic test of the model. Recall that in general the overidentifying test statistic is given by:

\[
J_T = T g_T(\hat{\mu}_m, \hat{\mu}_r, \hat{\varphi}; S^{-1} g_T(\hat{\mu}_m, \hat{\mu}_r, \hat{\varphi}).
\]  

17
Substituting for $g_T(\hat{\mu}_m, \hat{\mu}_r, \hat{x}_{t+1}, \varphi)$ from (23) and replacing $S$ with our consistent estimate yields:

$$J_T \equiv T(\hat{\mu}_r - \hat{\gamma}_0 1_N - \hat{\lambda} \hat{\sigma}_{rm})' \hat{S}^{-1}(\hat{\mu}_r - \hat{\gamma}_0 1_N - \hat{\lambda} \hat{\sigma}_{rm})$$  \hspace{1cm} (33)

with $\hat{\gamma}_0$ and $\hat{\lambda}$ given by (28). Under the null hypothesis that the Black CAPM prices all of the assets correctly the criterion shown in (33) converges to a central chi-square random variable with $N - 2$ degrees of freedom.

3.4. Analytic Tests in General

Now that the basics of the analytic approach to estimation and testing are established we can extend the analysis to cover more complicated situations. First we illustrate how to handle conditional pricing models. We will assume, as is the conventional practice, that the exists an asset whose return $\tilde{R}_{t+1}^f$ is known conditional on $\mathcal{F}_t$.\footnote{Note that this assumption does not limit the generality of the subsequent results. The analysis for the case where there is no conditionally risk free asset is virtually identical.}

Because conditional model must price this asset, it must be the case that $\gamma_{0t} = \tilde{R}_{t+1}^f$. Let's take the conditional version of the CAPM as our example. The pricing relation for this model is:

$$E[\tilde{r}_{t+1} | \mathcal{F}_t] = \lambda_t \text{cov}_t(\tilde{r}_{t+1}, \tilde{r}_{t+1}^m),$$  \hspace{1cm} (34)

where $\tilde{r}_{t+1} \equiv (\tilde{R}_{t+1} - \tilde{R}_{t+1}^f 1_N)$ is an $N \times 1$ vector of excess asset returns, $\tilde{r}_{t+1}^m \equiv (\tilde{R}_{t+1}^m - \tilde{R}_{t+1}^f)$ is the excess return on the market, and $\lambda_t \equiv E[\tilde{r}_{t+1}^m | \mathcal{F}_t]/\text{var}_t(\tilde{r}_{t+1}^m)$ denotes the conditional price of covariance risk.

To keep the notation manageable we will assume that $\lambda_t$, $E[\tilde{r}_{t+1}^m | \mathcal{F}_t]$, and the elements of $E[\tilde{r}_{t+1} | \mathcal{F}_t]$ are modelled as linear functions of the set of instrumental variables. This has no effect on the generality of the results. The analysis using polynomial series expansions proceeds along exactly the same lines. In either case we obtain a pricing relation of the form:

$$E \left[ \tilde{r}_{t+1} - (\tilde{r}_{t+1} - \Delta_r' \tilde{Z}_t)(\tilde{r}_{t+1}^m - \delta_m' \tilde{Z}_t) \tilde{Z}_t' \psi_m | \mathcal{F}_t \right] = 0,$$  \hspace{1cm} (35)

where the $L \times N$ matrix $\Delta_r$, the $L \times 1$ vector $\delta_m$, and the $L \times 1$ vector $\psi_m$ represent our unknown parameters. Applying the law of iterated expectations to (35) yields
the unconditional restriction:

$$E[\tilde{r}_{t+1} - (\tilde{\epsilon}_{t+1} \tilde{u}_{t+1}) \tilde{Z}_t' \psi_m] = 0,$$

(36)

where $\tilde{\epsilon}_{t+1} \equiv (\tilde{r}_{t+1} - \Delta_r \tilde{Z}_t)$ and $\tilde{u}_{t+1} \equiv (\tilde{r}_{t+1} - \delta_m \tilde{Z}_t)$.

Let $\tilde{x}_{t+1}$ denote the $(N + L + 1) \times 1$ vector $[\tilde{r}_{t+1}^{\prime}, \tilde{r}_{t+1}', \tilde{Z}_t']$. We can estimate the parameters of the above system using the set of sample moment conditions:

$$g_T(\tilde{u}_{t+1}, \tilde{\epsilon}_{t+1}, \tilde{x}_{t+1}, \psi_m) = \frac{1}{T} \sum_{t=1}^{T} (\tilde{r}_{t+1} - (\tilde{\epsilon}_{t+1} \tilde{u}_{t+1}) \tilde{Z}_t' \psi_m),$$

(37)

where $\tilde{\epsilon}_{t+1}$ and $\tilde{u}_{t+1}$ denote the fitted values of $\tilde{\epsilon}_{t+1}$ and $\tilde{u}_{t+1}$ obtained by using ordinary least squares (OLS) to estimate $\Delta_r$ and $\delta_m$. If our model is valid this form of sequential estimation has no effect on the asymptotic distribution of the GMM estimator of $\psi_m$. To see why note that by assumption both $\tilde{\epsilon}_{t+1}$ and $\tilde{u}_{t+1}$ are conditionally independent of $\tilde{Z}_t$. Thus we can easily verify that the $D$ matrix for the full set of sample moment conditions:

$$g_T(\tilde{x}_{t+1}, \Delta_r, \delta_m, \psi_m) = \frac{1}{T} \sum_{t=1}^{T} \begin{pmatrix}
\tilde{r}_{t+1} - \Delta_r \tilde{Z}_t \\
\tilde{r}_{t+1}^m - \delta_m \tilde{Z}_t \\
\tilde{r}_{t+1} - (\tilde{\epsilon}_{t+1} \tilde{u}_{t+1}) \tilde{Z}_t' \psi_m
\end{pmatrix},$$

(38)

has a block diagonal structure and it follows, therefore, that the sequential estimator for $\psi_m$ is fully efficient.

The analytic estimator of $\psi_m$ is the solution to the set of linear equations:

$$\hat{D}' \hat{S}^{-1} \left( \frac{1}{T} \sum_{t=1}^{T} \tilde{r}_{t+1} - (\tilde{\epsilon}_{t+1} \tilde{u}_{t+1}) \tilde{Z}_t' \psi_m \right) = 0,$$

(39)

where $\hat{D}$ and $\hat{S}$ denote consistent estimators of the $D$ and $S$ matrices. The $D$ matrix for (37) is simply $-E[(\tilde{\epsilon}_{t+1} \tilde{u}_{t+1}) \tilde{Z}_t']$. Thus (39) can be written as:

$$\hat{D}' \hat{S}^{-1} \hat{\mu}_r + \hat{D}' \hat{S}^{-1} \hat{D} \psi_m = 0,$$

(40)

which implies that:

$$\hat{\psi}_m = - (\hat{D}' \hat{S}^{-1} \hat{D})^{-1} \hat{D}' \hat{S}^{-1} \hat{\mu}_r.$$

(41)

Of course this is the same formula that we derived earlier. The only difference is that the $\hat{D}$ and $\hat{S}$ matrices take a different form than before.
The analysis for conditional models with more than one factor is just as simple. First we let each element of the $K \times 1$ vectors $E[\tilde{s}_{t+1} \mid \mathcal{F}_t]$ and $\lambda_t$ be linear in the instruments:

$$\tilde{s}_{t+1} = \Delta'_t \hat{Z}_t + \tilde{u}_{t+1}$$
$$\lambda_t = (I_K \otimes \hat{Z}'_t)\psi$$

(42)

Then we form the set of sample moment conditions:

$$g_T(\hat{u}_{t+1}, \hat{e}_{t+1}, \tilde{x}_{t+1}, \psi) = \frac{1}{T} \sum_{t=1}^{T} (\tilde{r}_{t+1} - (\hat{e}_{t+1} \hat{u}'_{t+1}) \otimes \hat{Z}'_t)\psi,$$

(43)

where $\hat{e}_{t+1}$ and $\hat{u}_{t+1}$ denote the fitted values of $\tilde{e}_{t+1}$ and $\tilde{u}_{t+1}$ from OLS. Armed with (43) it is straightforward to compute the analytic estimator and test statistic using the procedures outlined earlier.

The final issue that we want to address is how to deal with the returns to dynamic trading strategies within our analytic framework for estimating and testing factor pricing models. When we incorporate dynamic trading strategies into the analysis the pricing restriction takes the form:

$$E[\tilde{r}_{t+1} \otimes \hat{Z}_t - ((\hat{e}_{t+1} \otimes \hat{Z}_t)\hat{u}'_{t+1}) \otimes \hat{Z}'_t)\psi = 0.$$  

(44)

We can estimate the parameters of (44) using the set of sample moment conditions:

$$g_T(\hat{u}_{t+1}, \hat{e}_{t+1}, \tilde{x}_{t+1}, \psi) = \frac{1}{T} \sum_{t=1}^{T} (\tilde{r}_{t+1} \otimes \hat{Z}_t - ((\hat{e}_{t+1} \otimes \hat{Z}_t)\hat{u}'_{t+1}) \otimes \hat{Z}'_t)\psi.$$  

(45)

Although the specification shown in (45) is a bit more complicated than that in (43) it is still linear in the parameters. Thus the efficient GMM estimator of $\psi$ and the analytic test statistic take the same form as before with the $D$ matrix given by $-E[((\hat{e}_{t+1} \otimes \hat{Z}_t)\hat{u}'_{t+1}) \otimes \hat{Z}'_t]$ and $g_T(\hat{u}_{t+1}, \hat{e}_{t+1}, \tilde{x}_{t+1}, \psi)$ defined accordingly.

4. Data and Empirical Results

Our data consists of monthly holding periods returns on portfolios of common stock over the period from July 1963 to December 1993 (366 observations). The first set of 25 portfolios is constructed from firms listed on the New York Stock Exchange (NYSE) using data provided by the Center for Research in Security Prices (CRSP). We sort
the NYSE firms into 25 industry groups based on their two and three digit standard industrial classification (SIC) codes. A firm is included in its respective industry for every month that a return, price per share, and number of shares outstanding are available. The return on each portfolio represents a value-weighted average of the returns for individual firms.

We also employ a second set of 25 portfolios that are formed on the basis of market capitalization and book equity to market equity (BE/ME). The criteria used to form these portfolios are identical to those used by Fama and French (1993). Thus, the only difference between our portfolios and those of Fama and French (1993) is that we have an additional two years of data. A brief description of both the size-BE/ME portfolios and the industry portfolios is contained in Panel A of Table 1.

4.1. The Factors

Our empirical tests focus on the CAPM and the three factor Fama and French model. The Fama and French (1993) factors are the excess return on the value-weighted market portfolio, the return on a zero-investment portfolio that is designed to capture systematic differences in risk between small capitalization and large capitalization firms, and the return a zero-investment portfolio that is designed to capture systematic differences in risk between high and low book-to-market equity firms.

4.2. The Instruments

We use three instrumental variables in our tests: a term premium, a default risk spread, and a dividend yield. The justification for choosing these three instruments is that a number of studies have shown that they appear to have the ability to predict monthly holding period returns on stock and bond portfolios. Our term premium is the one-month return on a three-month Treasury bill less the return on the one-month Treasury bill (TERM). The data used to construct this instrument are provided by CRSP. Our measure of default risk is the yield spread between bonds rated Baa and Aaa (JUNK) by Moody’s Investor Services, and our dividend yield is the annualized dividend yield on Standard and Poor’s 500 stock index (XDIV) measured in excess of
the one-month Treasury bill rate. These two instruments are constructed using data drawn from the Federal Reserve Board.

4.3. Summary Statistics

Table 2 provides descriptive statistics for the industry portfolios, the size-BE/ME portfolios, the factors, and instrumental variables. Note that the portfolio returns in Panel A of this table are measured in excess of the one-month Treasury bill rate. Sorting the firms into different industries produces a fairly large dispersion in both average returns and standard deviations. The highest and lowest average returns of 0.84 and 0.22 percent per month correspond to the Entertainment and Metals industries. The highest and lowest standard deviations of 7.75 and 3.95 percent per month are associated with the Construction and Utility industries respectively.

The size-BE/ME portfolios also exhibit a wide dispersion in average returns and standard deviations. As expected, the size and book-to-market effects noted by Fama and French (1991) are clearly evident. Small capitalization firms appear to have higher returns on average than large capital firms, and firms with high BE/ME seem to have higher returns on average than firms with low BE/ME. The mean and standard deviations of the factor realizations shown in Panel B of Table 2 are very similar to those reported by Fama and French (1993).

4.4. Return Predictability

Before we consider the results of the estimation, it is useful to briefly examine the role of predictability in the analytic tests of the factor pricing models. Panel A of Table 3 summarizes the results from fitting linear forecasting models to the excess portfolio returns using ordinary least squares. The regression R-square for the industry portfolios ranges from a high of 11.9% for Motor Vehicles to a low of 5.0% for Utilities. For the size-BE/ME portfolios it ranges from a high of 12.2% to a low of 7.2%. Note that in every case the heteroskedasticity-consistent Wald test rejects the null of no predictability at standard levels of significance.

The results of fitting linear forecasting models to the factors are shown in Panel
B of Table 3. The excess return on the NYSE index has an R-square of 10.2%. This is considerably higher than that for either the mimicking portfolio for size (4.5%) or the mimicking portfolio for BE/ME (0.7%). Based on these results the BE/ME factor does not appear to be predictable. On balance, however, the empirical evidence indicates that excess portfolio returns are to some extent predictable over time. As a consequence, dynamic trading strategies that exploit this predictability are likely to play an important role in tests of factor pricing models.

4.5. Analytic Tests of the CAPM

Table 4 reports the results of the analytic estimation and tests of the CAPM. Panel A is for the specification where the price of market covariance risk (λ) is restricted to be constant, and dynamic trading strategies are excluded from the set of test assets. The industry portfolios yield an estimate of 1.69 for λ, and the analytic test for overidentifying restrictions provides no evidence against the model. The size-BE/ME portfolios, on the other hand, yield an estimate of 2.72 for λ, and we can reject the model at the 1% level.

We can get a better feel for why the results are different for the two data sets by examining Figure 1. This figure shows the relation between the mean monthly excess return on the portfolios and their covariance with the market. For the industry portfolios the relation appears to be generally consistent with the CAPM. High risk, as measured by the portfolio's covariance with the market, is typically associated with a high average return. But this does appear to be the case for the size-BE/ME portfolios. As in Fama and French (1991), there does not seem to be any clear relation between the covariance and the average return.

Panel B of Table 4 shows the results of adding the dynamic trading strategies associated with instrumental variables to the set of assets under consideration. The findings for the industry portfolios are much the same as before. Our estimate for λ increases to 2.60 with a standard error of 0.4 and the analytic test for overidentifying restrictions provides no evidence against the model. The estimate of 2.98 for λ for size-BE/ME portfolios is quite close to that shown in Panel A. Unlike before, however,
the analytic test for overidentifying restrictions does not reject the model.

We can shed additional light on these results by examining the tradeoff between risk and average return for the portfolios and dynamic trading strategies. This tradeoff is illustrated in Figure 2.\textsuperscript{10} Overall, adding the dynamic trading strategies to the analysis does little to alter the relation described earlier. For the industry portfolios, the covariance with the market appears, at least to some extent, to explain differences in average returns. But for the size-BE/ME portfolios the relation implied by the CAPM does not appear to hold. In seems, therefore, that the test for overidentifying may suffer from a lack of power when dynamic trading strategies are included in the set of test assets.

The final panel of Table 4 looks at a specification of the CAPM where $\lambda$ is allowed to vary over time. More specifically, $\lambda$ is modelled as a linear function of the set of three instrumental variables. Of course this version of the CAPM nests the constant $\lambda$ model, so it is not surprising that the analytic test for overidentifying restrictions shows no evidence of misspecification. There is ample evidence, however, that the market price of covariance risk changes through time. Each of the coefficients for the three instruments is greater than two standard errors from zero for both the industry and size-BE/ME portfolios.

Figure 3 illustrates the time-series behavior of the fitted price of market covariance risk. Although the test for overidentifying reveals no evidence of misspecification, the fact that the fitted price of risk is often negative is clearly inconsistent with the implications of the CAPM. Note also that fitted price of risk for the industry portfolios does not appear to coincide with that for the size-BE/ME portfolios. If the CAPM were valid, then in general we would expect that the fitted price of risk to be quite similar for the two data sets.

On balance, the empirical evidence does not provide a lot of support for the CAPM. Although it is true that in most cases we cannot reject the model based on the test for overidentifying restrictions, it appears that this is in large part due to

\textsuperscript{10}To construct this plot, the instrumental variables are scaled so that they have a unit mean. This implies that the payoffs on the dynamic trading strategies can be interpreted as gross returns on managed portfolios.
a lack of power. Now we want to explore whether the situation improves when we consider a multi-factor specification.

4.6. Analytic Tests of the Three Factor Model

Table 5 reports the results of the analytic estimation and tests of the three factor model. Panel A is for the specification where the price of market covariance risk ($\lambda_1$), the price of size covariance risk ($\lambda_2$), and the price of BE/ME covariance risk ($\lambda_3$) are restricted to be constant. Note also that dynamic trading strategies are excluded from the set of assets under consideration. The industry portfolios yield an estimates for $\lambda_1$, $\lambda_2$, and $\lambda_3$ of 1.79, -0.52, and 0.14 respectively. None of these is more that two standard errors from zero, and the analytic test for overidentifying restrictions provides no evidence against the model. When we move to the size-BE/ME portfolios, the corresponding estimates are 3.80, 3.69, and 8.10, each of which is more than two standard errors from zero. But the test for overidentifying restrictions rejects the model at the 10% level.

Panel B of Table 5 shows the results of adding the dynamic trading strategies to the test assets. This leads to more precise estimates of the prices of risk. Although we cannot reject the model for either set of portfolios, there is some evidence of misspecification. The estimated price of BE/ME covariance risk for the industry portfolios is reliably negative while that for the size-BE/ME portfolios is reliably positive. This is inconsistent with the view that BE/ME represents a legitimate proxy for a priced factor.

Panel C of Table 5 looks at a specification of the CAPM where all three prices of risk are allowed to vary over time. Again, each $\lambda_i$ is modelled as a linear function of the set of three instrumental variables. The results again provide strong indications of time-varying risk premia. Figures 4 through 6 illustrate the time-series behavior of the fitted prices of covariance risk for the three factors. As in the case for the CAPM, the fitted price of market covariance risk is frequently negative. This holds true for the other two prices of risk as well.

Of course the size and BE/ME factors are returns on zero investment portfolios,
so it is not clear that a negative price of risk is necessarily an indication of model misspecification. Still, if the three factor is well specified, then the fitted prices of risk should exhibit similar time series behavior for the two set of portfolios. In Figures 4 through 6 this does not appear to be the case. Thus, the overall picture for the three factor model is the same as that for the CAPM. Although the analytic test for overidentifying restrictions does not reject the model, there are other indications that the model is misspecified.

5. Conclusions

It is clear that the GMM approach to estimation and testing has been a great benefit to asset pricing research. The one potential drawback of using GMM is that it requires a two-stage, nonlinear numerical minimization. Thus, for large systems, the computational burden of GMM estimation is often nontrivial, and problems with the minimization algorithm, such as a failure to converge, are not uncommon. This paper develops a simple procedure for constructing GMM tests of both linear and nonlinear factor pricing models that bypass the numerical minimization procedure. No restrictions on conditionally heteroskedasticity or serial correlation are required, and the payoffs to dynamic trading strategies can be handled in a straightforward manner. As a consequence, we are able to construct analytic tests of factor pricing models that cover almost any scenario that is likely to be encountered in practice.
References


Table 1. The Portfolios, Factors, and Instruments

Panel A: The Portfolios

<table>
<thead>
<tr>
<th>Number</th>
<th>Industry</th>
<th>SIC Codes</th>
<th>Size Quintile</th>
<th>BE/ME Quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aerospace</td>
<td>372, 376</td>
<td>Small</td>
<td>Low</td>
</tr>
<tr>
<td>2</td>
<td>Transportation</td>
<td>40, 45</td>
<td>Small</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Banking</td>
<td>60</td>
<td>Small</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Building Materials</td>
<td>24, 32</td>
<td>Small</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Chemicals &amp; Plastics</td>
<td>281, 282, 286-289, 308</td>
<td>Small</td>
<td>Big</td>
</tr>
<tr>
<td>6</td>
<td>Construction</td>
<td>15-17</td>
<td>2</td>
<td>Low</td>
</tr>
<tr>
<td>7</td>
<td>Entertainment</td>
<td>365, 483, 484, 78</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>Food &amp; Beverages</td>
<td>20</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>Healthcare</td>
<td>283, 384, 385, 80</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Industrial Machinery</td>
<td>351-356</td>
<td>2</td>
<td>Big</td>
</tr>
<tr>
<td>11</td>
<td>Insurance &amp; Real Estate</td>
<td>63-65</td>
<td>3</td>
<td>Low</td>
</tr>
<tr>
<td>12</td>
<td>Investments</td>
<td>62, 67</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>Metals</td>
<td>33</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>Mining</td>
<td>10, 12, 14</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>Motor Vehicles</td>
<td>371, 551, 552</td>
<td>3</td>
<td>Big</td>
</tr>
<tr>
<td>16</td>
<td>Paper</td>
<td>26</td>
<td>4</td>
<td>Low</td>
</tr>
<tr>
<td>17</td>
<td>Petroleum</td>
<td>13, 29</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>Printing &amp; Publishing</td>
<td>27</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>Professional Services</td>
<td>73, 87</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>Retailing</td>
<td>53, 56, 57, 59</td>
<td>4</td>
<td>Big</td>
</tr>
<tr>
<td>21</td>
<td>Semiconductors</td>
<td>357, 367</td>
<td>Large</td>
<td>Low</td>
</tr>
<tr>
<td>22</td>
<td>Telecommunications</td>
<td>366, 381, 481, 482, 489</td>
<td>Large</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>Textiles &amp; Apparel</td>
<td>22, 23</td>
<td>Large</td>
<td>3</td>
</tr>
<tr>
<td>24</td>
<td>Utilities</td>
<td>49</td>
<td>Large</td>
<td>4</td>
</tr>
<tr>
<td>25</td>
<td>Wholesaling</td>
<td>50, 51</td>
<td>Large</td>
<td>Big</td>
</tr>
</tbody>
</table>

Panel B: Factors and Instruments

| Factor 1      | Excess return on NYSE index |
| Factor 2      | Return on small minus big (SML) portfolio |
| Factor 3      | Return on high BE/ME minus low BE/ME (HML) portfolio |
| Instrument 1  | Excess return on 3-month T-bill |
| Instrument 2  | Yield on Baa bonds minus yield on Aaa bonds |
| Instrument 3  | Excess dividend yield on S&P composite stock index |

The industry portfolios are constructed from firms listed on the NYSE using data provided by CRSP. We sort these firms into 25 industry groups based on their two and three digit SIC codes. A firm is included in its respective industry for every month that a return, price per share, and number of shares outstanding are available. Portfolio returns are a value-weighted average of the monthly returns for the individual firms. The size-BE/ME portfolios, the size factor, and the BE/ME factor are constructed as in Fama and French (1993). The data cover the period from July 1963 to December 1993 (366 observations).
Table 2. Summary Statistics

Panel A: Monthly Excess Returns

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Industry Portfolios</th>
<th>Size – BE/ME Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>1</td>
<td>0.57</td>
<td>6.56</td>
</tr>
<tr>
<td>2</td>
<td>0.48</td>
<td>6.56</td>
</tr>
<tr>
<td>3</td>
<td>0.35</td>
<td>6.41</td>
</tr>
<tr>
<td>4</td>
<td>0.50</td>
<td>6.19</td>
</tr>
<tr>
<td>5</td>
<td>0.42</td>
<td>5.32</td>
</tr>
<tr>
<td>6</td>
<td>0.64</td>
<td>7.75</td>
</tr>
<tr>
<td>7</td>
<td>0.84</td>
<td>6.74</td>
</tr>
<tr>
<td>8</td>
<td>0.68</td>
<td>4.57</td>
</tr>
<tr>
<td>9</td>
<td>0.61</td>
<td>5.29</td>
</tr>
<tr>
<td>10</td>
<td>0.40</td>
<td>5.98</td>
</tr>
<tr>
<td>11</td>
<td>0.51</td>
<td>5.91</td>
</tr>
<tr>
<td>12</td>
<td>0.49</td>
<td>4.60</td>
</tr>
<tr>
<td>13</td>
<td>0.22</td>
<td>6.16</td>
</tr>
<tr>
<td>14</td>
<td>0.61</td>
<td>5.44</td>
</tr>
<tr>
<td>15</td>
<td>0.46</td>
<td>5.92</td>
</tr>
<tr>
<td>16</td>
<td>0.45</td>
<td>5.42</td>
</tr>
<tr>
<td>17</td>
<td>0.51</td>
<td>5.26</td>
</tr>
<tr>
<td>18</td>
<td>0.69</td>
<td>5.97</td>
</tr>
<tr>
<td>19</td>
<td>0.62</td>
<td>7.06</td>
</tr>
<tr>
<td>20</td>
<td>0.61</td>
<td>6.09</td>
</tr>
<tr>
<td>21</td>
<td>0.26</td>
<td>5.66</td>
</tr>
<tr>
<td>22</td>
<td>0.37</td>
<td>4.16</td>
</tr>
<tr>
<td>23</td>
<td>0.53</td>
<td>6.25</td>
</tr>
<tr>
<td>24</td>
<td>0.32</td>
<td>3.95</td>
</tr>
<tr>
<td>25</td>
<td>0.60</td>
<td>6.26</td>
</tr>
</tbody>
</table>

Panel B: Factors and Instruments

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>0.43</td>
<td>4.39</td>
<td>-22.35</td>
<td>16.01</td>
</tr>
<tr>
<td>Factor 2</td>
<td>0.28</td>
<td>2.89</td>
<td>-9.80</td>
<td>10.77</td>
</tr>
<tr>
<td>Factor 3</td>
<td>0.46</td>
<td>2.59</td>
<td>-9.99</td>
<td>8.91</td>
</tr>
<tr>
<td>Instrument 1</td>
<td>0.06</td>
<td>0.11</td>
<td>-0.30</td>
<td>0.65</td>
</tr>
<tr>
<td>Instrument 2</td>
<td>0.09</td>
<td>0.04</td>
<td>0.03</td>
<td>0.22</td>
</tr>
<tr>
<td>Instrument 3</td>
<td>-0.20</td>
<td>0.17</td>
<td>-0.93</td>
<td>0.04</td>
</tr>
</tbody>
</table>

The data cover the period from July 1963 to December 1993 (366 observations). All returns are measured in excess of the one month T-bill rate. The industry portfolios are constructed from firms listed on the NYSE using data provided by CRSP. We sort these firms into 25 industry groups based on their two and three digit SIC codes. A firm is included in its respective industry for every month that a return, price per share, and number of shares outstanding are available. Portfolio returns are a value-weighted average of the monthly returns for the individual firms. The size-BE/ME portfolios, the size factor, and the BE/ME factor are constructed as in Fama and French (1993).
### Table 3. Predictability of Portfolio Returns and Factors

#### Panel A: Monthly Excess Returns

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>R-Square</th>
<th>( \chi^2 )</th>
<th>Prob &gt; ( \chi^2 )</th>
<th>R-Square</th>
<th>( \chi^2 )</th>
<th>Prob &gt; ( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.2</td>
<td>31.7</td>
<td>0.000</td>
<td>9.1</td>
<td>30.5</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>8.8</td>
<td>23.4</td>
<td>0.000</td>
<td>9.8</td>
<td>32.3</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>5.1</td>
<td>15.1</td>
<td>0.002</td>
<td>11.3</td>
<td>37.6</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>10.7</td>
<td>29.5</td>
<td>0.000</td>
<td>11.5</td>
<td>38.4</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>8.9</td>
<td>24.7</td>
<td>0.000</td>
<td>11.3</td>
<td>35.8</td>
<td>0.000</td>
</tr>
<tr>
<td>6</td>
<td>10.4</td>
<td>35.9</td>
<td>0.000</td>
<td>8.9</td>
<td>29.2</td>
<td>0.000</td>
</tr>
<tr>
<td>7</td>
<td>6.1</td>
<td>19.1</td>
<td>0.000</td>
<td>10.7</td>
<td>37.8</td>
<td>0.000</td>
</tr>
<tr>
<td>8</td>
<td>6.3</td>
<td>22.2</td>
<td>0.000</td>
<td>10.7</td>
<td>34.5</td>
<td>0.000</td>
</tr>
<tr>
<td>9</td>
<td>3.2</td>
<td>8.0</td>
<td>0.030</td>
<td>11.7</td>
<td>38.4</td>
<td>0.000</td>
</tr>
<tr>
<td>10</td>
<td>9.5</td>
<td>32.8</td>
<td>0.000</td>
<td>10.0</td>
<td>31.5</td>
<td>0.000</td>
</tr>
<tr>
<td>11</td>
<td>5.2</td>
<td>19.6</td>
<td>0.000</td>
<td>9.3</td>
<td>31.2</td>
<td>0.000</td>
</tr>
<tr>
<td>12</td>
<td>8.0</td>
<td>23.2</td>
<td>0.000</td>
<td>12.2</td>
<td>40.3</td>
<td>0.000</td>
</tr>
<tr>
<td>13</td>
<td>5.8</td>
<td>16.9</td>
<td>0.001</td>
<td>11.7</td>
<td>38.8</td>
<td>0.000</td>
</tr>
<tr>
<td>14</td>
<td>8.2</td>
<td>22.6</td>
<td>0.000</td>
<td>10.7</td>
<td>32.1</td>
<td>0.000</td>
</tr>
<tr>
<td>15</td>
<td>11.9</td>
<td>46.4</td>
<td>0.000</td>
<td>8.0</td>
<td>20.5</td>
<td>0.000</td>
</tr>
<tr>
<td>16</td>
<td>8.5</td>
<td>22.8</td>
<td>0.000</td>
<td>9.0</td>
<td>28.7</td>
<td>0.000</td>
</tr>
<tr>
<td>17</td>
<td>5.4</td>
<td>14.3</td>
<td>0.030</td>
<td>10.9</td>
<td>34.2</td>
<td>0.000</td>
</tr>
<tr>
<td>18</td>
<td>9.4</td>
<td>34.0</td>
<td>0.000</td>
<td>10.8</td>
<td>36.8</td>
<td>0.000</td>
</tr>
<tr>
<td>19</td>
<td>7.4</td>
<td>26.7</td>
<td>0.000</td>
<td>9.1</td>
<td>27.3</td>
<td>0.000</td>
</tr>
<tr>
<td>20</td>
<td>8.8</td>
<td>29.1</td>
<td>0.000</td>
<td>9.2</td>
<td>26.8</td>
<td>0.000</td>
</tr>
<tr>
<td>21</td>
<td>9.2</td>
<td>32.5</td>
<td>0.000</td>
<td>7.2</td>
<td>24.6</td>
<td>0.000</td>
</tr>
<tr>
<td>22</td>
<td>5.5</td>
<td>15.1</td>
<td>0.002</td>
<td>8.6</td>
<td>29.4</td>
<td>0.000</td>
</tr>
<tr>
<td>23</td>
<td>10.8</td>
<td>38.4</td>
<td>0.000</td>
<td>8.0</td>
<td>23.6</td>
<td>0.000</td>
</tr>
<tr>
<td>24</td>
<td>5.0</td>
<td>13.7</td>
<td>0.003</td>
<td>7.7</td>
<td>20.1</td>
<td>0.000</td>
</tr>
<tr>
<td>25</td>
<td>9.1</td>
<td>31.8</td>
<td>0.000</td>
<td>7.2</td>
<td>22.3</td>
<td>0.000</td>
</tr>
</tbody>
</table>

#### Panel B: Monthly Factor Realizations

<table>
<thead>
<tr>
<th>R-Square</th>
<th>( \chi^2 )</th>
<th>Prob &gt; ( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>10.2</td>
<td>31.1</td>
</tr>
<tr>
<td>Factor 2</td>
<td>4.5</td>
<td>16.4</td>
</tr>
<tr>
<td>Factor 3</td>
<td>0.7</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Table 3 summarizes the observed ability to predict the excess portfolio returns and factors using a linear forecasting model. The forecasting model takes the form:

\[
\hat{r}_{i,t+1} = \delta_0 + \delta_{11}(TERM_t) + \delta_{i2}(JUNK_t) + \delta_{i3}(XDIV_t) + \tilde{e}_{i,t+1},
\]

where TERM is the return on a three-month T-bill less the one-month T-bill rate, JUNK is the yield spread between Baa and Aaa bonds, and XDIV is the dividend yield on the S&P composite index less the one-month T-bill rate. The data cover the period from July 1963 to December 1993 (366 observations).
Table 4. Analytic Tests of the CAPM

Panel A: The CAPM

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Industry Portfolios</th>
<th></th>
<th>Size – BE/ME Portfolios</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>1.69</td>
<td>1.08</td>
<td>2.72</td>
<td>1.13</td>
</tr>
<tr>
<td>Analytic Test Statistic</td>
<td>13.6</td>
<td></td>
<td>43.6</td>
<td></td>
</tr>
<tr>
<td>Probability Value</td>
<td>0.955</td>
<td></td>
<td>0.009</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: The CAPM with Dynamic Trading Strategies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>2.60</td>
<td>0.40</td>
<td>2.98</td>
<td>0.56</td>
</tr>
<tr>
<td>Analytic Test Statistic</td>
<td>77.1</td>
<td></td>
<td>86.6</td>
<td></td>
</tr>
<tr>
<td>Probability Value</td>
<td>0.950</td>
<td></td>
<td>0.809</td>
<td></td>
</tr>
</tbody>
</table>

Panel C: The Conditional CAPM with Dynamic Trading Strategies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_0$</td>
<td>13.0</td>
<td>1.9</td>
<td>14.1</td>
<td>2.7</td>
</tr>
<tr>
<td>$\psi_1$</td>
<td>-1196.0</td>
<td>317.9</td>
<td>-8503.3</td>
<td>864.1</td>
</tr>
<tr>
<td>$\psi_2$</td>
<td>-12545.0</td>
<td>1789.4</td>
<td>34185.5</td>
<td>2940.4</td>
</tr>
<tr>
<td>$\psi_3$</td>
<td>-1201.2</td>
<td>420.2</td>
<td>16395.1</td>
<td>1026.4</td>
</tr>
<tr>
<td>Analytic Test Statistic</td>
<td>77.3</td>
<td></td>
<td>79.6</td>
<td></td>
</tr>
<tr>
<td>Probability Value</td>
<td>0.919</td>
<td></td>
<td>0.888</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 reports the results of the analytic GMM tests of the CAPM. The moment conditions take the form:

$$h_{t+1} \equiv [\tilde{r}_{t+1} - (\tilde{r}_{t+1} - \Delta_\tau \tilde{Z}_t)(\tilde{r}_{t+1} - \delta_m \tilde{Z}_t)\tilde{Z}_t'\psi_m] \otimes \tilde{Z}_t,$$

where the $4 \times 25$ matrix $\Delta_\tau$, the $4 \times 1$ vector $\delta_m$, and the $4 \times 1$ vector $\psi_m$ are unknown parameters. Panel A covers the most restrictive scenario where the price of market covariance risk, $\lambda_t \equiv \tilde{Z}_t'\psi_m$ is assumed to be constant, and the dynamic trading strategies associated with instrumental variables are excluded from the set of test assets (i.e., we only use the first 25 elements of $h_{t+1}$). In Panel B, we add the dynamic trading strategies. Panel C covers the most general scenario where the price of market covariance risk is allowed to vary over time. The parameters are estimated sequentially. The data set covers the period from July 1963 to December 1993 (366 observations).
Table 5. Analytic Tests of the Three Factor Model

Panel A: The Three Factor Model

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>1.79</td>
<td>1.21</td>
<td>3.80</td>
<td>1.16</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-0.52</td>
<td>2.50</td>
<td>3.69</td>
<td>1.70</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.14</td>
<td>2.77</td>
<td>8.10</td>
<td>1.94</td>
</tr>
<tr>
<td>Analytic Test Statistic</td>
<td>13.2</td>
<td></td>
<td>31.1</td>
<td></td>
</tr>
<tr>
<td>Probability Value</td>
<td>0.927</td>
<td></td>
<td>0.094</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: The Three Factor Model with Dynamic Trading Strategies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>2.31</td>
<td>0.52</td>
<td>3.92</td>
<td>0.60</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-1.86</td>
<td>1.39</td>
<td>2.93</td>
<td>0.82</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>-2.78</td>
<td>0.95</td>
<td>10.11</td>
<td>0.80</td>
</tr>
<tr>
<td>Analytic Test Statistic</td>
<td>76.9</td>
<td></td>
<td>83.2</td>
<td></td>
</tr>
<tr>
<td>Probability Value</td>
<td>0.935</td>
<td></td>
<td>0.840</td>
<td></td>
</tr>
</tbody>
</table>

Panel C: The Conditional Three Factor Model with Dynamic Trading Strategies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_{10}$</td>
<td>4.2</td>
<td>2.6</td>
<td>-20.2</td>
<td>3.6</td>
</tr>
<tr>
<td>$\psi_{11}$</td>
<td>1912.6</td>
<td>723.6</td>
<td>-898.0</td>
<td>703.6</td>
</tr>
<tr>
<td>$\psi_{12}$</td>
<td>-11989.3</td>
<td>2708.5</td>
<td>50005.8</td>
<td>4080.8</td>
</tr>
<tr>
<td>$\psi_{13}$</td>
<td>-6624.4</td>
<td>524.2</td>
<td>11279.4</td>
<td>1026.4</td>
</tr>
<tr>
<td>$\psi_{20}$</td>
<td>20.9</td>
<td>5.5</td>
<td>4.9</td>
<td>4.6</td>
</tr>
<tr>
<td>$\psi_{21}$</td>
<td>19930.2</td>
<td>1018.9</td>
<td>-18755.5</td>
<td>1828.0</td>
</tr>
<tr>
<td>$\psi_{22}$</td>
<td>-40079.0</td>
<td>6431.5</td>
<td>15844.6</td>
<td>4355.0</td>
</tr>
<tr>
<td>$\psi_{23}$</td>
<td>1646.2</td>
<td>948.4</td>
<td>7.8</td>
<td>1223.5</td>
</tr>
<tr>
<td>$\psi_{30}$</td>
<td>-30.0</td>
<td>6.0</td>
<td>2.8</td>
<td>5.8</td>
</tr>
<tr>
<td>$\psi_{31}$</td>
<td>3791.1</td>
<td>1265.9</td>
<td>-6099.9</td>
<td>1584.3</td>
</tr>
<tr>
<td>$\psi_{32}$</td>
<td>11570.8</td>
<td>5335.2</td>
<td>-1676.6</td>
<td>3859.8</td>
</tr>
<tr>
<td>$\psi_{33}$</td>
<td>-8393.5</td>
<td>674.1</td>
<td>-3711.5</td>
<td>1138.3</td>
</tr>
<tr>
<td>Analytic Test Statistic</td>
<td>67.2</td>
<td></td>
<td>70.5</td>
<td></td>
</tr>
<tr>
<td>Probability Value</td>
<td>0.951</td>
<td></td>
<td>0.914</td>
<td></td>
</tr>
</tbody>
</table>

Table 5 reports the results of the analytic GMM tests of the three factor model. The moment conditions take the form:

$$ h_{t+1} = [\hat{r}_{t+1} - (\Delta r, \hat{z}_t)(\hat{s}_{t+1} - \Delta s, \hat{z}_t)(I_3 \otimes \hat{z}_t)\psi] \otimes \hat{z}_t, $$

where the $4 \times 25$ matrix $\Delta r$, the $4 \times 3$ matrix $\Delta s$, and the $12 \times 1$ vector $\psi$ are unknown parameters. Panel A covers the most restrictive scenario where the price of market covariance risk, the price of size covariance risk, and the price of BE/ME covariance risk are assumed to be constant, and the dynamic trading strategies associated with instrumental variables are excluded from the set of test assets (i.e., we only use the first 25 elements of $h_{t+1}$). In Panel B, we add the dynamic trading strategies. Panel C covers the most general scenario where all three prices of risk are allowed to vary over time. The parameters are estimated sequentially. The data set covers the period from July 1963 to December 1993 (366 observations).
Figure 1. Risk Versus Average Return in the CAPM

- **Industry Portfolios**
  - Mean Excess Monthly Return (%) vs. Covariance with Market

- **Size-BE/ME Portfolios**
  - Mean Excess Monthly Return (%) vs. Covariance with Market
Figure 2. Risk Versus Average Return in the CAPM with Dynamic Trading Strategies

Size-BE/ME Portfolios

Mean Excess Monthly Return (%)

Industry Portfolios

Mean Excess Monthly Return (%)
Figure 3. Fitted Price of Covariance Risk for Conditional CAPM
Figure 4. Fitted Price of Market Covariance Risk for the Conditional Three-Factor Model

Industry Portfolios

Size-BE/ME Portfolios
Figure 5. Fitted Price of Size Covariance Risk for the Conditional
Three-Factor Model

Industry Portfolios

Size-BE/ME Portfolios

Fitted Price of Covariance Risk

Fitted Price of Covariance Risk

63 66 69 72 75 78 81 84 87 90 93
Year

63 66 69 72 75 78 81 84 87 90 93
Year
Figure 6. Fitted Price of BE/ME Covariance Risk for the Conditional Three-Factor Model