A Cross-Sectional Test of an Investment-Based Asset Pricing Model

John H. Cochrane

A Cross-Sectional Test of an Investment-Based Asset Pricing Model

John H. Cochrane

University of Chicago, Federal Reserve Bank of Chicago, and National Bureau of Economic Research

I examine a factor pricing model for stock returns. The factors are returns on physical investment, inferred from investment data via a production function. I examine the model's ability to explain variation in expected returns across assets and over time. The model is not rejected. It performs about as well as the CAPM and the Chen, Roll, and Ross factor model, and it performs substantially better than a simple consumption-based model. I also provide an easy technique for estimating and testing dynamic, conditional asset pricing models—one simply includes factors and returns scaled by instruments in an unconditional estimate—and for comparing such models.

I. Introduction

The *investment return* is the marginal rate at which a firm can transfer resources through time by increasing investment today and decreasing it at a future date, leaving its production plan unchanged at all other dates. I examine whether cross-sectional and time-series variation in expected stock returns can be explained by investment returns, inferred from investment data via an adjustment cost production function.

Why? The identity of the macroeconomic risks that drive asset

prices and expected returns is a central question of finance, and an important question for macroeconomics. There is a wealth of tantalizing empirical evidence for a link between macroeconomic events and asset returns: many of the variables that forecast stock and bond returns also forecast investment growth or growth in gross national product; stock and bond returns are correlated with contemporaneous and subsequent economic activity; and expected returns are related to the covariances of returns with macroeconomic variables. But there is as yet no accepted economic explanation for this evidence.

Ideally, the consumption-based asset pricing model should provide a framework for digesting this empirical evidence and for identifying the macroeconomic shocks that drive asset returns. But the empirical performance of the consumption-based model has been disappointing, despite a two-decade specification search. Mechanically, this poor performance results from the fact that nondurable consumption growth barely moves over the business cycle, and it is poorly correlated with stock returns. Economically, the poor performance may be due to measurement error in consumption data; a poor understanding of the representative agent's utility function; or taxes, transactions costs, borrowing constraints, and other frictions that can "delink" many consumers' intertemporal consumption choices from high-frequency asset market movements.[1]

Much of finance studies reduced-form models that explain an asset's expected return by its covariance with other assets' returns, rather than covariance with macroeconomic risks. Though these models may successfully *describe* variation in expected returns, they will never *explain* it. To say that an asset's expected return varies over the business cycle because (say) the market expected return varies leaves unanswered the question, What real risks cause the market expected return to vary? Furthermore, fishing for asset return factors with no explicit connection to real risks can result in models that price assets by construction in a given data set (ex post mean-variance efficient portfolios).

In this context, the basic idea of this paper is to infer the presence of real macroeconomic shocks by watching firms' investment decisions, just as the consumption-based model tries to infer the presence of systematic shocks by watching consumption decisions.

This paper extends the work in Cochrane (1991). That paper explained time-series variation in the market return with a single investment return, inferred from gross fixed investment data with an ad-

---

[1] Cochrane and Hansen (1992) give a literature review and a summary of "puzzles" that characterize the empirical failure of the consumption-based model. Cochrane (1989) and Luttmer (1992) calculate effects of small frictions.

justment cost production function. It showed that some adjustment cost is necessary to produce investment return variation anything like that observed in market returns. It showed that variation in the expected market return is largely matched by variation in the expected investment return and that market returns and investment returns have the same association with subsequent economic activity. This paper explains *cross-sectional* as well as time-series variation in expected stock returns by reference to investment returns.

However, the factor pricing model studied in this paper is not a pure production-based asset pricing model. A pure production-based model uses *no* assumptions on preferences or restrictions on the space of asset returns and reads any asset return off a producer's first-order conditions, just as the consumption-based model uses no technology assumptions (i.e., is valid for any production technology) and reads asset prices from a consumer's first-order conditions. The standard production functions I use here do not have general cross-sectional asset pricing implications, since there is nothing a producer can do to transform goods *across states*.

The techniques I use to estimate and test dynamic, conditional factor models are derived from the work of Hansen (1982), Hansen and Singleton (1982), Hansen and Richard (1987), and Hansen and Jagannathan (1991*b*). Knez (1991) and Snow (1991) use similar techniques to study factor pricing models; Braun (1991) uses them to investigate consistent pricing of asset and investment returns and tests whether the inverse of a single investment return can explain a cross section of assets; and De Santis (1992) uses them to study international capital market integration.

II. Investment Returns: Definition and Construction

To construct investment returns from production data, I use adjustment cost technologies of the form

$$y_t = f(k_t, l_t) - c(i_t, k_t), \tag{1}$$

$$k_{t+1} = (1 - \delta)(k_t + i_t), \tag{2}$$

where y_t is output, $f(k_t, l_t)$ is the production function, k_t is capital stock, l_t is labor input, i_t is investment, δ is the depreciation rate, and $c(i_t, k_t)$ is the adjustment cost function. The adjustment cost reflects the fact that it is hard to produce in periods of high investment. For example, it is hard to write papers while a new computer is being installed. This is the standard sort of production function used to justify the *q*-theory of investment.

The one-period investment return is the amount of extra output the firm can sell at $t + 1$ if it invests an additional unit at t, leaving all variables at $t + 2, t + 3, \ldots$ unchanged. The Appendix goes through the algebra to show that the one-period investment return for the technology specified in (1)–(2) is given by

$$r^i_{t+1} = (1 - \delta) \frac{1 + f_k(t + 1) + c_i(t + 1) - c_k(t + 1)}{1 + c_i(t)}, \qquad (3)$$

where

$$c_i(t) \equiv \frac{\partial c(i_t, k_t)}{\partial i_t},$$

and so forth.

The denominator $1 + c_i(t)$ in equation (3) reflects the fact that some output is lost to adjustment costs when investment is increased at time t. This term is equal to marginal q: the marginal rate of transformation between installed and uninstalled capital. The extra time t investment gives rise to extra capital stock at $t + 1$: $f_k(t + 1)$ is the extra output that results from the extra capital stock, and $c_k(t + 1)$ captures the effect of extra capital at $t + 1$ on time $t + 1$ adjustment costs. At $t + 1$, the firm must *lower* investment to restore the capital stock at $t + 2$ to its original value. The lowered investment lowers $t + 1$ adjustment costs, and this means that more can be sold. The term $1 + c_i(t + 1)$ captures these effects.

The investment return is random: it depends on events at $t + 1$ as well as events at t. A positive productivity shock at time $t + 1$ implies an unexpectedly high return to investment from t to $t + 1$. Do not confuse the investment return with the expected investment return, the required return, or other ex ante concepts.

I use the following parametric specification of technology:

$$y_t = mpk\, k_t + mpl\, l_t - \frac{\eta}{2} \left(\frac{i_t}{k_t} \right) i_t. \qquad (4)$$

In this case, the investment return (3) becomes

$$r^i_{t+1} = (1 - \delta) \frac{1 + mpk + \eta(i_{t+1}/k_{t+1}) + (\eta/2)(i_{t+1}/k_{t+1})^2}{1 + \eta(i_t/k_t)}. \qquad (5)$$

Though this function is not pretty, the investment return it specifies is approximately proportional to growth in the investment/capital ratio or, since capital does not vary much, growth in investment.

For given values of the parameters $\{\eta, \delta, mpk\}$, I form investment/capital ratios by accumulating capital according to equation (2) start-

ing from the steady-state investment/capital ratio

$$\frac{i_0}{k_0} = \left[E\left(\frac{i_t}{i_{t-1}}\right) \Big/ (1 - \delta) \right] - 1.$$

Then, given η and mpk, I construct the investment returns from their definition (5).

One might object to the excessive simplicity of this technology. But the most obvious complications—such as taxes, declining marginal products, substitutability between capital and labor, and marginal productivity shocks—affect the dividend or one-period cash flow portion of the investment return. The adjustment cost influences the price change term ($1 + c_i$ or $1 + \eta[i/k]$, the marginal rate of transformation between installed and uninstalled capital), which dominates the investment return as it does stock returns.[2] Furthermore, most of these complications have low-frequency effects on the level of prices or q. By looking at investment and stock *returns*, we essentially first-difference out the effects of these complications. On the other hand, complications to the adjustment cost technology such as gestation lags may have first-order effects on the results.

III. Factor Pricing Model

What can one do with investment returns? If there are no arbitrage opportunities, then there is a stochastic discount factor m, such that any asset return r_j and investment return r_k^i obey (see, e.g., Ross 1978; Hansen and Richard 1987; Hansen and Jagannathan 1991b)

$$1 = E(mr_j); \quad 1 = E(mr_k^i). \tag{6}$$

The marginal utility growth of a nonsatiated owner of the firm is one such m; E can be interpreted as a conditional or unconditional expectation. I shall be specific about conditioning information below.

Braun (1991) tests two immediate implications of this fact. First, one can expand the space of returns on which one tests any asset pricing model (model for m) to include investment returns. In this way, one tests whether the asset pricing model can account for macroeconomic events as well as stock market events. Second, one can test for the absence of arbitrage or consistent pricing between the set of asset and investment returns by trying to construct nonnegative m's that satisfy equations (6).

---

[2] Braun (1991) found that tests of producer first-order conditions were insensitive to these issues in detailed experiments, for just the reasons given above. Sharathchandra (1991) models a concave technology with production function shocks but no adjustment costs and obtains an essentially constant investment return.

I concentrate on *asset* pricing. What can one learn about asset returns r from investment returns r^i? I study the hypothesis that a factor pricing model holds; namely, *the investment returns are factors for the asset returns*. More formally, an investment return factor pricing model says that there exists a discount factor that is a function of only the investment returns and yet prices both asset and investment returns:

$$m = \sum_k b_k r_k^i \tag{7}$$

satisfies equation (6) for all asset returns r_j and investment returns r_k^i.

The law of one price implies that there is always a discount factor m that is a linear combination of the investment and asset returns and that prices both. Mechanically, one can choose b's to construct

$$m = \sum_j b_j r_j + \sum_k b_k r_k^i$$

so as to satisfy equations (6) exactly, in sample or in population. The factor pricing model restriction is that the asset returns can be excluded from this construction. This restriction is equivalent to the statement that expected excess returns are proportional to the covariance or betas of any return with the investment returns. It is also equivalent to the statement that the investment returns span the mean-variance frontier of investment and asset returns.

Why should investment returns be factors for asset returns? Factor pricing models are derived by arbitrage assumptions or by preference assumptions. We can assume that the firms on the New York Stock Exchange (NYSE) are claims to different combinations of N production technologies, plus idiosyncratic components that have small prices. Alternatively, we can invoke preference assumptions under which the returns on the N active production processes, which are the only nondiversifiable payoffs in the economy and add up to aggregate wealth, drive marginal utility growth and hence price assets (see, e.g., Brock 1982; Cox, Ingersoll, and Ross 1985).[3]

The number and nature of the intertemporal technologies that drive asset returns or, equivalently, the appropriate level of aggregation of the capital stock are a modeling choice. "My car" and "your

---

[3] Time-separable preferences are the central assumption. If preferences are not time-separable, then past investment returns could affect current asset returns. One could, of course, account for potential nonseparabilities by including past investment returns as additional factors. With general preferences in discrete time, nonlinear functions of investment returns might also enter m; one can regard linearity either as an assumption on preferences or as a first-order approximation as in the log utility example below.

car" are both ways of getting consumption services from today to tomorrow, but one hopes that their behavior across states of nature that affect asset returns is sufficiently similar that we can aggregate them into "cars." However, there is no reason to believe a priori that all the intertemporal investment opportunities in the economy can be summarized by one or two aggregated production functions. I follow the "spirit of the arbitrage pricing theory" and hope that only a few investment return factors will suffice, but this is an additional modeling assumption, not a prediction of theory. Models with highly disaggregated investment opportunities may turn out to be more useful for some purposes.

Following the factor pricing tradition, I estimate the loadings of the investment return factors in the discount factor—the b's—as free parameters. In complete general equilibrium models—models in which we can solve for consumption and asset returns ex post, not just state $1 = E(mR)$—the b's can be derived from economic theory as well. For example, in the standard one-sector stochastic growth model with log utility, Cobb-Douglas production, and full depreciation, we have[4]

$$m_{t+1} = \beta \frac{c_t}{c_{t+1}} = \frac{1}{r^i_{t+1}} \approx 1 - (r^i_{t+1} - 1),$$

where β is the subjective discount factor. This model predicts $b_0 = 2$, $b_1 = -1$. I estimate the b's rather than construct a complete general equilibrium model, in order to focus on production technologies and firm behavior rather than preferences and sources of shocks. More theory is better only if it is the right theory.

IV.   Empirical Methods

A.   A GMM Test of Factor Pricing Models

The statement of the factor pricing model above maps naturally into the generalized method of moments (GMM) framework for estima-

---

[4] The model is

$$\max E \sum_{j=0}^{\infty} \beta^j \ln(c_j)$$

subject to $c_t + i_t = y_t = \lambda_t i_{t-1}^\alpha$, $\ln \lambda_t = \rho \ln \lambda_{t-1} + \epsilon_t$.

The investment return is $r^i_{t+1} = \alpha \lambda_{t+1} i_t^{\alpha-1} = \alpha y_{t+1}/i_t$. The solution to the model gives $c_t = (1 - \alpha\beta)y_t$ and $i_t = \alpha\beta y_t$. Substituting this solution into the investment return, we obtain

$$r^i_{t+1} = \frac{1}{\beta} \frac{c_{t+1}}{c_t} = \frac{1}{m_{t+1}}.$$

tion and testing. Let r_{t+1} denote a vector of returns and p_t denote a vector of prices. (The "price" of a return is one, and that of an excess return or difference between two returns is zero.) I suppress time subscripts when they are clear. The natural set of moment conditions to exploit is $p = E(mr)$ or

$$E(mr - p) = 0. \tag{8}$$

Let f_{t+1} denote a vector of pricing factors (one factor may be a constant). A linear factor pricing model is $m = f'b$, where b is a vector of coefficients.

Following the standard GMM procedure (Hansen 1982; Hansen and Singleton 1982), I estimate the parameters b to minimize a weighted combination of the sample moments (8). Using Hansen's notation, let E_T denote the sample mean, $E_T \equiv (1/T) \Sigma_{t=1}^T$; let W denote a weighting matrix; and denote the sample moments g_T,

$$g_T \equiv E_T(mr - p) = E_T(rf')b - E_T(p). \tag{9}$$

The GMM objective is to choose b to minimize a weighted sum of squares of the pricing errors across assets,

$$J_T = g_T'Wg_T. \tag{10}$$

Econometric issues aside, this objective is a natural and intuitive way to pick parameters in order to make the model fit as well as possible.

Since the parameters b enter linearly in the minimization, we can find their estimates analytically. Let D denote the matrix of cross–second moments between returns and factors:

$$D = \frac{\partial g_T}{\partial b} = E_T(rf').$$

Let $\hat{b}$ denote the estimate of b. Then the first-order conditions to the minimization of (10) are

$$\frac{\partial g_T'}{\partial b} Wg_T = D'W(D\hat{b} - E_Tp) = 0. \tag{11}$$

Solving, and assuming that at least one element of $E_t p$ is not equal to zero,[5] we get

$$\hat{b} = (D'WD)^{-1}D'WE_T(p). \tag{12}$$

---

[5] If all elements of p are zero, which occurs when only excess returns are used, then b is identified only up to a constant ($0 = E(mr) \Rightarrow 0 = E(2mr)$). In this case, one can impose one element of b arbitrarily (e.g., $b_0 = 1$) and solve for the others, one can add a normalization such as $1 = E(m)$ as an additional moment, or one can add a single return in levels to the system such as the real Treasury-bill return to obtain a $p \neq 0$. I follow the last strategy.

This estimate has a natural interpretation: $\hat{b}$ is the coefficient in a generalized least squares (GLS) cross-sectional regression of the mean price vector $E_T(\mathbf{p})$ on the second moments $\mathbf{D}$. Since the asset pricing model (8) or (9) says that prices should be proportional to second moments, this estimate is a natural way of choosing parameters to make the model hold as well as possible.

The GMM distribution theory (Hansen 1982) gives an asymptotic joint normal distribution for the estimates $\hat{\mathbf{b}}$. Denote by $\mathbf{S}$ a consistent estimate of the covariance matrix of the sample pricing errors $\mathbf{g}_T$, which is also the spectral density at zero of $m\mathbf{r} - \mathbf{p}$, or

$$\mathbf{S} = \text{consistent estimate of} \sum_{j=-\infty}^{\infty} E[(m_t\mathbf{r}_t - \mathbf{p}_t)(m_{t-j}\mathbf{r}_{t-j} - \mathbf{p}'_{t-j})].$$

Then Hansen (1982) shows that $\hat{\mathbf{b}}$ is asymptotically normal with variance-covariance matrix

$$\text{var}(\hat{\mathbf{b}}) = \frac{1}{T}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}. \tag{13}$$

If $\mathbf{W}$ is chosen equal to $\mathbf{S}^{-1}$, the GMM estimator is "optimal" or "efficient" in the sense that this variance matrix is as small as possible. In this case, the variance formula specializes to the more familiar form

$$\text{var}(\hat{\mathbf{b}}) = \frac{1}{T}(\mathbf{D}'\mathbf{S}^{-1}\mathbf{D})^{-1}. \tag{14}$$

It is interesting to know whether the pricing errors are in fact equal to zero, after one accounts for estimation and sampling error. Hansen (1982, lemma 4.1) also gives us a distribution theory for the pricing errors $\mathbf{g}_T$:

$$\text{var}(\mathbf{g}_T) = \frac{1}{T}[\mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}]\mathbf{S}[\mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}]'. \tag{15}$$

(While $\mathbf{S}$ is the variance-covariance matrix of the sample pricing errors evaluated at the true parameters, $\text{var}[\mathbf{g}_T(\mathbf{b})]$, the terms in brackets account for the fact that linear combinations of pricing errors are set to zero in parameter estimation, giving us $\text{var}[\mathbf{g}_T(\hat{\mathbf{b}})]$.) We can use this formula to construct standard errors for the pricing errors on individual assets or groups of pricing errors. In particular, we can test whether all pricing errors are zero by forming

$$\mathbf{g}'_T[\text{var}(\mathbf{g}_T)]^+\mathbf{g}_T \sim \chi^2(\#\text{moments} - \#\text{parameters}). \tag{16}$$

The "+" denotes pseudo-inversion,[6] since the variance-covariance matrix is singular with rank #moments − #parameters. When one uses the efficient estimator, $\mathbf{W} = \mathbf{S}^{-1}$, this test reduces to the celebrated J_T test of overidentifying restrictions:

$$TJ_T = T\mathbf{g}_T'\mathbf{S}^{-1}\mathbf{g}_T \sim \chi^2(\#\text{moments} - \#\text{parameters}). \qquad (17)$$

This is the basic test whether we can statistically reject a given observable factor model against a nonspecific alternative.

It can be more interesting to test a model against *specific* alternatives, that is, to ask, Given factors $\mathbf{f}_1$, are factors $\mathbf{f}_2$ important for pricing assets? There are two ways to perform such tests, corresponding to Wald and likelihood ratio philosophies. Start with a general model that includes both sets of factors, $m = \mathbf{b}_1'\mathbf{f}_1 + \mathbf{b}_2'\mathbf{f}_2$. First, we can use the sampling theory (13) or (14) to form t or χ^2 tests for $\mathbf{b}_2 = \mathbf{0}$. Second, we can compare the minimized objective $J_T = \mathbf{g}_T'\mathbf{S}^{-1}\mathbf{g}_T$ of a restricted system that excludes a given set of factors to the objective of the unrestricted system that includes all factors. If the excluded factors are not important for asset pricing, the J_T should not rise much. Precisely, if we use the same weighting matrix to estimate both systems (I use the weighting matrix from the unrestricted system), then

$$TJ_T(\text{restricted}) - TJ_T(\text{unrestricted}) \sim \chi^2(\#\text{ of restrictions})$$

(Newey and West 1987a). It is important *not* to simply compare J_T statistics from two estimates, but rather to use the same weighting matrix. A model can achieve a low J_T by simply blowing up the $\mathbf{S}$ matrix rather than improving the moment conditions.

To perform the GMM estimation, I start with an identity weighting matrix, $\mathbf{W} = \mathbf{I}$, which forms "first-stage" estimates of the parameters $\mathbf{b}$. I use these first-stage estimates to form an estimate of the matrix $\mathbf{S}$ and then use $\mathbf{S}^{-1}$ as the weighting matrix for "second-stage" estimates. I iterate this procedure, finding third-stage estimates and so forth. This does not change the asymptotic distribution theory, but Ferson and Foerster (1994) find that this procedure gives better small-sample performance. I also found that it produces results that are more stable across small variations in the model setup. Hansen, Heaton, and Luttmer (1995) suggest that one instead minimize $\mathbf{g}_T'(\mathbf{b})\mathbf{S}(\mathbf{b})^{-1}\mathbf{g}_T(\mathbf{b})$ directly as a function of parameters $\mathbf{b}$, and they find that this procedure can work well in small samples. This procedure

---

[6] One way to pseudo-invert a variance-covariance matrix is to perform an eigenvalue decomposition, $\mathbf{V} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$, with $\mathbf{\Lambda}$ a diagonal matrix of eigenvalues. Let $\mathbf{\Lambda}^+$ be a diagonal matrix with $1/\lambda_j$ for nonzero eigenvalues λ_j and zero for zero eigenvalues. Then $\mathbf{V}^+ = \mathbf{Q}\mathbf{\Lambda}^+\mathbf{Q}'$. In practice, it helps to multiply $\mathbf{V}$ by a large number initially to help distinguish small nonzero eigenvalues from rounding errors.

has the great advantage that it is invariant to normalization choices when **b** is not completely identified and to the choice of the initial weighting matrix. However, I have an analytic formula for $\hat{\mathbf{b}}$ only when the weighting matrix is held fixed, so the iterative strategy is much quicker.

When the factors are investment returns, I estimate production function parameters in addition to the factor weights **b**. Since these parameters enter nonlinearly, a search is required. The programming is harder, but the GMM methodology extends trivially.

B. Conditional Estimates and Conditional Factor Models

So far, I have considered *unconditional* factor models and estimates of unconditional moments. It is easy to include the effects of conditioning information by scaling the returns or the factors by instruments, as follows. (Scaling returns is Hansen and Singleton's [1982] use of instruments in a moment condition. Scaling factors is analogous to linear models of conditional betas, as in Ferson, Kandel, and Stambaugh [1987], Harvey [1989], and Shanken [1990].)

Scaling Returns

To test the conditional predictions of an asset pricing model,

$$\mathbf{p}_t = E(m_{t+1}\mathbf{r}_{t+1}|I_t), \tag{18}$$

we can expand the set of returns to include returns scaled by instruments and then proceed as before; that is, we use the moment conditions

$$E[\mathbf{p}_t \otimes \mathbf{z}_t] = E[m_{t+1}(\mathbf{r}_{t+1} \otimes \mathbf{z}_t)], \quad \mathbf{z}_t \in I_t, \tag{19}$$

where $\otimes$ denotes the Kronecker product (multiply every asset return by every instrument).

Equation (19) is an implication of equation (18): multiply both sides of (18) by z_t and take unconditional expectations. Conversely, if (19) holds for *all* variables z_t in an information set I_t, then (18) holds. Thus expanding the payoff space to include scaled returns as in (19) can test *all* the implications of (18), so that no generality is lost in principle. Of course, the usual instrument selection problem remains. If $z_t \in I_t$, then $z_t^2 \in I_t$; "every variable" in I_t means every variable and every measurable function of every variable, so in principle one has to include a lot of variables. In practice, one hopes to capture most of the predictability of $m\mathbf{r}$ with a few well-chosen and thoughtfully transformed instruments, as one hopes to capture the information in the thousands of available assets in a few well-chosen portfolios.

This scaling procedure also has an intuitive interpretation. The scaled returns $r_{t+1}z_t$ are the returns on managed portfolios in which the manager invests more or less according to the signal z_t. Thus we have shown that one can test all the *conditional* implications of an asset pricing model by performing *unconditional* tests on managed portfolio returns!

Scaling Factors

To test a model in which the factors are expected only to conditionally price assets, we can expand the set of factors to include factors scaled by instruments:

$$m_{t+1} = \mathbf{b}'(\mathbf{f}_{t+1} \otimes \mathbf{z}_t).$$

To motivate scaling factors, note that we have supposed so far that the discount factor m is a *fixed* linear combination of a given set of factors. However, the discount factor m might be a linear combination of factors with weights that vary as a vector of instruments z varies across different information sets:

$$m_{t+1} = \mathbf{b}(\mathbf{z}_t)' \mathbf{f}_{t+1}.$$

A conditional factor model does not imply an unconditional factor model: the model $0 = E_t[\mathbf{r}_{t+1}\mathbf{f}'_{t+1}\mathbf{b}(\mathbf{z}_t)]$ does not imply that there is a $\mathbf{b}$ such that $0 = E[\mathbf{r}_{t+1}\mathbf{f}'_{t+1}\mathbf{b}]$.

It is sufficient to consider $\mathbf{b}$'s that vary linearly with the instruments, since nonlinear functions can be expressed as linear functions of additional instruments. With one instrument z and one factor f then, the conditional factor model is

$$m_{t+1} = (b_0 + z_t b_1) f_{t+1}.$$

Scaling the factors f by the instruments z achieves the same result. The last equation is equivalent to

$$m_{t+1} = b_0 f_{t+1} + b_1 (f_{t+1} \times z_{t+1}).$$

Therefore, given the choice of instruments, performing the GMM estimation and testing with scaled factors is in principle a completely general test of a dynamic, conditional factor pricing model based on the instruments. Again, the only complaint one can make is that more or other instruments (or functions of instruments) should have been included.

Scaling

To keep scaling returns and scaling factors distinct, I refer to the former as "conditional estimates" and the latter as a "scaled factor

pricing model." The two are distinct: If one had a model that predicted constant b's, then it would be appropriate to scale the returns (perform conditional estimates) but not the factors. One can also examine the unconditional implications of a scaled factor pricing model, scaling the factors but not the returns.

Consumers may observe finer information sets, that is, more instruments z, than we do. This fact potentially reduces the power of tests that include scaled returns but does not bias them. Omitting instruments is exactly the same as omitting potential assets (managed portfolios). However, a conditional factor pricing model with respect to a fine information set does not imply a conditional factor pricing model with respect to a coarser information set, as it does not imply an unconditional factor model. Equivalently, conditional mean-variance efficiency does not imply unconditional mean-variance efficiency, though the converse is true (Hansen and Richard 1987). Thus a rejection of any factor model that is derived as a conditional factor pricing model with respect to consumers' information may still be attributed to an insufficiently rich set of instruments. But scaling factors does provide a very easy method for estimating and testing generally specified conditional factor pricing models given an information set.

C.   Relation to Traditional Statements and Tests of Factor Models

Factor Models

The statement that the discount factor m is a linear function of factors is equivalent to the conventional statements of factor pricing models in terms of betas and factor risk premia. (This fact has been known at least since Ross [1978] or Dybvig and Ingersoll [1982].) Precisely, the model

$$m = \mathbf{b}'\mathbf{f}; \quad 1 = E(mr) \tag{20}$$

implies the traditional statement of a factor pricing model,

$$E(r) = r^0 + \boldsymbol{\beta}'\boldsymbol{\lambda}, \tag{21}$$

where $\boldsymbol{\beta}$ is a vector of multiple regression coefficients of returns on the variable factors and r^0 is a constant across assets. Conversely, (21) implies that there exists a discount factor of the form $m = \mathbf{b}'\mathbf{f}$.

To prove this statement, define the riskless or zero beta rate

$$r^0 = \frac{1}{E(m)} = \frac{1}{E(\mathbf{f}')\mathbf{b}};$$

denote the variable factors $\tilde{\mathbf{f}}$, that is,

$$\mathbf{f} = \begin{bmatrix} 1 \\ \tilde{\mathbf{f}} \end{bmatrix};$$

define $\boldsymbol{\beta}$ as

$$\boldsymbol{\beta} \equiv \mathrm{cov}(\tilde{\mathbf{f}}, \tilde{\mathbf{f}}')^{-1} \mathrm{cov}(\tilde{\mathbf{f}}, r);$$

and define $\boldsymbol{\lambda}$ as the price of the demeaned variable factors, brought forward at the risk-free rate,

$$\boldsymbol{\lambda} = -r^0 E[m(\tilde{\mathbf{f}} - E\tilde{\mathbf{f}})].$$

With these definitions, one can simply manipulate either (20) or (21) to obtain the other. These definitions also allow one to obtain $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$ estimates from a GMM estimate of $\mathbf{b}$ together with the factor variance-covariance matrix or to obtain estimates and tests of $\mathbf{b}$ from cross-section or time-series regression estimates of $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$.[7]

One can rewrite an element λ_j of $\boldsymbol{\lambda}$ as

$$\lambda_j = -r^0[E(mf_j) - E(m)E(f_j)] = E(f_j) - r^0 E(mf_j).$$

If a factor f_j is an excess return, then $0 = E(mf_j)$, and we recover the familiar result that λ_j is the mean of the excess return factor $E(f_j)$. If a factor is a return, then $1 = E(mf_j)$, and we recover $\lambda_j = E(f_j) - r^0$.

The $\mathbf{b}$'s are *not* the same as the $\boldsymbol{\beta}$'s: $\mathbf{b}$ are the regression coefficients of m on $\mathbf{f}$, and $\boldsymbol{\beta}$ are the regression coefficients of r on $\mathbf{f}$. One tests whether "factor j is priced" by testing whether $\lambda_j = -r^0 E[m(\tilde{f}_j - E\tilde{f}_j)] = 0$. This hypothesis does *not* answer the question whether factor j is marginally useful in pricing other assets. To test whether factor j helps to price assets, one tests whether $b_j = 0$, that is, whether one can construct an m that prices the set of assets under examination without factor f_j. Since

$$\boldsymbol{\lambda} = -r^0 E[(\tilde{\mathbf{f}} - E\tilde{\mathbf{f}})\mathbf{f}'\mathbf{b}] = -r^0 \mathrm{cov}(\tilde{\mathbf{f}}, \tilde{\mathbf{f}}')\mathbf{b},$$

---

[7] Mechanically,

$$1 = E(mr) = E(r\mathbf{f}')\mathbf{b} = E(r)E(\mathbf{f}')\mathbf{b} + \mathrm{cov}(r, \mathbf{f}')\mathbf{b},$$

$$E(r) = \frac{1 - \mathrm{cov}(r, \mathbf{f}')\mathbf{b}}{E(\mathbf{f}')\mathbf{b}} = \frac{1 - \mathrm{cov}(r, \tilde{\mathbf{f}}')\tilde{\mathbf{b}}}{E(\mathbf{f}')\mathbf{b}}$$

$$= \frac{1 - \mathrm{cov}(r, \tilde{\mathbf{f}}')\mathrm{cov}(\tilde{\mathbf{f}}, \tilde{\mathbf{f}}')^{-1}\mathrm{cov}(\tilde{\mathbf{f}}, \tilde{\mathbf{f}}')\tilde{\mathbf{b}}}{E(\mathbf{f}')\mathbf{b}} = r^0 - r^0\boldsymbol{\beta}' \mathrm{cov}(\tilde{\mathbf{f}}, \mathbf{f}')\mathbf{b}$$

$$= r^0 - r^0\boldsymbol{\beta}' E[(\tilde{\mathbf{f}} - E\tilde{\mathbf{f}})\mathbf{f}'\mathbf{b}] = r^0 + \boldsymbol{\beta}'\boldsymbol{\lambda}.$$

The same steps backward prove "only if." Given either model, there is a model of the other form. They are not *unique*. We can add to m any random variable orthogonal to returns, and we can add risk factors with zero $\boldsymbol{\beta}$ or $\boldsymbol{\lambda}$, leaving pricing implications unchanged.

the hypotheses $b_j = 0$ and $\lambda_j = 0$ are equivalent only if the factors are orthogonal: if $\text{cov}(\tilde{\mathbf{f}}, \tilde{\mathbf{f}}')$ is diagonal.

The inclusion of scaled factors and scaled returns in an $m = \mathbf{b}'\mathbf{f}$, $\mathbf{p} = E(m\mathbf{r})$ model captures variation in conditional betas and factor risk premia $\boldsymbol{\lambda}$ in a very simple structure. Most tests of factor pricing models include auxiliary assumptions, such as constant conditional betas, constant conditional factor risk premia, constant conditional covariances, or complex time-series models for these quantities. Furthermore, the factors $\mathbf{f}$ do *not* have to be conditionally mean zero (white noise), conditionally or unconditionally orthogonal, or conditionally or unconditionally homoskedastic, as is often assumed.

The model $m = \mathbf{b}'\mathbf{f}$ is a factor *pricing* model. A factor *structure* on the covariance matrix of returns is sometimes used to derive factor pricing, but factor pricing does not imply or require a factor structure. The use of the same word "factor" for a pricing factor, a covariance factor structure, and a discount factor is unfortunate, but it is too late for me to try to change it.

Empirical Procedures

The moment conditions or pricing errors are proportional to expected return errors or α's. By the same argument as given above for the equivalence of $m = \mathbf{b}'\mathbf{f}$ and β, λ models, we have

$$\mathbf{g} = E(m\mathbf{r}) - \mathbf{p} = \frac{1}{r^0}[E(\mathbf{r}) - \boldsymbol{\beta}'\boldsymbol{\lambda} - r^0\mathbf{p}] = \frac{\boldsymbol{\alpha}}{r^0},$$

and conversely, we can recover expected return error α estimates from GMM estimates via

$$\boldsymbol{\alpha} = \frac{\mathbf{g}}{E(m)}. \tag{22}$$

The GMM objective is to minimize a weighted sum of squared pricing errors, which we can write in terms of α's as

$$\mathbf{g}_T(\mathbf{b})'\mathbf{W}\mathbf{g}_T(\mathbf{b}) = \frac{\boldsymbol{\alpha}_T(\mathbf{b})'\mathbf{W}\boldsymbol{\alpha}_T(\mathbf{b})}{(r^0)^2}.$$

Except for scaling by r^0, we can think of GMM as minimizing α's. The resulting χ^2 test has the form $\mathbf{g}_T' \times (\text{covariance matrix})^{-1} \times \mathbf{g}_T$, which is obviously analogous to the Gibbons, Ross, and Shanken (1989) test statistic, $\boldsymbol{\alpha}_T' \times (\text{covariance matrix})^{-1} \times \boldsymbol{\alpha}_T$.

Generalized method of moments minimizes the pricing errors in a way that is similar to a Fama-MacBeth (1973) cross-sectional regression. To see this, consider the special case in which the factors are

mean zero and only excess returns are considered, and ignore for the moment the distinction between sample and population moments. Then the model can be written as

$$0 = E(m\mathbf{r}^e); \quad m = 1 + \tilde{\mathbf{f}}'\mathbf{b}, \quad E(\tilde{\mathbf{f}}) = 0$$

($E(m)$ is not identified by excess returns, so I normalize to $E(m) = 1$). The first-order conditions for the GMM minimization are then

$$\text{cov}(\tilde{\mathbf{f}}, \mathbf{r}^{e'})\mathbf{W}[E(\mathbf{r}^e) - \text{cov}(\mathbf{r}^e, \tilde{\mathbf{f}}')\hat{\mathbf{b}}] = 0,$$

which we solve for

$$\hat{\mathbf{b}} = [\text{cov}(\tilde{\mathbf{f}}, \mathbf{r}^{e'})\mathbf{W}\,\text{cov}(\mathbf{r}^e, \tilde{\mathbf{f}}')]^{-1}\text{cov}(\mathbf{f}, \mathbf{r}^{e'})\mathbf{W}E(\mathbf{r}^e).$$

The first-stage GMM estimate ($\mathbf{W} = \mathbf{I}$) is thus a cross-sectional ordinary least squares (OLS) regression of expected returns on covariances. The second-stage GMM estimate is a corresponding GLS estimate.

In the more general case, the GMM estimate (eq. [12]) runs a cross-sectional regression of mean prices on second moments. This slight refinement allows the distribution theory to reflect the sampling error induced by estimating sample means of the factors.

V. Estimating and Testing the Investment Return Factor Model

A. Setup

I use two investment technologies, corresponding to gross private domestic nonresidential and residential investment. (The Appendix details the sources and transformations used for all data.) I assume that each investment series corresponds to a technology of the form (4), so that its investment returns are given by (5).

For asset returns, I use the 10 portfolios of NYSE stocks sorted by market value (size) maintained by the Center for Research in Security Prices (CRSP). There is a large spread in the mean returns of these portfolios: the small-firm decile's mean excess return is almost twice that of the large-firm decile. Any asset pricing model must explain this spread in mean returns by a spread in assets' covariance with risk factors. Since the investment returns are based on quarterly average investment, I transformed the asset returns to quarterly average returns rather than use end-of-quarter to end-of-quarter returns. I include moment conditions for investment returns along with the moment conditions generated by asset returns, since both sets of returns should be correctly priced by this or any model and since it is interesting and important to check whether the asset pricing model can ac-

count in this way for macroeconomic events. I created excess returns by subtracting the 3-month Treasury-bill return in each case, to focus on risk premia. I also include the level of the ex post real Treasury-bill rate as an asset, in order to identify the level of the discount factor m.[8]

I use two instruments: the term premium (yield on long-term government bonds less yield on 3-month Treasury bills) and the dividend/price ratio of the equally weighted NYSE portfolio. I also considered the default premium (yield on BAA corporate bonds minus yield on AAA corporate bonds); with the dividend/price ratio, it produces similar results. These instruments are popular forecasters of stock returns. In the first-stage estimation, the moments corresponding to scaled returns are treated equally with the nonscaled returns, so it is convenient that the scale of the two is roughly comparable. To this end, I used $1 + 100 \times [(d/p) - 0.04]$ in place of the raw dividend/price ratio.[9] To avoid overlap with the averaged return series, I lag the instruments twice: an instrument used for the return from the first to second quarter is known by the last day of December.

Scaling factors and assets by instruments can lead to an explosion of moment conditions and scaled factors. I prune this explosion in three ways, beyond the already limited set of assets and instruments. First, I do not scale the Treasury-bill return by the instruments. Such scaling asks whether the model can capture variation over time in the real Treasury-bill return. I want to focus on whether the model can capture the much larger variation over time in risk premia, represented by scaled excess stock returns. Second, I scale the variable factors by the instruments, but I do not include the instruments themselves (constant scaled by instruments) as factors. Such factors help the model to capture variation in risk-free rates (varying $E_t(m)$), but they do not help the model capture time-varying risk premia. Third, I use only deciles 1, 2, 5, and 10 in the conditional estimates (return times instrument). With all 10 deciles I would have 37 moment conditions in 186 data points. The iterated GMM estimates behaved badly with 37×37 covariance matrices. I hope that deciles 2, 5, and 10 capture most of the cross-sectional information in (span the frontier of) the original 10 deciles; I include decile 1 as well in order to examine the well-known size anomaly.

---

[8] Of course, the system $1 = E(mr^{tb})$; $0 = E[m(\mathbf{r} - r^{tb})]$ is equivalent to the system $1 = E(mr^{tb})$; $1 = E[m\mathbf{r}]$, since either set of moments is a linear combination of the other. In a one-step efficient GMM estimate, min $\mathbf{g}_T(b)'\mathbf{S}^{-1}(b)\mathbf{g}_T(b)$, the two setups would yield exactly the same result. However, iterated and first-stage GMM estimates are affected by the initial choice of moments.

[9] This transformation is supposed to be data-independent. If I were to use $[(d/p) - E(d/p)]/\sigma(d/p)$, I would have to adjust the distribution theory for estimation of the mean and variance of d/p, not a straightforward task. The estimates *are* sensitive to scaling choices of the instruments, so it is important to choose instruments with a reasonable scale.

If we allow all the production function parameters $\{\eta, \delta, mpk\}$ to vary, the system is overparameterized. Examining the definition of the investment return (5), one can see that the parameters δ and mpk basically affect the mean of the investment return, and η affects its mean and standard deviation. None of the parameters substantially affects the cross-correlations of investment returns with other variables; they are basically given by the cross-correlation of investment growth with the other variables. Furthermore, the mean and standard deviation of the factors are not separately identified, since the factors can be rescaled at will by choice of b. As a result, the minimization surface has a valley in it, and the minimization program soon crashes with a singular gradient matrix $\partial g_T / \partial [\eta\ \delta\ mpk]$. Therefore, I present results in which η and δ are held fixed, minimizing only over mpk. I tried choosing each of the parameters and choosing the parameters sequentially (first η, then mpk, etc.); both procedures produce similar results.

The tables below present results imposing no autocorrelation in the construction of the S matrix, since the null hypothesis of the central conditional models predicts that lagged mr should not predict mr. I also tried using four Newey-West (1987b) lags to construct standard errors, and the results overall are not much changed, indicating little autocorrelation.

B.   Iterated GMM Estimates and Tests of the Investment Return Model

Table 1 presents estimates and tests of the investment return factor model. Start with the simple unconditional estimates of the nonscaled factor model, panel A. The marginal product of capital parameters mpk are plausible and highly significant. They have about the same value (0.05–0.06) and are highly significant in all the following estimates, so I do not present them in the following tables, though they are estimated any time an investment return is estimated. The b's measure which factors are important for pricing assets. The residential factor prices significantly ($t = -2.61$), but the nonresidential factor does not ($t = 1.37$). They are jointly significant (p-value for joint $b_{nr}, b_r = 0$ is 3 percent). In interpreting the b's, keep in mind that the discount factor m is proportional to the minimum *second-moment* return, which is on the lower portion of the minimum variance frontier. Since the investment returns are typically on the upper portion of the minimum variance frontier, b's may be negative. Finally, the J_T test of overidentifying restrictions does not reject the model (p-value 24 percent).

In the conditional estimates (with scaled returns), panel B, both investment return factors are individually significant (t on $b = 1.94$

TABLE 1

Iterated GMM Estimates and Tests of Investment Return Factor Model

A. Nonscaled Model $m = b_0 + b_{nr} r_{nr}^i + b_r r_r^i$

	mpk_{nr}	mpk_r	b_0	b_{nr}	b_r
			PARAMETER ESTIMATES		
	Unconditional Estimates ($1 = E(mr^{tb})$, $0 = E(m\mathbf{r}^e)$ [13 moments])				
Coefficient	.058	.063	-19	71	-51
t-statistic	38	44	$-.40$	1.37	-2.61
	Conditional Estimates ($1 = E(mr^{tb})$, $0 = E(m\mathbf{r}^e \otimes \mathbf{z})$ [19 moments])				
Coefficient			-36	123	-86
t-statistic			$-.56$	1.94	-5.57

	All b	b_{nr}, b_r	J_T	Stock J_T
		TESTS		
	Unconditional Estimates			
χ^2	382	7.0	10.4	
Degrees of freedom	3	2	8	
p-value (%)	.00	3.1	24	
	Conditional Estimates			
χ^2		35	22	22
Degrees of freedom		2	14	12
p-value (%)		.00	7.4	4.0

B. Scaled Factor Model $m = b_0 + \mathbf{b}'(\mathbf{r}^i \otimes \mathbf{z})$: Conditional Estimates ($1 = E(mr^{tb})$, $0 = E(m\mathbf{r}^e \otimes \mathbf{z})$)

	b_0	b_{nr}	b_r	$b_{nr \cdot tp}$	$b_{r \cdot tp}$	$b_{nr \cdot dp}$	$b_{r \cdot dp}$
			PARAMETER ESTIMATES				
Coefficient	-29	157	-126	48	-48	-57	56
t-statistic	$-.46$	1.48	-2.00	3.42	-3.42	-2.12	2.11

	No b_0	Unscaled	Scaled	All nr	All r	J_T
			TESTS			
			Joint $b = 0$			
Wald χ^2	32.8	4.1	18.4	12.2	13.6	
p-value (%)	.001	13	.1	.7	.4	
$\Delta J_T \chi^2$	54	4.2	26	11.1	10.9	10.6
p-value (%)	.00	12	.003	1.1	1.2	39
Degrees of freedom	6	2	4	3	3	10

NOTE.—In the unconditional estimates, $\mathbf{r}^e$ is the 10 CRSP size decile portfolio and two investment excess returns, and r^{tb} is the real Treasury-bill return. The conditional estimates use the deciles 1, 2, 5, 10, and investment excess returns, scaled by instruments, and the real Treasury-bill return. Investment returns are functions of nonresidential (nr) and residential (r) gross fixed investment, eq. (5) with parameters $\eta = 3.0$, $\delta = .05$ throughout. mpk_{nr} and mpk_r are always estimated, even when not shown. Instruments are the constant, term premium (tp), and equally weighted dividend/price ratio (dp). The p-value is the percentage probability of obtaining a χ^2 value as high or higher.

and -5.57) and jointly highly significant (p-value < 0.00 percent). Adding moments should sharpen the precision of estimates, and it does. However, the J_T statistic now provides some evidence against the model (p-value 7 percent). A J_T test formed using only the stock returns, explained in some detail below, produces slightly more evidence against the model, with a 4 percent p-value.

The natural response to the rejection of the conditional estimates is to include scaled factors, panel B. The scaled factors are individually and jointly significant. The table presents Wald tests, based on the joint distribution of the b in equation (14), as well as tests on the increase in J_T when groups of the factors are omitted from the model. It is comforting that the two procedures yield quite similar results. The unscaled factors are only marginally jointly significant (p-value 13 percent), but the scaled factors are highly jointly significant. Also, the residential and nonresidential factors are significant as subgroups. Finally, the J_T test does not reject the overall model (p-value 39 percent).

C. First-Stage Estimates and Tests

Every first-year econometrics student is advised that GLS is best but OLS is pretty good. The GLS estimates can be more efficient with a good estimate of the covariance matrix, but GLS estimates can be terrible if the covariance matrix is poorly modeled. The OLS estimates are consistent and robust to many misspecifications. Thus one is often advised to make sure that GLS estimates are not too different from OLS estimates, or even to estimate parameters by the inefficient but robust OLS, correcting standard errors for residual correlation or heteroskedasticity.

The same advice applies to GMM. Efficient GMM estimates use the estimated covariance matrix of the sample moments to find linear combinations of those moments that are the most precisely measured. Generalized method of moments weights those linear combinations more highly in estimation, in order to improve efficiency, and then evaluates the model by testing whether those most precisely estimated linear combinations of moments are in fact zero. The dangers of this procedure if the S matrix is poorly measured are the same as those of GLS. The estimate may pay too much attention to portfolios that spuriously seem nearly risk-free in a small sample and hence seem to have well-measured pricing errors. Statistical issues aside, efficient GMM may pay close attention to economically uninteresting but statistically well-measured moments.

To make this observation precise, diagonalize the S matrix, $S = Q\Lambda Q'$, where Q is an orthonormal matrix with eigenvectors in its

columns and Λ is a diagonal matrix of eigenvalues. Since $\mathbf{S}^{-1} = (\mathbf{Q}'\Lambda\mathbf{Q})^{-1} = \mathbf{Q}'\Lambda^{-1}\mathbf{Q}$, we can define

$$\mathbf{r}^* = \mathbf{Q}'\mathbf{r}; \quad \mathbf{p}^* = \mathbf{Q}'\mathbf{p}$$

and write the efficient GMM objective as

$$\min\{[E(m\mathbf{r}^* - \mathbf{p}^*)]'\Lambda^{-1}[E(m\mathbf{r}^* - \mathbf{p}^*)]\}.$$

The efficient GMM estimate first forms portfolios of the original assets with weights given by the eigenvalues of $\mathbf{S}$; then it pays more attention to the portfolios corresponding to small eigenvalues of $\mathbf{S}$— the ones whose pricing errors are most precisely measured.

The smallest four eigenvalues of $\mathbf{S}$ in this case are 1/64,000, 1/22,600, 1/14,200, and 1/7,800, so the first eigenvector is by far the most important in the estimation and testing. Figure 1 presents the corresponding portfolio weights: the first column of $\mathbf{Q}$. (Figure 1 rescales the weights so that they sum to one.) By far, the most important assets in this portfolio are the two investment returns. The investment returns have a good deal less variance than the stock returns, and so their means are more precisely measured. The portfolio places practically no weight on the Treasury-bill return. The asset and investment return moments are formed from m times excess returns, or m times a number typically around 0.02 (investment) or 0.10 (asset). The Treasury-bill rate moment is formed from m times a
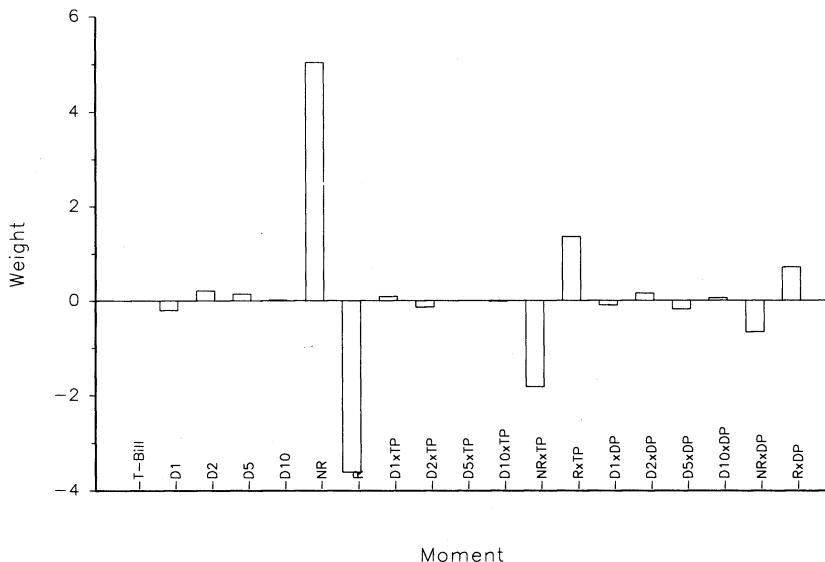

FIG. 1.—Portfolio (eigenvector) corresponding to largest eigenvalue of $\mathbf{S}^{-1}$

number typically around 1.01. Therefore, the quite high volatility of the discount factor m translates into a high variance of mr^{tb} and thus a high sampling variability for this moment.

The portfolio graphed in figure 1 also features large long and short positions in similar assets. This is a common feature of portfolios formed to minimize variance in a sample. The original portfolios are highly correlated, so their sample variance-covariance matrix and hence the S matrix are nearly singular. Small sampling errors in means and covariances make it look like there are nearly riskless portfolios.

So, GMM does exactly what we ask it to. We ask it to measure parameters "efficiently," which means to weight more heavily sources of information with less estimated sampling variation. Then GMM evaluates the model (via the J_T statistic) by asking whether the pricing errors on these low sampling error portfolios are in fact zero, after accounting for sampling information.

But is this what we want GMM to do? Perhaps not.

Evaluation: J_T Tests on a Restricted Set of Moments

Perhaps we do not want to accept or reject the model on the basis of how well it prices the portfolio graphed in figure 1. One alternative is to use only a subset of assets in the overidentifying restrictions test. Using the expression (15) for the variance-covariance matrix of the moment conditions, we can form an analogue to the J_T test in equations (16) and (17) using only an interesting subset of moments.

I conduct such tests for all moments excluding the Treasury-bill rate, and for the stock returns alone, to check that the model is not evaluated only on its ability to price investment returns. In all but one case, these tests produce numerically the same values as the J_T test in table 1. To see how this is possible, recall that the GMM estimate sets some linear combinations of moments to zero in sample in order to estimate parameters. Suppose that the first moment of a two-moment GMM estimate is set to zero in estimation, that is,

$$\left[\left(\frac{\partial \mathbf{g}_T}{\partial \mathbf{b}}\right)' \mathbf{W}\right]\mathbf{g}_T = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} g_{1T} \\ g_{2T} \end{bmatrix} = 0.$$

Now, the J_T test is based only on the second moment. If we test only the second moment, we obtain exactly the same test statistic and degrees of freedom as the J_T test. In the one case in table 1 in which the stock return J_T test was different from the J_T test, marked "Stock J_T," the stock return only test gives slightly more statistical evidence against the model, lowering the p-value from 7 percent to 4 percent.

Thus the iterated GMM estimates and tests survive an important robustness check: *evaluating* this model on only the asset moments does not make much difference, even though the *estimates* concentrate on the investment return moments.

Estimation: First-Stage GMM

If we are uncomfortable estimating and evaluating a model on the basis of portfolios such as figure 1, perhaps we should instead ask GMM to weight assets more evenly or, in an economically more interesting way, in estimation as well as in testing. I use an identity weighting matrix, so that all moments are weighted equally. This is the first-stage GMM estimate and test.[10] The only tricky part of first-stage or fixed weighting matrix estimates is that one must use the corresponding standard error, moment variance, and overidentifying restrictions test formulas, equations (13) and (16), rather than their more familiar special cases for an optimal weighting matrix, equations (14) and (17). In addition, the test based on the increase in GMM objective no longer has a χ^2 distribution, so I examine only Wald tests of joint b's for model comparisons.

Table 2 presents *first-stage* estimates and tests of the investment return factor model. The estimates and test statistics are quite similar to their iterated GMM counterparts in table 1, with the decrease in precision that one expects of a less efficient estimate: the significance levels of the b's and χ^2 tests are lower. The one qualitative difference is that the evidence against the nonscaled factor model is much weaker; I return to this point below. Overall, the iterated GMM esti-

---

[10] Hansen and Jagannathan (1991*a*) advocate an alternative prespecified weighting matrix, the second-moment matrix of returns,

$$\min\{E(m\mathbf{r} - \mathbf{p})' E(\mathbf{rr'})^{-1} E(m\mathbf{r} - \mathbf{p})\}.$$

This weighting matrix, like the one-step efficient GMM outlined in n. 8, has an important advantage over the identity matrix: it produces an estimate that is invariant to units and portfolio formation. In particular, it is invariant to the choice of units of the instruments; using $2z$ in place of z makes no difference, where it would double the attention the identity weighting matrix pays to returns scaled by z ratios. Because of this sensitivity, I had to carefully choose the units of the instruments. On the other hand, the return second-moment matrix is very nearly singular: $E(\mathbf{rr'}) = E(\mathbf{r})E(\mathbf{r'}) + \text{cov}(\mathbf{r}, \mathbf{r'})$, so one takes an already near-singular covariance matrix and adds a singular matrix, typically one with much larger elements. Therefore, its eigenvectors feature much stronger long and short positions, and no more economic justification, than the eigenvectors of the **S** matrix as graphed in fig. 1. Estimates with this weighting matrix can produce large pricing errors on the original assets. Also, I use the root mean square error (RMSE) expected return errors or α's and plots of α's as a diagnostic. Using the identity weighting matrix, GMM picks parameters to do best by this diagnostic; we avoid the confusion that a model might fit better by the estimate but look worse in the RMSE α diagnostic.

TABLE 2

FIRST-STAGE GMM ESTIMATES AND TESTS OF INVESTMENT RETURN FACTOR MODEL

A. NONSCALED MODEL

	PARAMETER ESTIMATES		
	b_0	b_{nr}	b_r
	Unconditional Estimates		
Coefficient	−9.5	99	−88
t-statistic	−.11	1.10	−3.27
	Conditional Estimates		
Coefficient	−293	408	−109
t-statistic	−1.70	2.23	−2.93
	TESTS		
	b_{nr}, b_r	J_T	Stock J_T
	Unconditional Estimates		
χ^2	1.2	8.4	
Degrees of freedom	2	8	
p-value (%)	27	40	
	Conditional Estimates		
χ^2	10.7	12.4	11.3
Degrees of freedom	2	14	12
p-value (%)	.47	57	51

B. SCALED FACTOR MODEL: CONDITIONAL ESTIMATES

	PARAMETER ESTIMATES						
	b_0	b_{nr}	b_r	$b_{nr \cdot tp}$	$b_{r \cdot tp}$	$b_{nr \cdot dp}$	$b_{r \cdot dp}$
Coefficient	−38	131	−90	50	−50	−44	44
t-statistic	−.30	.74	−1.04	2.03	−2.03	−1.07	1.06

	TESTS					
	Joint $b = 0$					
	No b_0	Unscaled	Scaled	All nr	All r	J_T
χ^2	17.4	1.1	9.3	5.3	5.4	12.3
Degrees of freedom	6	2	4	3	3	10
p-value (%)	.8	58	5.5	15	15	27

NOTE.—The same estimates and tests as given in table 1, except first-stage GMM (identity weighting matrix) rather than iterated or optimal GMM estimates and tests.

mates and tests pass the robustness check that they are not too different from first-stage estimates and tests.

D. Pricing Errors

Expected return pricing errors or α's are a useful characterization of a model's performance. Examining them helps to guard against accepting an uninteresting model: one that prices assets badly but produces large enough standard errors not to be rejected by the J_T statistic. It also helps to guard against the equally dangerous possibility of rejecting a good model: one that produces economically tiny pricing errors, but such small standard errors that the model is still statistically rejected.

Figure 2 presents the predicted versus actual mean excess returns for the nonscaled model (fig. 2a), the conditional estimates of the nonscaled model (fig. 2b), and the scaled model (fig. 2c). These values are calculated from equation (22). The straight line in each panel is the 45° line, along which all the assets should lie. Starting from the lower left of figure 2a, we have the Treasury-bill rate and two investment excess returns. The placement of the investment returns is not an essential feature of the model. It is easy to produce investment returns that lie farther apart or at different places along the line in figure 2, yet price about as well, by different choices of the fixed parameters η and δ. The group of assets up and to the right are the decile portfolios, with the smaller-size firm deciles farther out on the graph. In figure 2b and c, there are three triangles for each asset, corresponding to the asset and the asset scaled by each of the two instruments.

Figure 2 presents α's calculated from the *first*-stage estimates. The first-stage GMM estimate basically minimizes RMSE α's, so if we compare models by α plots or RMSE α's, we know that each model will fit as well as possible along this dimension. For comparison, figure 3 presents the *iterated* GMM predicted versus actual mean excess returns. The Treasury-bill pricing error is dramatic: its excess return is predicted at -4.6 percent, but the actual value is, of course, zero. This occurs because, as discussed above regarding figure 1, the Treasury-bill moment condition is very imprecisely measured; iterated GMM therefore pays little attention to the Treasury-bill return in order to better price the other assets. The iterated GMM estimates do a reasonably good job of pricing the remaining portfolios. Still, the spread is larger than in the first-stage estimates. The iterated GMM estimate would of course produce smaller pricing errors for the **S** eigenvalue portfolios such as shown in figure 1, whose pricing errors it minimizes.
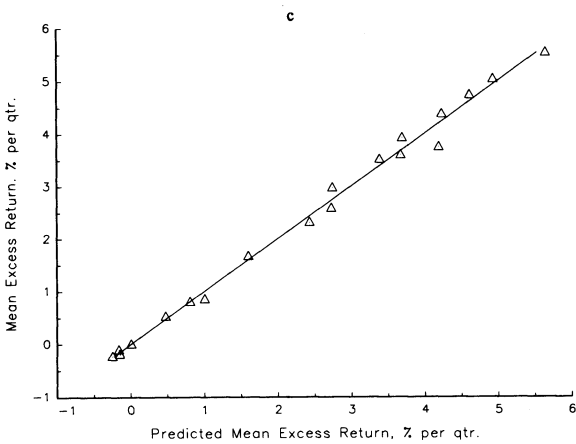
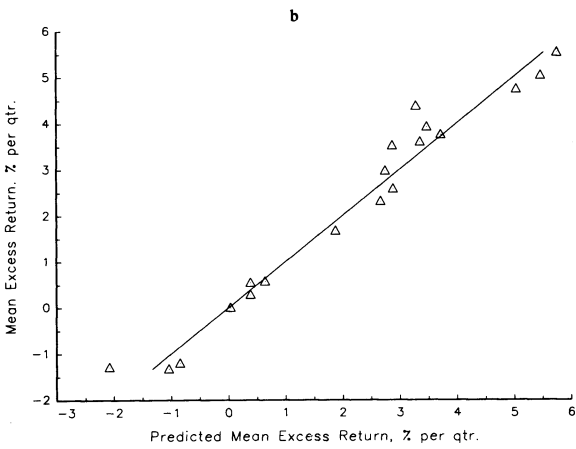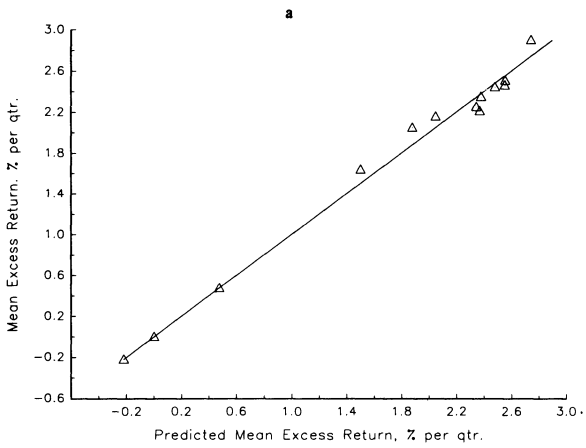FIG. 2.—Predicted vs. actual mean excess returns, investment return model: *a*, non-scaled model; *b*, nonscaled model, conditional estimates; *c*, scaled model. First-stage GMM estimates.
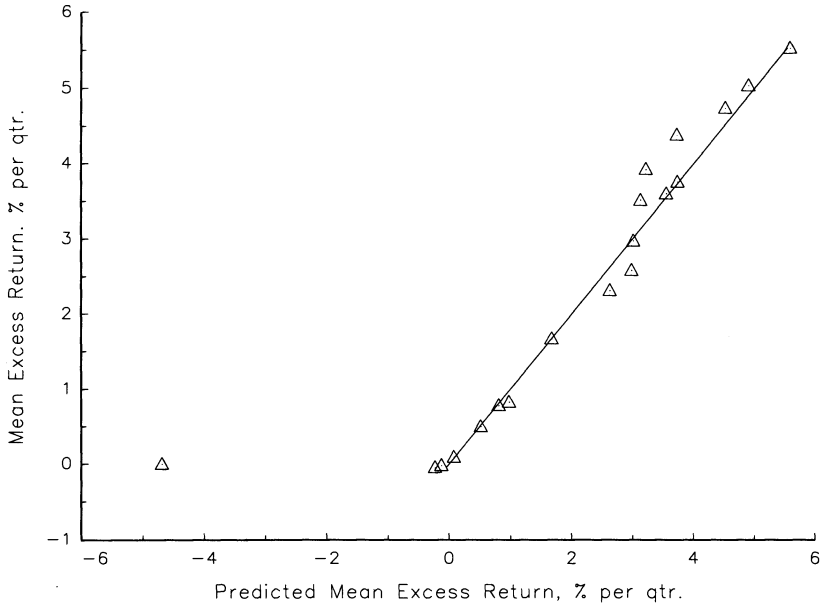
Fig. 3.—Predicted vs. actual mean excess returns, scaled investment return factor pricing model, iterated GMM estimate.

It is a good check that the iterated GMM pricing errors are not that different from the first-stage pricing errors, at least for the interesting assets. If the iterated GMM estimates produced wild pricing errors for the original assets in order to minimize the pricing errors of the S eigenvalue portfolios, we might suspect that the GMM estimates are not that reliable or that something is wrong with the model. Hence, the iterated GMM estimates pass another important robustness check.

The difference between figure 2 and figure 3 dramatizes a point made by Kandel and Stambaugh (1995): pricing error graphs are not robust to portfolio reformation. As long as the pricing errors are not zero, one can find a repackaging of portfolios to make graphs like figures 2 and 3 look arbitrarily good or bad. Therefore, it is important to examine a model's ability to explain the expected returns of economically interesting portfolios.

The main difference between first-stage estimates of table 2 and iterated estimates of table 1 is that the scaled factors seem only marginally statistically useful in the first-stage estimates. The conditional estimate of the nonscaled model fails to reject with p-values around 50 percent instead of the 4–7 percent p-values from the iterated estimate, and the scaled factors have only a 5.5 percent p-value in the scaled factor model.

However, the pricing errors or α's presented in figure 2 suggest that scaled factors are indeed important, since the pricing errors of the scaled factor model (fig. 2c) are decidedly smaller. Table 3 collects RMSE α's for a variety of models and tells the same story: the scaled model achieves an RMSE α of 0.15 percent, whereas the nonscaled model achieves only 0.42 percent, more than twice as much.

Part of the story for the decline in pricing errors with a scaled model is degrees of freedom: one always lowers the objective by adding more factors. The other part of the story is the danger of accepting models with large pricing errors but even larger standard errors. The sampling variance of the moments estimated in the nonscaled factor model is a good deal larger than for the scaled factor model. Thus larger pricing errors are less statistically significant. One way to see this point is in the fact that the nonscaled factor model is not rejected with a 57 percent p-value, yet we reject elimination of the scaled factors from the scaled factor model—using the *scaled* factor model variance-covariance matrix in the test—with a 5.5 percent p-value.

All the sample pricing errors are far from individually statistically significant (I use the diagonal elements of eq. [15] to calculate individual moment standard errors). The J_T test is based on a weighted sum of squares of the pricing errors plotted in figure 2. When it does not reject (tables 1 and 2), the pricing errors are jointly insignificant as well.

VI. Comparison with Other Models

The overidentifying restrictions (J_T) test the investment return model against no specific alternative. But *all* currently available nontrivial models can undoubtedly be statistically rejected if one uses a sufficiently rich set of assets and instruments and a long enough sample. Therefore, it may be more interesting to compare a given model to plausible competitors rather than simply reject or fail to reject it.

In this section, I compare the investment return model to the capital asset pricing model (CAPM), the Chen, Roll, and Ross (1986) factor model, the consumption-based model, and two ad hoc macro factor models. In each case, I estimate, test, and examine the competing model, in the style of tables 1 and 2 and figure 2. Then I estimate models that include *both* investment return and the other factors, to see which set of factors can be deleted in the presence of the other. The RMSE pricing errors and pricing error graphs like figure 2 provide an economic counterpart to the statistical comparison.

TABLE 3

EXPECTED RETURN PRICING ERRORS OR α's

ASSETS		MODEL								
	r^i	CAPM	r^i + CAPM	CRR	r^i + CRR	$\Delta c^{-\gamma}$	$r^i + \Delta c^{-\gamma}$	Δc	$r^i + \Delta c$	ΔI
				Nonscaled Models, Unconditional Estimates						
Stocks, r^i, r^{tb}	.099		.069		.028		.094		.098	
Stocks	.113	.094	.079	.037	.031	.054	.112	.301	.112	.096
				Nonscaled Models, Conditional Estimates						
Stocks, r^i, r^{tb}	.42		.41		.21		.30		.32	
Stocks	.46	.94	.45	.33	.24	2.86	.30	1.81	.33	.51
				Scaled Models (CRR and $\Delta c^{-\gamma}$), Conditional Estimates						
Stocks, r^i, r^{tb}	.15		.12		.11		.14		.14	
Stocks	.19	.49	.15	.33	.13	2.86	.16	.59	.17	.18

NOTE.—Entries are root mean square α's—mean return less predicted mean return—expressed in percentage per quarter. Entries are calculated as $\alpha_z = 100 \times E(mr_j - p_j)/E(m)$. They are based on first-stage GMM estimates. Unconditional estimates use deciles 1–10 and investment returns and real Treasury-bill return where indicated. Conditional estimates use deciles 1, 2, 5, and 10 scaled by the constant, term premium, and dividend/price ratio, plus the Treasury-bill rate, investment returns, and scaled investment returns where indicated.

A. CAPM

Estimates and Tests of the CAPM

The CAPM is a single-factor model with the market return r^m as the factor, $m = b_0 + b_m r^m$.[11] Thus it trivially maps into the factor pricing–GMM framework outlined above.

   Table 4 presents GMM estimates and tests of the CAPM. This estimation does not include any investment returns. The pattern of results is similar to that of the investment return factor model. In the unconditional estimates of the nonscaled model, the market return is a significant factor, with t-statistics of -3.2 (first-stage) and -3.5 (iterated). The J_T test does not reject with a 95 percent p-value. In the conditional estimates, the market return prices even more significantly, but the J_T test soundly rejects the model with a 0.7–1.6 percent p-value. However, in many derivations, the CAPM is a one-period or conditional model, so we should include scaled factors. When we do so (panel B), the b's are individually and jointly significant. In contrast to the investment model, however, the overidentifying restrictions of the scaled model are now rejected at a 2.6 percent p-value in the iterated GMM estimates, and nearly rejected at a 7.7 percent p-value even at the first stage. Overall, the first-stage and iterated estimates are similar in this table, with the iterated estimates giving slightly stronger statistical results.

Comparison Tests

Do the investment returns drive out the market return or vice versa? In a factor model that includes *both* the market and the investment returns, which factors are significant for pricing assets?

   Table 5 collects such comparison tests. There are three tests: Wald tests for joint $b = 0$ conducted with the first- and second-stage estimates and χ^2 difference tests conducted by eliminating each set of factors in turn during the iterated GMM estimates. I compare scaled and nonscaled models separately. All the comparison tests are based on conditional estimates, including scaled returns.

   The estimates in this table *do* include investment returns, for two reasons. First, including investment return moments is important for estimating the production function parameters mpk_{nr}, mpk_r. Second, we are interested in finding models that not only price financial assets but relate asset prices to events in the macroeconomy. To guard

---

[11] The CAPM can also be specified with the excess market return as the factor, $m = b_0 + b_1(r^m - r^f)$, or with the market return and risk-free rate or zero beta rate as two factors, $m = b_0 r^f + b_1 r^m$.

TABLE 4

GMM Estimates and Tests of CAPM

A. Nonscaled Model $m = b_0 + b_m r^m$

Parameter Estimates

	Unconditional Estimates		Conditional Estimates	
	b_0	b_m	b_0	b_m
First-stage:				
Coefficient	6.5	−5.4	9.5	−8.4
t-statistic	3.74	−3.21	5.53	−5.05
Iterated:				
Coefficient	6.7	−5.6	9.8	−8.6
t-statistic	4.08	−3.53	5.94	−5.42

Tests

	Unconditional Estimates	Conditional Estimates
	J_T	J_T
First-stage:		
χ^2	3.3	26
Degrees of freedom	9	11
p-value (%)	95	.71
Iterated:		
χ^2	3.3	23
Degrees of freedom	9	11
p-value (%)	95	1.55

B. Scaled Model $m = b_0 + b_m r^m + b_{tp}(r^m \times tp) + b_{dp}(r^m \times dp)$: Conditional Estimates

Parameter Estimates

	b_0	b_m	b_{tp}	b_{dp}
First-stage:				
Coefficient	4.56	−2.66	−.33	−.39
t-statistic	1.48	−.80	−1.32	−2.05
Iterated:				
Coefficient	5.88	−4.62	.24	−.36
t-statistic	3.51	−2.70	2.26	−3.62

Tests

	b_m, b_{tp}, b_{dp}	Scaled b	J_T
First-stage:			
χ^2	59	4.9	15.6
Degrees of freedom	3	2	9
p-value (%)	.00	8.6	7.7
Iterated:			
χ^2	67	15	18.9
Degrees of freedom	3	2	9
p-value (%)	.00	.06	2.6

Note.—r^m is the value-weighted NYSE return. No investment returns are included.

TABLE 5

MODEL COMPARISON TESTS

A. NONSCALED INVESTMENT RETURN MODEL

	ALTERNATIVE MODEL							
	Static CAPM		CRR		$\Delta c^{-\gamma}$		Δc	
	r^i	r^m	r^i	crr	r^i	c	r^i	c
Omitted factors	r^i	r^m	r^i	crr	r^i	c	r^i	c
Degrees of freedom	2	1	2	5	2	1	2	1
First-stage χ^2	5.0	.09	3.8	2.3	11	.01	8.3	2.5
p-value (%)	8.2	76	15	81	.4	91	1.6	12
Iterated χ^2	13.5	.68	2.0	5.4	26	.01	32	29
p-value (%)	.1	41	37	37	.00	94	.00	.00
$\Delta J_T \chi^2$	11.5	.68	1.7	5.4	27	.15	28	20
p-value (%)	.3	41	43	37	.00	70	.00	.00

B. SCALED INVESTMENT RETURN MODEL

	ALTERNATIVE MODEL							
	Scaled CAPM		CRR		$\Delta c^{-\gamma}$		Scaled Δc	
	r^i	r^m	r^i	crr	r^i	c	r^i	c
Omitted factors	r^i	r^m	r^i	crr	r^i	c	r^i	c
Degrees of freedom	6	3	6	5	6	1	6	3
First-stage χ^2	8.8	.96	6.4	1.9	24	.00	9.8	.44
p-value (%)	19	81	38	85	.05	95	13	93
Iterated χ^2	14.2	7.6	5.9	1.9	29	.06	13	2.4
p-value (%)	2.8	5.6	43	87	.01	80	4.4	50
$\Delta J_T \chi^2$	13	9.5	17	1.9	29	2.6	12	3.1
p-value (%)	4.5	.4	1.1	87	.01	11	5.3	37

NOTE.—Tests for joint $b_j = 0$ in models $m_0 = b_0 + b_1' r^i + b_2' f$, where r^i denotes investment returns and f denotes additional factors listed in each column. All tests are based on conditional moments: deciles 1, 2, 5, and 10, the Treasury-bill return, and investment returns, scaled by the constant, term premium, and dividend/price ratio. First-stage and iterated give Wald tests based on the indicated GMM estimate; ΔJ_T gives rise in the GMM objective when one set of factors is excluded.

against the danger that the comparison tests are driven by the investment returns, I include first-stage estimates and I compare stock return pricing errors.

The nonscaled model tests favor the investment return model. In the first stage, there is an 8.2 percent p-value for dropping the investment return factors, compared to 76 percent for dropping the market return. Iterated estimates and the ΔJ_T test raise this to 0.1–0.3 percent p-values for dropping the investment return factors, against 41 percent for dropping the market. Here and below, the similarity of Wald and ΔJ_T tests is comforting.

When we compare the scaled investment return model to the scaled CAPM, the table suggests that each set of factors is statistically important in the presence of the others. In the first stage, we cannot reject

eliminating either set of factors, though the investment factors have somewhat stronger evidence (p-value 19 percent) than the market factors (p-value 81 percent). The iterated GMM Wald tests reject exclusion of the investment return factors (p-value 2.8 percent) but borderline reject exclusion of the scaled market factors (p-value 5.6 percent); but the ΔJ_T tests neatly reverse the pattern of these p-values, and small differences in p-values based on asymptotic distributions are a dangerous decision criterion.

Pricing Errors

Figure 4 plots the pricing errors or α's for the CAPM. In the unconditional estimates of the nonscaled model (fig. 4a), we see visually the nice fit suggested by the statistics in table 4. The outlier on the top right is the well-known small-firm effect. The point estimate gives a small-firm α of 0.23 percent per quarter, or about 1 percent per year.

The static CAPM has more difficulty with conditioning information, as seen in figure 4b. The four assets below the 45° line are the unscaled portfolios, with the scaled portfolios above the line. These pricing errors are much larger than the small-firm effect found in the unconditional estimates in figure 4a.

The scaled CAPM does a somewhat better job of handling conditioning information, as seen in figure 4c. However, the pricing errors are still fairly large, and this lies behind the statistical rejection shown in table 4. Even with scaled factors, the CAPM cannot price the scaled returns. This observation helps to give us some confidence that the investment model was not performing well only because of scaling.

Note also that the small-firm effect disappears once we include scaled market returns as factors. Thus the apparent small-firm effect may simply be due to inadequate treatment of conditioning information. Most derivations of the CAPM specify that the market is conditionally, but not unconditionally, mean-variance efficient, so this result is not too surprising. (It may also be due to a specific failure of the CAPM: none of the other models displays a small-firm effect.)

The pricing error in table 3 confirms the better visual fit of the investment return model in figure 2 versus the CAPM in figure 4. The unconditional estimate of the nonscaled CAPM produces a 0.09 percent RMSE α, about the same as the 0.11 percent value from the corresponding investment return model. However, the investment return model produces better than half the RMSE pricing errors in the conditional estimate: 0.46 percent versus 0.94 percent in the nonscaled model and 0.19 percent versus 0.49 percent in the scaled model.

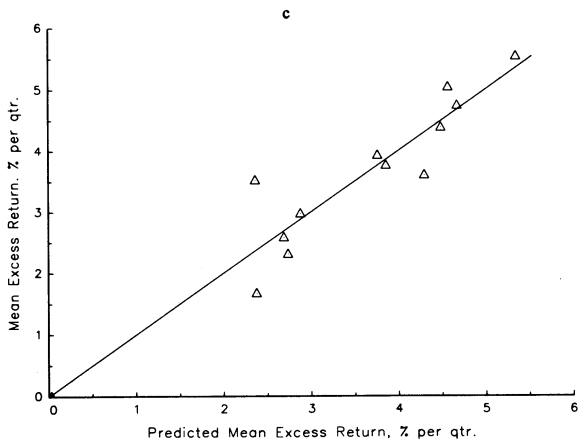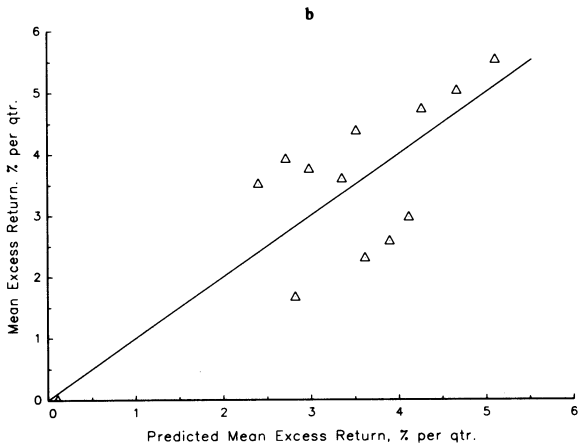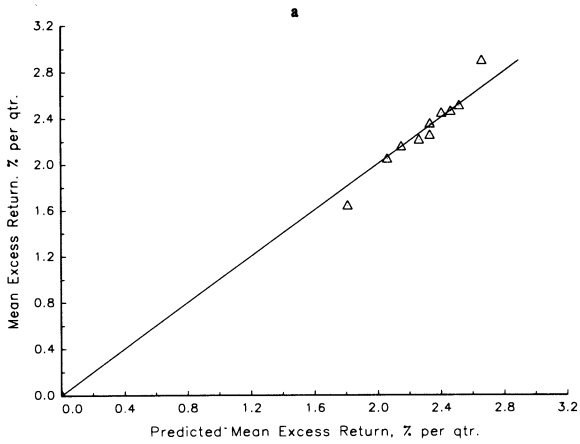Table 3 includes the RMSE pricing error for a model containing

FIG. 4.—Predicted vs. actual mean excess returns, CAPM: *a*, static CAPM; *b*, static CAPM, conditional estimates; *c*, scaled CAPM, conditional estimates. All estimates first-stage GMM.

both investment return and market factors, so we can see the effect of dropping either set of factors in the style of the model comparison tests of table 5. The investment return *plus* market model does not do meaningfully better than the investment return model taken alone: reductions from 0.42 percent to 0.41 percent or 0.46 percent to 0.45 percent for the nonscaled model and 0.15 percent to 0.12 percent or 0.19 percent to 0.15 percent for the scaled model. However, the investment return plus market return model seems to do meaningfully better than the CAPM, reducing pricing errors by one-half.

In summary, the pricing errors confirm the statistical tests for the nonscaled models: the investment return model does better than the static CAPM and drives out the CAPM factors. For the scaled models, the statistical tests found both investment and market factors important. However, we find that the rise in pricing error when investment return factors are omitted is economically large, even if statistically small, and the rise in pricing error when market return factors are omitted is not economically meaningful, even if statistically significant.

B.   Chen, Roll, and Ross Model

The Chen, Roll, and Ross (1986) (CRR) model was explicitly designed to link stock returns to economic fluctuations, and Chen, Roll, and Ross claim that their model drives out the market return. Thus it is an important alternative model to examine. Chen, Roll, and Ross advocate a five-factor model, in which the factors are MP (growth in industrial production), DEI (change in inflation forecast), UI (inflation forecast residual), UPR (return on corporate bonds minus return on 10-year government bonds), and UTS (return on 10-year government bonds minus return on bills). All but MP are based on bond returns (the inflation forecasts are based on Treasury-bill returns).

Table 6 presents GMM estimates and tests of the CRR model. The table presents only iterated estimates, since first-stage estimates were not different enough to warrant an extra set of numbers. Similarly, the ΔJ_T tests were almost identical to Wald tests, so I omitted them from the table.

In the unconditional estimates, only one of the CRR factors is individually significant, though they are jointly marginally significant with a 6.1 percent p-value. However, in the more efficient conditional estimate, two factors are individually significant, and the factors together are jointly significant with a 3.6 percent p-value. The model is comfortably not rejected by the J_T test, with 95 percent and 64 percent p-values.

It is not clear whether Chen, Roll, and Ross intend their model as

TABLE 6

CHEN, ROLL, AND ROSS MODEL

A. PARAMETER ESTIMATES

	b_0	b_{mp}	b_{dei}	b_{ui}	b_{upr}	b_{uts}
	Unconditional Estimates					
Coefficient	1.0	−1.4	−7.5	41	−54	−19
t-statistic	4.58	−.25	−.11	1.14	−1.88	−2.43
	Conditional Estimates					
Coefficient	1.8	−25	−68	−51	75	−.37
t-statistic	4.36	−2.30	−.51	−.82	2.39	−.03

B. TESTS

	$b_{mp-uts} = 0$	J_T
	Unconditional Estimates	
χ^2	10.6	1.15
Degrees of freedom	5	5
p-value (%)	6.1	95
	Conditional Estimates	
χ^2	11.9	5.13
Degrees of freedom	5	7
p-value (%)	3.6	64

NOTE.—Asset returns are deciles 1–10 in the unconditional estimates and deciles 1, 2, 5, and 10 scaled by the constant, term premium, and dividend/price ratio in the conditional estimates. Assets do not include investment returns.

a conditional or unconditional factor model. Their test allows some variation in β's but imposes constant factor risk premia (λ's), and they only attempt to explain unconditional expected returns. Pragmatically, I do not include a scaled Chen, Roll, and Ross model since it would have $1 + (3 \times 5) = 16$ b parameters and I use only 13 asset return moments. One can use more moments, of course, but a comfortable moments/parameters ratio would leave us with an uncomfortable data points/moments ratio.

The column marked "CRR" in table 5 presents a comparison of the investment model with the CRR model. The ΔJ_T test of the scaled model strongly rejects dropping the investment return factors in the presence of the CRR factors. Otherwise, the table suggests that either set of factors can be dropped in the presence of the others. This is good news: it means that the investment return model can explain the relatively good fit of the CRR model.

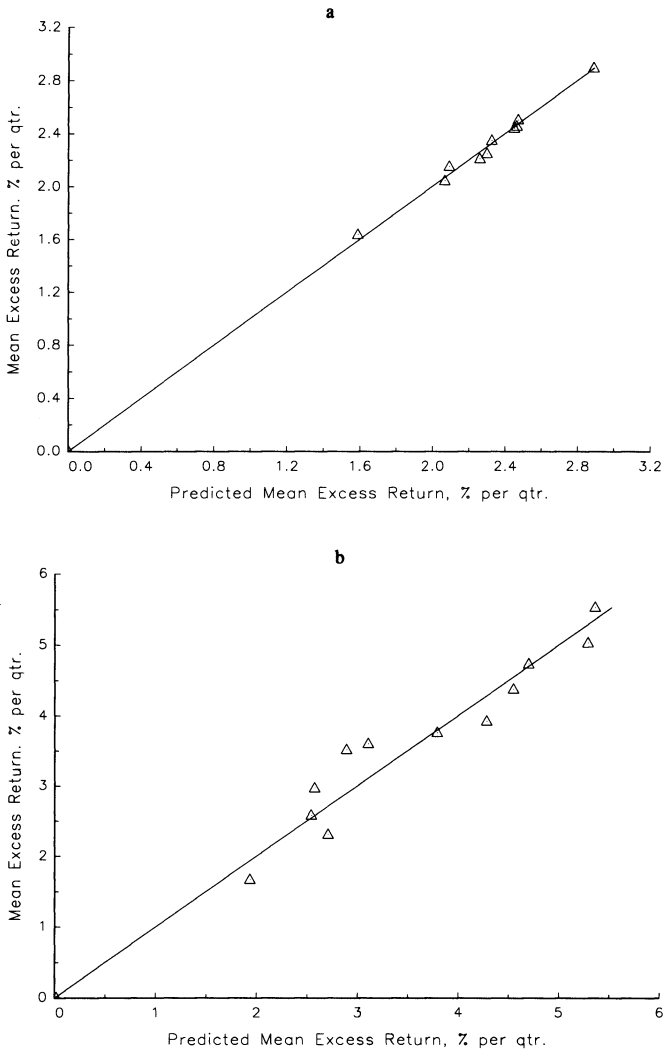Figure 5 plots the first-stage pricing errors and confirms the nice

FIG. 5.—Predicted vs. actual mean excess returns, Chen et al. model, first-stage estimates: *a*, unconditional estimates; *b*, conditional estimates.

fit, at least for the unconditional estimates. Keep in mind, though, that the model has six factors to fit 11 moments in the unconditional estimates and 13 moments in the conditional estimates. The fit in these plots does not correct for degrees of freedom.

Comparing figure 5 with the investment return pricing errors in figure 2 and examining the pricing error in table 3, we see that the CRR model gives stock return pricing errors of 0.33 percent, about

halfway between the nonscaled (0.46 percent) and scaled (0.19 percent) investment return factor model. Thus the scaled investment return factor may provide an economically interesting improvement over the CRR model.

C.   Consumption-Based Model

The consumption-based model is based on a measure of consumers' intertemporal marginal rate of *substitution*, where the investment model is based on a measure of firms' intertemporal marginal rate of *transformation*. It is perhaps the most appropriate comparison for the investment return factor model. Like the investment model, the consumption-based model relates asset returns strictly to macroeconomic data rather than other asset returns. It also provides an explicit link between asset return and macroeconomic events.

I limit my comparisons to the standard time-separable constant relative risk aversion formulation, which is about the same level of simplicity as this investment return model. It is possible that one of the many variations on the consumption-based model, such as habit persistence, durability, and so forth, may perform better than this simple model. (Campbell and Cochrane [1995] certainly hope so!) But it is also possible that one of the many possible variations on the investment model, such as production shocks, gestation lags, and so forth, performs better still.

Table 7 presents GMM estimates and tests of the basic consumption-based model. The model predicts that

$$m_t = \beta \left(\frac{c_t}{c_{t-1}}\right)^{-\gamma}$$

regardless of conditioning information, so I do not consider scaled consumption factors. The γ estimates are huge and the β's are often larger than one. This is the equity premium puzzle, a familiar pattern when this model tries to explain the cross section of asset returns.

The pricing errors for the unconditional estimates in figure 6a are huge. However, the imprecision of the estimate is so high that the unconditionally estimated model cannot be rejected by the J_T test—a graphic illustration of the dangers of looking only at J_T tests and not also the underlying pricing errors. Figure 6b shows that, like the CAPM, this consumption-based model has great trouble reconciling conditional and unconditional moments. In this case the pricing errors are so large that the model is resoundingly rejected by the J_T test with p-values of 0.3 percent in the first-stage estimate and 0.04 percent in the iterated estimate. Table 3 confirms the visual picture:

TABLE 7

CONSUMPTION-BASED MODEL

	PARAMETER ESTIMATES			
	Unconditional Estimates		Conditional Estimates	
	β	γ	β	γ
First-stage:				
Coefficient	.98	241	1.29	153
t-statistic	.49	.61	6.39	1.56
Iterated:				
Coefficient	1.27	71	1.29	116
t-statistic	10.9	2.17	13.9	3.36

	TESTS	
	Unconditional Estimates	Conditional Estimates
	J_T	J_T
First-stage:		
χ^2	6.17	28
Degrees of freedom	9	11
p-value (%)	72	.30
Iterated:		
χ^2	11.3	33.9
Degrees of freedom	9	11
p-value (%)	26	.04

NOTE.—GMM estimates and tests of consumption-based model: $m_{t+1} = \beta(c_{t+1}/c_t)^{-\gamma}$. Asset returns are deciles 1–10 in the unconditional estimates and deciles 1, 2, 5, and 10 scaled by the constant, term premium, and dividend/price ratio in the conditional estimates. Assets do not include investment returns.

the RMSE pricing error of the consumption-based model is 2.86 percent, 10 times higher than that of any of the other models.

The column marked $\Delta c^{-\gamma}$ in table 5 presents a comparison of the investment model against the consumption-based model. The unrestricted model here is $m = b_0 + \mathbf{b}_1' \mathbf{r}^I + \beta \Delta c^{-\gamma}$. Not surprisingly, all the tests decisively reject omitting the investment return factors, with p-values less than 1 percent and mostly less than 0.01 percent, while never rejecting dropping the $\Delta c^{-\gamma}$ factor.

D. Consumption Growth Factor Model

Perhaps *consumption* is not at fault, but the tight structure implied by the utility function and absence of free b parameters in the consumption-based model. Linearized versions of the consumption-based model such as Brown and Gibbons (1985) have been more successful. In this spirit, consider the model

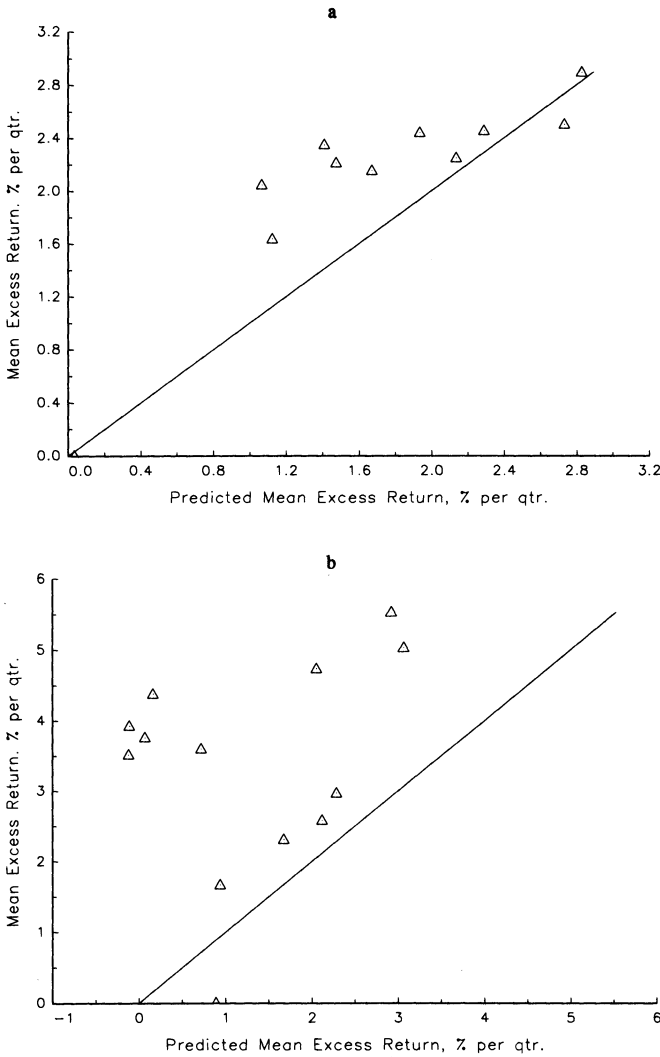$$m_{t+1} = b_0 + b_{\Delta c} \Delta c_{t+1},$$

FIG. 6.—Predicted vs. actual mean excess returns, consumption-based model, first-stage estimates: a, unconditional estimates; b, conditional estimates.

where Δc_{t+1} denotes consumption growth, and a similar model with scaled factors.

Table 8 presents iterated GMM estimates of this model (the first stage was not different enough to warrant an extra set of numbers). A familiar pattern emerges: in the unconditional estimates, consumption growth significantly prices assets and the model is not rejected. In the conditional estimates, consumption growth prices more signifi-

TABLE 8

CONSUMPTION GROWTH FACTOR MODEL

A.   NONSCALED FACTOR MODEL $m = b_0 + b_{\Delta c}\Delta c$

	PARAMETER ESTIMATES			
	Unconditional Estimates		Conditional Estimates	
	Constant	Δc	Constant	Δc
Coefficient	92	−90	108	−107
t-statistic	2.32	−2.30	2.84	−2.81

	TESTS	
	Unconditional Estimates	Conditional Estimates
	J_T	J_T
χ^2	9.88	34.8
Degrees of freedom	9	11
p-value (%)	36	.03

B.   SCALED FACTOR MODEL $m = b_0 + b_{\Delta c}\Delta c + b_{tp}(\Delta c \times tp) + b_{dp}(\Delta c \times tp)$:
CONDITIONAL ESTIMATES

	PARAMETER ESTIMATES			
	b_0	$b_{\Delta c}$	b_{tp}	b_{dp}
Coefficient	105	−102	−.31	−.53
t-statistic	2.73	−2.68	−1.76	−4.74

	TESTS	
	Scaled $b = 0$	J_T
χ^2	33	17
Degrees of freedom	2	9
p-value (%)	.00	4.6

NOTE.—Iterated GMM tests of consumption growth factor model. Asset returns are deciles 1–10 in the unconditional estimates and deciles 1, 2, 5, and 10 scaled by the constant, term premium, and dividend/price ratio in the conditional estimates. Assets do not include investment returns. Wald and ΔJ_T joint b tests give the same results, so they are not separately presented.

cantly, but the model is decisively rejected, with a 0.03 percent p-value that is much lower than corresponding rejections of the investment return factor model and the CAPM. Scaling helps: the scaled factors are mostly individually significant and jointly highly significant, and the model is now only borderline rejected with a 4.6 percent p-value.

The consumption-based factor model does not do all that well when compared to the investment return model in comparison tests in table 5, in pricing error comparisons in table 3, and by examination of the pricing errors in figure 7, however. In table 5, we reject exclusion of
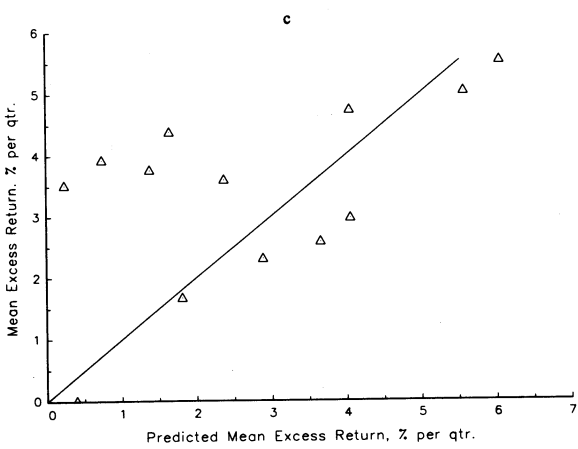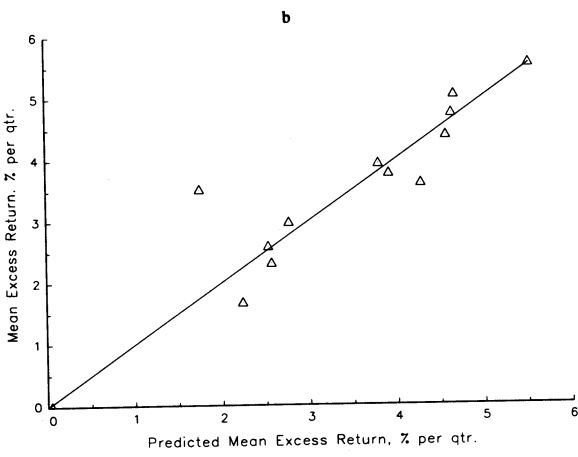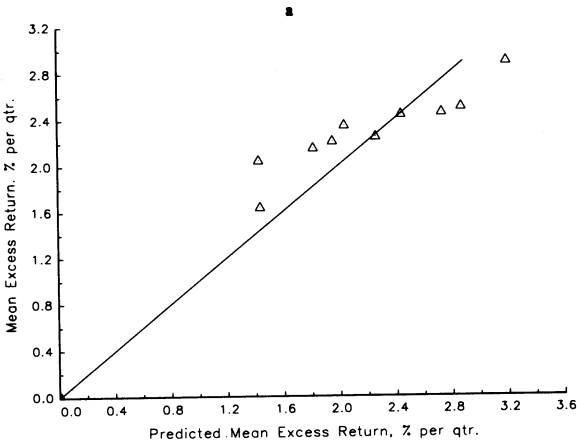
FIG. 7.—Predicted vs. actual mean excess returns, consumption growth factor model, first-stage estimates: *a*, unconditional estimates, nonscaled model; *b*, conditional estimates, nonscaled model; *c*, conditional estimates, scaled model.

the consumption growth factor only in the iterated estimate of non-scaled models; otherwise the evidence against dropping the investment return factors is much stronger than that against dropping the consumption growth factors. Figure 7 reveals pricing errors almost as large as those of the consumption-based model and dramatically worse than the fit of the other models. Table 3 shows that the RMSE pricing errors are three times greater than those of their investment return model and larger than the CAPM as well, whereas the investment return plus consumption growth model has pricing errors not much lower than those of the investment return model alone.

In summary, the problem does *not* seem to be the tight structure imposed by the consumption-based model, but that consumption data are less informative about stock returns than investment data.

E. Investment Growth Model

How much of the success of the investment return factor model has to do with the precise functional forms used to infer investment returns from investment data? To investigate this question, I consider an ad hoc factor model based on investment *growth*,

$$m = b_0 + b_{nr}\Delta i_{nr} + b_r\Delta i_r,$$

and its scaled extension.

Table 9 presents the usual GMM estimates and tests of the investment growth factor model. Again, I present only second-stage estimates since the first stage was quite similar. The performance is overall a little *better* than the investment return factor model of table 1. However, this estimate uses fewer moments, since it tries to price only asset returns and not asset and investment returns simultaneously. As before, only the residential investment factor prices assets significantly in the unconditional estimates, but both factors significantly price assets in the conditional estimates. The unconditional estimate of the nonscaled model is not rejected (p-value 26 percent), but now the conditional estimate is not rejected as well, with a p-value of 65 percent. Given this fact, it is unsurprising that scaling seems not to be statistically necessary; the scaled investment growth factors are individually and jointly insignificant. Though the scaled model is not rejected either with a 32 percent p-value, the J_T declines only from 7.8 to 7.0 with a loss of four degrees of freedom.

Figure 8 presents pricing errors. The fit of the 10 deciles is quite good. The scaled returns are not priced quite as well. As with the investment return factor model, the *statistically* marginal or insignificant improvement from adding scaled factors masks a definite *economic* improvement in the sample pricing errors. The RMSE pricing errors in table 3 show that this model does about as well as the invest-

TABLE 9

INVESTMENT GROWTH (Not Return) FACTOR MODEL

A. NONSCALED MODEL

| | PARAMETER ESTIMATES | | | | | |
| | Unconditional Estimates | | | Conditional Estimates | | |
	b_0	b_{nr}	b_r	b_0	b_{nr}	b_r
Coefficient	−4.2	12.3	−7.1	−51	62	−9.8
t-statistic	−.55	1.56	−2.59	−2.35	2.86	−2.62

| | TESTS | |
| | Unconditional Estimates | Conditional Estimates |
	J_T	J_T
χ^2	10	7.8
Degrees of freedom	8	10
p-value (%)	26	65

B. SCALED FACTOR MODEL: CONDITIONAL ESTIMATES

| | PARAMETER ESTIMATES | | | | | | |
	Constant	b_{nr}	b_r	$b_{nr \cdot tp}$	$b_{n \cdot tp}$	$b_{nr \cdot dp}$	$b_{r \cdot dp}$
Coefficient	−41	50	−8.3	3.4	−3.2	.15	−.18
t-statistic	−1.18	1.38	−.78	.81	−.77	.02	−.03

| | TESTS | | |
	Unscaled b	Scaled b	J_T
χ^2	2.0	1.9	7.0
Degrees of freedom	2	4	6
p-value (%)	37	75	32

NOTE.—Iterated GMM tests of investment growth factor model, using residential (r) and nonresidential (nr) gross fixed investment. Asset returns are deciles 1–10 in the unconditional estimates and deciles 1, 2, 5, and 10 scaled by the constant, term premium, and dividend/price ratio in the conditional estimates. Assets do not include investment returns. Wald and ΔJ_T joint b tests give the same results, so they are not separately presented.

ment return factor model, producing 0.51 percent RMSE versus 0.46 percent for the investment return model with nonscaled factors and 0.18 percent versus 0.19 percent for the scaled factor model.

In summary, the good behavior of the investment return factor model does not depend on the specific functional form.

VII. Concluding Remarks

The simple investment return model performs surprisingly well. The investment return factors significantly price assets, the model is not rejected, and it is able to explain a wide spread in expected returns,
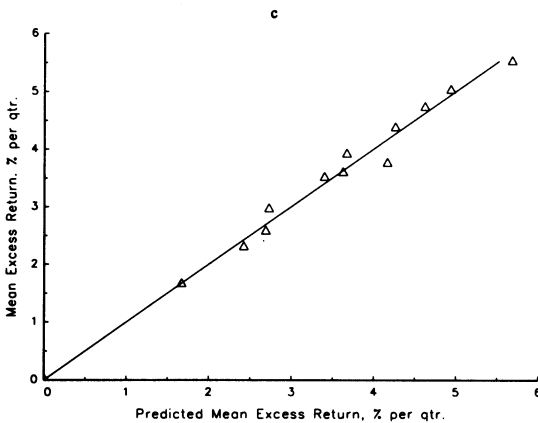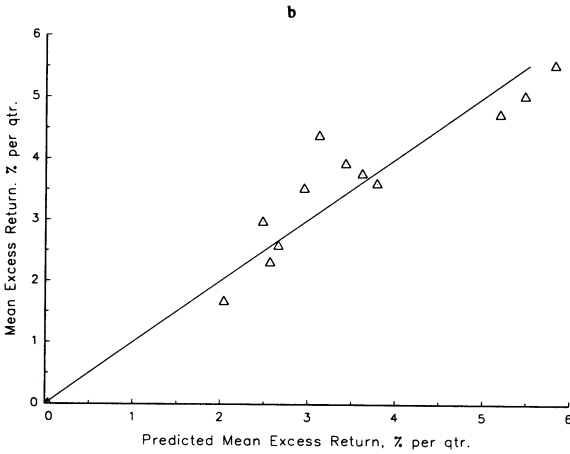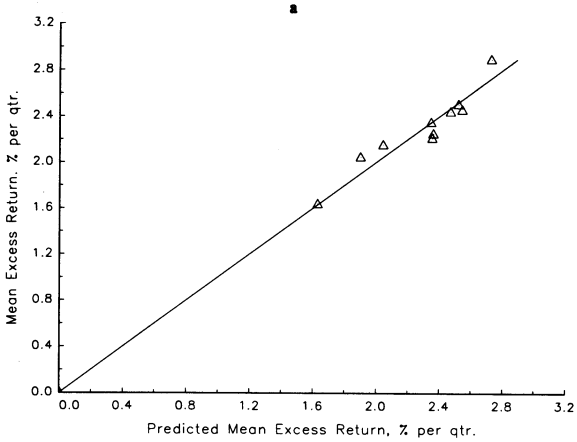
FIG. 8.—Predicted vs. actual mean excess returns, investment growth factor model, first-stage estimates: *a*, unconditional estimates, nonscaled model; *b*, conditional estimates, nonscaled model; *c*, conditional estimates, scaled model.

including managed portfolio returns formed by multiplying returns with instruments. The model performs about as well as two standard finance models, the CAPM and the Chen, Roll, and Ross factor model. The investment return model performs substantially better than the standard consumption-based model and an ad hoc consumption growth factor model. It is robust; an investment *growth* model performs about as well.

The fact that *any* model whose factors are related to economic theory and are based solely on quantity data is even in a position to challenge the empirical success of traditional finance models may be regarded as an encouraging initial success. Since any model can be expressed in terms of its mimicking portfolios and the latter are better measured, models based on quantity data or measures of real risk factors are always at a statistical disadvantage relative to models based on asset returns. And since one can always construct portfolios that perfectly price any set of assets ex post, the difficulty of obtaining a good fit depends entirely on the discipline one imposes in the search for factors.

The scaled factor models typically perform substantially better than the nonscaled factor models. This suggests that time variation in the parameters of asset pricing models, which can be handled by the simple expedient of including scaled factors, is an important ingredient for their empirical success.

A comparison of this paper and Cochrane (1991) with the empirical *q*-theory literature suggests that investment responds to changes in *risk premia* that the empirical finance literature has found to dominate changes in expected returns. Most *q*-theory models specify constant risk premia and try, without much success, to explain changes in investment from changes in risk-free rates. The relative success of the model presented here may help to rehabilitate the *q*-theory view of investment, amended to include substantial changes in risk premia over time.

More generally, macroeconomists are interested in the links between asset returns and fluctuations for the information they can provide about preferences, technologies, and market structures that will be useful in the construction of macroeconomic models. One lesson of these papers is that an adjustment cost (or some wedge between the price of installed and uninstalled capital), currently not included in most real business cycle models, is useful in order to reconcile investment and asset returns.

Appendix

A. Derivation of Investment Returns from the Production Function

This section derives the investment return from the production technology and shows that the firm's first-order conditions direct the firm to remove

arbitrage opportunities between investment and asset returns. This deriva-
tion follows that of Braun (1991); Cochrane (1991) presents a derivation of
the investment return directly from its definition as the marginal extra sale
possible tomorrow from a marginal investment today.

The firm maximizes its present value,

$$\max_{\{i_t\}} E_t \sum_{j=0}^{\infty} m_{t,t+j}(y_{t+j} - i_{t+j} - w_{t+j}l_{t+j}), \tag{A1}$$

subject to (1) and (2).

In a complete market, m are the contingent claims prices divided by proba-
bilities, so this present value is the firm's time t contingent claim value. If
markets are less than complete, the firm still maximizes (A1), but m is now
an extension of the stochastic discount factor that prices asset returns rather
than *the* stochastic discount factor for the whole economy. The marginal
utility of a nonsatiated owner of the firm who can also trade assets is one
such m.

I derive the first-order condition by varying i_t. Note that $\partial k_{t+j}/\partial i_t = (1 - \delta)^j$. Hence,

$$\frac{\partial y_{t+j}}{\partial i_t} = \frac{\partial y_{t+j}}{\partial k_{t+j}} \frac{\partial k_{t+j}}{\partial i_t} = (1 - \delta)^j[f_k(t + j) - c_k(t + j)].$$

The notation $f_k(t)$ means partial derivative with respect to k, evaluated with
respect to the appropriate arguments at time t; $f_k(t) \equiv \partial f(k_t, l_t)/\partial k_t$. The
first-order condition is then

$$1 + c_i(t) = E_t \sum_{j=1}^{\infty} m_{t,t+j}(1 - \delta)^j[f_k(t + j) - c_k(t + j)]. \tag{A2}$$

The left-hand side is the relative price of a unit of installed capital versus
output today; the right-hand side is the present value of its benefits.

We desire a model of returns, rather than price and present value. Using
$m_{t,t+j} = m_{t,t+1}m_{t+1,t+j}$, break the right-hand side of (A2) into two pieces:

$$1 + c_i(t) = E_t m_{t,t+1}(1 - \delta)[f_k(t + 1) - c_k(t + 1)]$$

$$+ E_t m_{t,t+1}(1 - \delta) \sum_{j=1}^{\infty} m_{t+1,t+1+j}(1 - \delta)^j$$

$$\times [f_k(t + 1 + j) - c_k(t + 1 + j)].$$

Substituting (A2) at time $t + 1$ for the sum in the right-hand side, we get

$$1 + c_i(t) = E_t[m_{t,t+1}(1 - \delta)[f_k(t + 1) - c_k(t + 1) + 1 + c_i(t + 1)]],$$

$$1 = E_t \left[m_{t,t+1} \frac{(1 - \delta)[1 + f_k(t + 1) + c_i(t + 1) - c_k(t + 1)]}{1 + c_i(t)c} \right]$$

or

$$1 = E_t[m_{t,t+1}r_{t+1}^i],$$

with

$$r_{t+1}^i = \frac{(1 - \delta)[1 + f_k(t + 1) + c_i(t + 1) - c_k(t + 1)]}{1 + c_i(t) c}.$$

For some production technologies it is not possible to summarize the price versus present value relation (A2) in a single-period investment return. For example, if the adjustment cost depends on p lags of investment, then a p-period investment strategy must be considered.

B. Data Description

All asset return data are taken from CRSP. National Income and Product Accounts data and yield data are taken from Citibase. The two investment returns are based on Citibase series GINQ and GIRQ. The stock return series are based on CRSP series EWRETD and VWRETD and the size decile return series DECRET1 . . . DECRET10. The default premium is based on Citibase series FYBAAC–FYAAAC. Quarterly data are obtained by using the last month of the quarter. The dividend/price ratio is based on CRSP EWRETD and EWRETX, the equally weighted portfolio returns with and without dividends. The returns are cumulated for a year to avoid the seasonal in dividends; then d/p = (annual EWRETD/annual EWRETX) − 1. Again, the last monthly observation in each quarter is the quarterly observation.

The investment data are quarterly averages, and the asset return data are point-to-point. As an ad hoc correction for this difference, I averaged monthly asset returns over the quarter to correspond with the investment returns (I thank Campbell Harvey for suggesting this transformation). Thus the second-quarter return is an average of returns from the last day in December to the last day in March, the last day in January to the last day in April, and the last day in February to the last day in May. Instruments for the second-quarter return are all observed at the end of December (i.e., all instruments are lagged twice).

I constructed Chen, Roll, and Ross factors as follows: MP is the growth rate of industrial production. Chen, Roll, and Ross lead this variable by one month to take account of the fact that industrial production (IP) is a monthly average and returns are end-of-month to end-of-month. To make the same adjustment for quarterly data, I average IP growth in a similar way to returns. For example, the second-quarter MP is

$$\text{MP} = \ln[\text{IP(Apr)}\,\text{IP(May)}\,\text{IP(Jun)}] - \ln[\text{IP(Jan)}\,\text{IP(Feb)}\,\text{IP(Mar)}].$$

UI is unexpected inflation and DEI is the change in expected inflation. These variables require an expected inflation series. Chen, Roll, and Ross take their values from Fama and Gibbons (1982). Therefore, I replicated the Fama and Gibbons procedure to extend the data set. Fama and Gibbons start with the Fisher equation

$$E_{t-1}(\pi_t) = TB_{t-1} - E_{t-1}(R_t),$$

where π is Consumer Price Index inflation, TB is the Treasury-Bill rate, r^{tb} is the ex post real rate, and $r_t^{tb} = TB_{t-1} - \pi_t$. They add a univariate time-

series model for ex post real rates

$$r_t^{tb} - r_{t-1}^{tb} = u_t + \theta u_{t-1}.$$

Substituting, we get

$$E_{t-1}(\pi_t) = TB_{t-1} - r_{t-1}^{tb} - \theta u_t.$$

To construct this series, I take Fama and Gibbons's value of $\theta = 0.9223$. I start with $u_1 = 0$. Then I construct u_t by

$$r_2^{tb} - r_1^{tb} = u_2,$$

$$r_3^{tb} - r_2^{tb} = u_3 + \theta u_2,$$

and so forth. The expected real return on Treasury bills is then given by

$$E_{t-1}r_t^{tb} = E_{t-2}r_{t-1}^{tb} + (1 - 0.9223)u_{t-1}.$$

References

Braun, Phillip. "Asset Pricing and Capital Investment: Theory and Evidence." Manuscript. Evanston, Ill.: Northwestern Univ., 1991.

Brock, William A. "Asset Prices in a Production Economy." In *The Economics of Information and Uncertainty*, edited by John J. McCall. Chicago: Univ. Chicago Press (for NBER), 1982.

Brown, David P., and Gibbons, Michael R. "A Simple Econometric Approach for Utility-Based Asset Pricing Models." *J. Finance* 40 (June 1985): 359–81.

Campbell, John Y., and Cochrane, John H. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." Working Paper no. 4995. Cambridge, Mass.: NBER, January 1995.

Chen, Nai Fu; Roll, Richard; and Ross, Stephen A. "Economic Forces and the Stock Market." *J. Bus.* 59 (July 1986): 383–403.

Cochrane, John H. "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives." *A.E.R.* 79 (June 1989): 319–37.

———. "Production-Based Asset Pricing and the Link between Stock Returns and Economic Fluctuations." *J. Finance* 46 (March 1991): 209–37.

Cochrane, John H., and Hansen, Lars Peter. "Asset Pricing Lessons for Macroeconomics." In *NBER Macroeconomics Annual 1992*, edited by Olivier J. Blanchard and Stanley Fischer. Cambridge, Mass.: MIT Press, 1992.

Cox, John C.; Ingersoll, Jonathan E., Jr.; and Ross, Stephen A. "An Intertemporal General Equilibrium Model of Asset Prices." *Econometrica* 53 (March 1985): 363–84.

De Santis, Giorgio. "Volatility Bounds for Stochastic Discount Factors: Tests and Implications from International Stock Returns." Manuscript. Los Angeles: Univ. Southern California, 1992.

Dybvig, Philip H., and Ingersoll, Jonathan E., Jr. "Mean-Variance Theory in Complete Markets." *J. Bus.* 55 (April 1982): 233–51.

Fama, Eugene F., and Gibbons, Michael R. "Inflation, Real Returns and Capital Investment." *J. Monetary Econ.* 9 (May 1982): 297–323.

Fama, Eugene F., and MacBeth, James D. "Risk, Return, and Equilibrium: Empirical Tests." *J.P.E.* 81 (May/June 1973): 607–36.

Ferson, Wayne E., and Foerster, Stephen R. "Finite Sample Properties of the Generalized Method of Moments Tests of Conditional Asset Pricing Models." *J. Financial Econ.* 36 (August 1994): 29–55.

Ferson, Wayne E.; Kandel, Shmuel; and Stambaugh, Robert F. "Tests of Asset Pricing with Time-Varying Expected Risk Premiums and Market Betas." *J. Finance* 42 (June 1987): 201–20.

Gibbons, Michael R.; Ross, Stephen A.; and Shanken, Jay. "A Test of the Efficiency of a Given Portfolio." *Econometrica* 57 (September 1989): 1121–52.

Hansen, Lars Peter. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50 (July 1982): 1029–54.

Hansen, Lars Peter; Heaton, John; and Luttmer, Erzo. "Econometric Evaluation of Asset Pricing Models." *Rev. Financial Studies* 8 (Summer 1995): 237–74.

Hansen, Lars Peter, and Jagannathan, Ravi. "Assessing Specification Errors in Stochastic Discount Factor Models." Manuscript. Chicago: Univ. Chicago; Minneapolis: Univ. Minnesota, 1991. (*a*)

———. "Implications of Security Market Data for Models of Dynamic Economies." *J.P.E.* 99 (April 1991): 225–62. (*b*)

Hansen, Lars Peter, and Richard, Scott F. "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models." *Econometrica* 55 (May 1987): 587–613.

Hansen, Lars Peter, and Singleton, Kenneth J. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica* 50 (September 1982): 1269–86.

Harvey, Campbell R. "Time-Varying Conditional Covariances in Tests of Asset Pricing Models." *J. Financial Econ.* 24 (October 1989): 289–317.

Kandel, Shmuel, and Stambaugh, Robert F. "Portfolio Inefficiency and the Cross-Section of Expected Returns." *J. Finance* 50 (March 1995): 157–84.

Knez, Peter J. "Pricing Money Market Securities with Stochastic Discount Factors." Manuscript. Madison: Univ. Wisconsin, 1991.

Luttmer, Erzo. "Implications of Asset Market Data for Economies with Transaction Costs." Manuscript. Evanston, Ill.: Northwestern Univ., 1992.

Newey, Whitney K., and West, Kenneth D. "Hypothesis Testing with Efficient Method of Moments Estimation." *Internat. Econ. Rev.* 28 (October 1987): 777–87. (*a*)

———. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (May 1987): 703–8. (*b*)

Ross, Stephen A. "A Simple Approach to the Valuation of Risky Streams." *J. Bus.* 51 (July 1978): 453–75.

Shanken, Jay. "Intertemporal Asset Pricing: An Empirical Investigation." *J. Econometrics* 45 (July–August 1990): 99–120.

Sharathchandra, Gopalakrishnan. "Asset Pricing and Production: Theory and Empirical Tests." Manuscript. Dallas: Southern Methodist Univ., Cox School Bus., Finance Dept., 1991.

Snow, Karl N. "Diagnosing Asset Pricing Models Using the Distribution of Asset Returns." *J. Finance* 46 (July 1991): 955–83.