NBER WORKING PAPER SERIES

COVARIANCES VERSUS CHARACTERISTICS IN GENERAL EQUILIBRIUM

Xiaoji Lin
Lu Zhang

Working Paper 17285
http://www.nber.org/papers/w17285

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2011

We thank Frederico Belo, Jonathan Berk, Michael Brennan, Andrew Chen, Bob Dittmar, Wayne Ferson, Adlai Fisher, Kewei Hou, Lars-Alexander Kuehn, Mike Lemmon, Stavros Panageas, Jules van Binsbergen, Toni Whited, and seminar participants at the University of British Columbia's Phillips, Hager, and North Centre for Financial Research Summer Finance Conference in 2011 for helpful discussions. The Matlab and SAS programs used to produce the results reported in the paper are available on the authors' Web sites. All remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Covariances versus Characteristics in General Equilibrium
Xiaoji Lin and Lu Zhang
NBER Working Paper No. 17285
August 2011
JEL No. D51,D53,D58,E22,E44,G12,G14,G31

ABSTRACT

We question a deep-ingrained doctrine in asset pricing: If an empirical characteristic-return relation is consistent with investor "rationality," the relation must be "explained" by a risk factor model. The investment approach changes the big picture of asset pricing. Factors formed on characteristics are not necessarily risk factors: Characteristics-based factor models are linear approximations of firm-level investment returns. The evidence that characteristics dominate covariances in horse races does not necessarily mean mispricing: Measurement errors in covariances are more likely to blame. Most important, the investment approach completes the consumption approach in general equilibrium, especially for cross-sectional asset pricing.

Xiaoji Lin
Department of Finance
Fisher College of Business
The Ohio State University
2100 Neil Avenue
Columbus OH 43210
x.lin6@lse.ac.uk

Lu Zhang
Fisher College of Business
The Ohio State University
2100 Neil Avenue
Columbus, OH 43210
and NBER
zhanglu@fisher.osu.edu

1 Introduction

It is time to question a deep-ingrained doctrine in asset pricing: If an empirical characteristic-return relation is consistent with investor "rationality," the relation must be "explained" by a risk factor model. In particular, Fama and French (1993, 1996) argue that common factors formed on characteristics such as their size and value factors are risk factors in the context of Merton's (1973) intertemporal capital asset pricing model or Ross's (1976) arbitrage pricing theory. Building on the same doctrine as the basic premise, Daniel and Titman (1997) argue that the evidence that characteristics dominate covariances in asset pricing tests must mean mispricing.

We start in Section 2 by reviewing how the deep-rooted doctrine permeates critical writings in asset pricing and capital markets research in accounting. In Section 3, we use the simplest example of a general equilibrium production economy from Fisher (1930) to show that the consumption approach and the investment approach are two parallel ways of asset pricing in the earliest days of modern economics. We then construct a general equilibrium production economy to define the investment approach to asset pricing as conceptually equivalent to the weighted average cost of capital approach to capital budgeting in corporate finance. While the consumption approach of Lucas (1978) connects expected stock returns to consumption betas, the investment approach of Cochrane (1991) connects stock returns to firm characteristics.

In Section 4, we challenge the doctrine in four steps. First, building on Brock (1982), we show that, conceptually, only sources of aggregate uncertainty, such as shocks to total factor productivity and government policy shocks, are risk factors in stock returns. To the extent that the investment approach implies that characteristics — as components of firm-level investment returns — should forecast stock returns, we interpret characteristics-based factor models as linear approximations to firm-level investment returns. Clearly, this characteristics-based interpretation from the investment approach differs from the risk factors interpretation from the consumption approach (Section 4.1).

Second, we argue that in general equilibrium, covariances and characteristics are conceptually equivalent “explanatory” variables of expected returns. In particular, the investment approach predicts that characteristics are sufficient statistics of expected returns: Once characteristics are controlled for, covariances should not affect the cross section expected returns. This “outrageous” statement is in effect the dual counterpart to the equally strong statement that covariances are sufficient statistics of expected returns from the standard consumption approach (Section 4.2).

If covariances and characteristics are equivalent in theory, why do characteristics dominate covariances in asset pricing tests? We argue that mispricing is not the only possibility. In particular, we demonstrate quantitatively, using simulations from the Zhang (2005) economy, how measurement errors in *estimated* covariances can make covariances lose the Daniel and Titman (1997) horse races to characteristics, even in a world embedded with a conditional risk structure (Section 4.3).

Third, we articulate the (lack of) causality among covariances, expected returns, and characteristics in general equilibrium. Causality runs from covariances to expected returns *only* in the Lucas tree economy, in which quantities are fixed. However, causality runs *from characteristics to expected returns* in a linear technologies economy such as Cox, Ingersoll, and Ross (1985), in which quantities are freely adjusted. In a more realistic general equilibrium economy with adjustment costs, covariances, expected returns, and characteristics are all endogenous variables that are determined *simultaneously* by a system of equilibrium conditions. *No* causality runs from covariances to expected returns, from characteristics to expected returns, or vice versa. As such, the investment approach is as “causal” as the consumption approach in “explaining” anomalies (Section 4.4).

Fourth, we address several common critiques on the investment approach. We argue that: (i) The critiques largely originate from treating the consumption approach as the only approach in asset pricing; and (ii) the investment approach changes the big picture of asset pricing (Section 4.5).

In Section 5, we conclude that the covariances versus characteristics dichotomy of the world should be abandoned, and that the investment approach is a new basis for asset pricing research.

2 The Deep-Instilled Doctrine

The doctrine permeates critical writings in asset pricing: If an empirical characteristic-return relation is consistent with investor "rationality," the relation must be "explained" by a risk factor model. In a stream of extremely influential articles, Fama and French (1993, 1996, 1997) argue for the risk-based interpretation of their common factors formed on size and book-to-market equity. In particular, Fama and French (1993, p. 4–5) write:

> "[I]f assets are priced rationally, variables that are related to average returns, such as size and book-to-market equity, must proxy for sensitivity to common (shared and thus undiversifiable) risk factors in returns. The time-series regressions give direct evidence on this issue. In particular, the slopes and R^2 values show whether mimicking portfolios for risk factors related to size and [book-to-market] capture shared variation in stock and bond returns not explained by other factors."

Fama and French (1996, p. 57) further claim:

> "[T]he empirical successes of [the three-factor model] suggest that it is an equilibrium pricing model, a three-factor version of Merton's (1973) intertemporal CAPM (ICAPM) or Ross's (1976) arbitrage pricing theory (APT). In this view, SMB and HML mimic combinations of two underlying risk factors or state variables of special hedging concern to investors."

The Fama and French's risk factor interpretation of their size and book-to-market factors has been controversial. Disputing this interpretation, Daniel and Titman (1997) study whether portfolios with similar characteristics but different loadings on the Fama-French factors have different expected returns. Finding that expected returns do not correlate positively with the factor loadings after controlling for characteristics, Daniel and Titman (p. 4) write:

> "Our results are disturbing in that, like Fama and French (1992), they suggest that

> traditional measures of risk do not determine expected returns. In equilibrium asset pricing models the covariance structure of returns determines expected returns. Yet we find that variables that reliably predict the future covariance structure do not predict future returns. Our results indicate that high book-to-market stocks and stocks with low capitalizations have high average returns whether or not they have the return patterns (i.e., covariances) of other small and high book-to-market stocks. Similarly, after controlling for size and book-to-market ratios, a common share that 'act like' a bond (i.e., has a low market beta) has the same expected return as other common shares with high market betas."

The Daniel and Titman (1997) covariances versus characteristics test has become the workhorse for "disentangling" risk versus mispricing in asset pricing. Davis, Fama, and French (2000) conduct the Daniel-Titman test in a longer sample from 1929 to 1997, and find that covariances seem to do better than characteristics in forecasting returns. Daniel, Titman, and Wei (2001) conduct their tests in Japanese stocks, and find that book-to-market is more related to stock returns than HML loadings. Gebhardt, Hvidkjaer, and Swaminathan (2005) conduct similar tests in corporate bond returns. After finding that default betas are significantly correlated to average bond returns even after controlling for duration, ratings, and yield-to-maturity, Gebhardt et al. conclude that systematic risk matters for corporate bonds.

The doctrine is alive and well today. Hirshleifer, Hou, and Teoh (2011, p. 1) apply the covariances versus characteristics test in the context of the accrual anomaly of Sloan (1996):

> "According to rational frictionless asset pricing models, the ability of accruals to predict returns should come from the loadings on this accrual factor-mimicking portfolio. However, our tests indicate that it is the accrual characteristic rather than the accrual factor loading that predicts returns. These findings suggest that investors misvalue the accrual characteristic and cast doubt on the rational risk explanation."

Hou, Karolyi, and Kho (2011, p. 4) use global stock returns to examine:

> "whether the cross-sectional explanatory power of our global factor-mimicking portfolios is directly related to the firm-level characteristics on which they are based — for such reasons as investor over- or under-reaction or illiquidity — or whether it derives from the covariance structure of returns that is related to these characteristics."

The evidence that characteristics dominate covariances in the Daniel and Titman (1997) test has often been interpreted as a sign of mispricing. Surveying the behavioral finance literature, Daniel, Hirshleifer, and Teoh (2002, p. 152) write:

> "[F]or the factors in these [characteristics-based factor] models to represent risk factors, it would have to be the case that the factor realizations have a strong covariance with investors' marginal utility across states. For example, the empirical evidence shows that growth (low book-to-market) stocks have had consistently low returns given their CAPM beta. For these low returns to be consistent with a rational asset pricing model, the distribution of returns provided by a portfolio of growth stocks must be reviewed by investors as 'insurance'; it must provide high returns in bad (high marginal utility) states and low returns in good (low marginal utility) states."

Barberis and Thaler (2003, p. 1091–1092) also argue:

> "One general feature of the rational approach is that it is loadings or betas, and not firm characteristics, that determine average returns." "[Daniel and Titman's (1997)] results appear quite damaging to the rational approach." "[Rational] models typically measure risk as the covariance of returns with marginal utility of consumption. Stocks are risky if they fail to pay out at times of high marginal utility—in 'bad' times—and instead pay out when marginal utility is low—in 'good' times. The problem is that for many of the above [anomalies], there is little evidence that the portfolios with anomalously

> *high* average returns do poorly in bad times, whatever plausible measure of bad times is used (original emphasis)."

The literature on capital markets research in accounting has also been deeply influenced by the Daniel and Titman (1997) study. Richardson, Tuna, and Wysocki (2010, p. 430) write:

> "[I]t is important to empirically distinguish (1) the covariance between stock returns and a given attribute from (2) the returns attributable to the characteristic. Finding evidence in support of (1) is consistent with a risk based explanation for the return relation, whereas finding (2) would suggest mispricing. Specifically, [Hirshleifer, Hou, and Teoh (2011)] find that high (low) accrual firms earn lower (higher) returns after controlling for the extent to which high (low) accrual firms load on a factor mimicking portfolio reflecting the accrual return. This evidence suggests that the relation between the accrual measure and future returns is due to the characteristic and not comovement, and tends to support the mispricing explanation over the risk-based explanation for the accrual anomaly (p. 430)."

Dechow, Khimich, and Sloan (2011) echo further:

> "Risk-based explanations are the natural default explanation of efficient market aficionados for any anomaly. The basic idea is that stocks with predictably higher (lower) returns must be more (less) risky. Investors are assumed to have already figured this out and have priced the stocks accordingly. In order to make a compelling case that the accrual anomaly is attributable to risk, one first has to come up with a story as to why investors find low (high) accrual stock more (less) risky. Ideally, one would like to identify the underlying risk factor and show that it subsumes accruals in predicting future returns (p. 14)" "[T]he evidence in Hirshleifer et al. (2011) suggests that it is the accrual characteristic rather than the accrual factor that predicts future stock returns.

This makes it difficult to attribute any related variation in discount rates to rationally priced risk (p. 23)."

3 Two Model Economies

To question the doctrine, we start from the Fisherian economy in Section 3.1.[1] In Section 3.2, we use a two-period general equilibrium production economy à la Long and Plosser (1983) to define the investment approach and to clarify its relation with the consumption approach to asset pricing.

3.1 The Fisherian Economy

Using the first intertemporal general equilibrium model, Fisher (1930) reconciles two theories of the interest rate, one based on consumer time preferences and the other based on real investment opportunity. We use the Fisherian economy to show that the basic idea of the investment approach was already evident in the earliest days of modern economics.

We follow Rubinstein's (2006, p. 57–60) presentation of the Fisherian economy. Each agent is both the consumer and the producer of a single aggregate consumption good under certainty. There are two dates, 0 and 1. $U(C_0)$ and $U(C_1)$ are the utility of consumption with $U' > 0$ and $U'' < 0$. ρ is time preference. W_0 is the initial endowment of the consumption good. I_0 is the amount of W_0 used in production so that $C_0 = W_0 - I_0$. $f(I_0)$ is the output from production of date 1's consumption with $f' > 0$ and $f'' < 0$, so that $C_1 = f(I_0)$. With certainty, the real interest rate is a constant, r. Asset pricing in this economy aims to pin down the equilibrium interest rate.

The optimal consumption-saving problem for the consumer is:

$$\max_{\{C_0, C_1\}} U(C_0) + \rho U(C_1) \quad \text{subject to} \quad W_0 = C_0 + \frac{C_1}{r}. \tag{1}$$

---

[1] Rubinstein (2006) writes: "Fisher (1930) is the seminal work for most of the financial theory of investments during the twentieth century." "Fisher develops the first formal equilibrium model of an economy with both intertemporal exchange and production. In so doing, at one swoop, he not only derives present value calculations as a natural economic outcome in calculating wealth, he also justifies the maximization of present value as the goal of production and derives determinants of the interest rates that are used to calculate present value (p. 55)."

Consumption first-order condition quickly says:

$$r = 1/\left[\rho \frac{U'(C_1)}{U'(C_0)}\right]. \tag{2}$$

This equation provides the standard consumption-based equation of the equilibrium interest rate: $r = 1/E[M]$ in which $M = \rho U'(C_1)/U'(C_0)$ is the stochastic discount factor.

However, the consumer's problem is only one half of the general equilibrium economy. The other half is the producer's optimal investment problem:

$$\max_{\{I_0\}} \; -I_0 + \frac{f(I_0)}{r}, \tag{3}$$

in which the producer takes the equilibrium interest rate as given when choosing optimal investment, I_0. Investment first-order condition quickly says:

$$r = f'(I_0), \tag{4}$$

which gives the investment-based equation of the equilibrium interest rate.

Fisher (1930) reconciles the two theories of the equilibrium interest rate, the consumption-based equation (2) and the investment-based equation (4). Take the social planner's problem:

$$\max_{\{C_0, C_1\}} \; U(C_0) + \rho U(C_1) \quad \text{subject to} \quad C_0 = W_0 - I_0 \text{ and } C_1 = f(I_0). \tag{5}$$

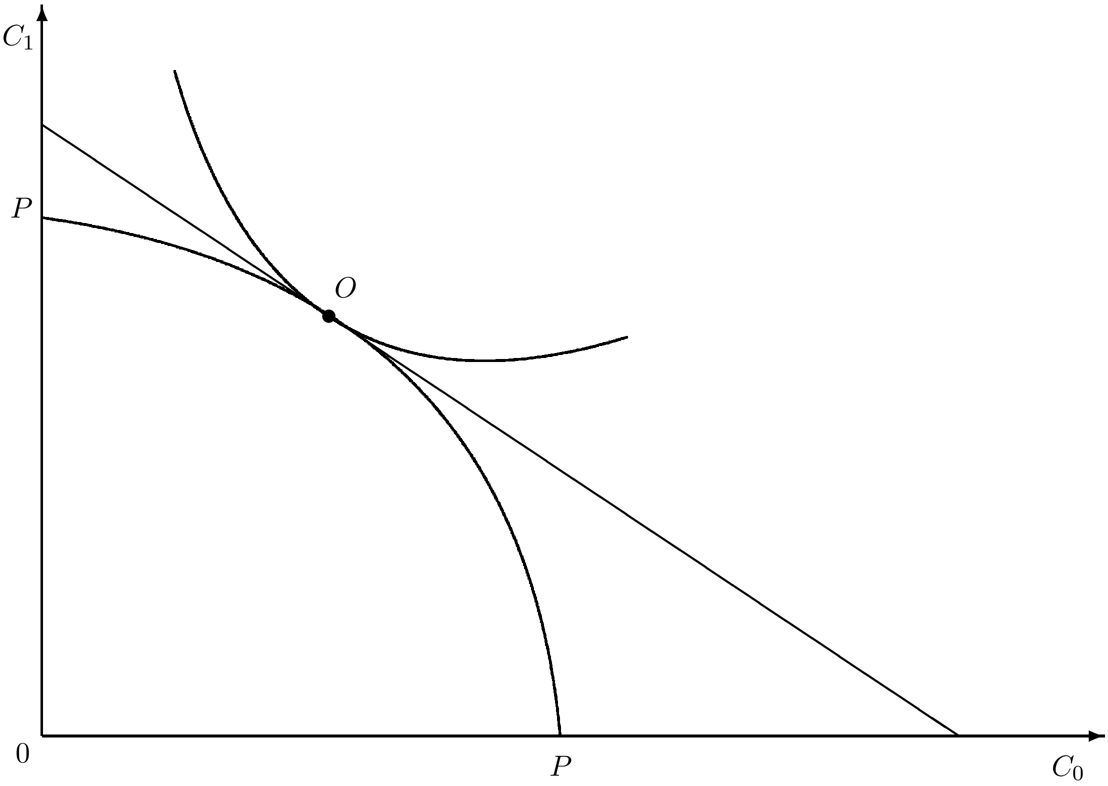
The first-order condition says that:

$$1/\left[\rho \frac{U'(C_1)}{U'(C_0)}\right] = f'(I_0), \tag{6}$$

meaning that the consumption-based and the investment-based theories give the identical answer.

The Fisherian economy is so stylized that it can be analyzed using graphical tools, as done in Hirshleifer (1970). Figure 1 shows the representative consumer-producer's optimum. The curved locus PP represents the individual's production frontier, and its concavity means diminishing returns

Figure 1 : Optimum for the Representative Individual in the Fisherian Economy



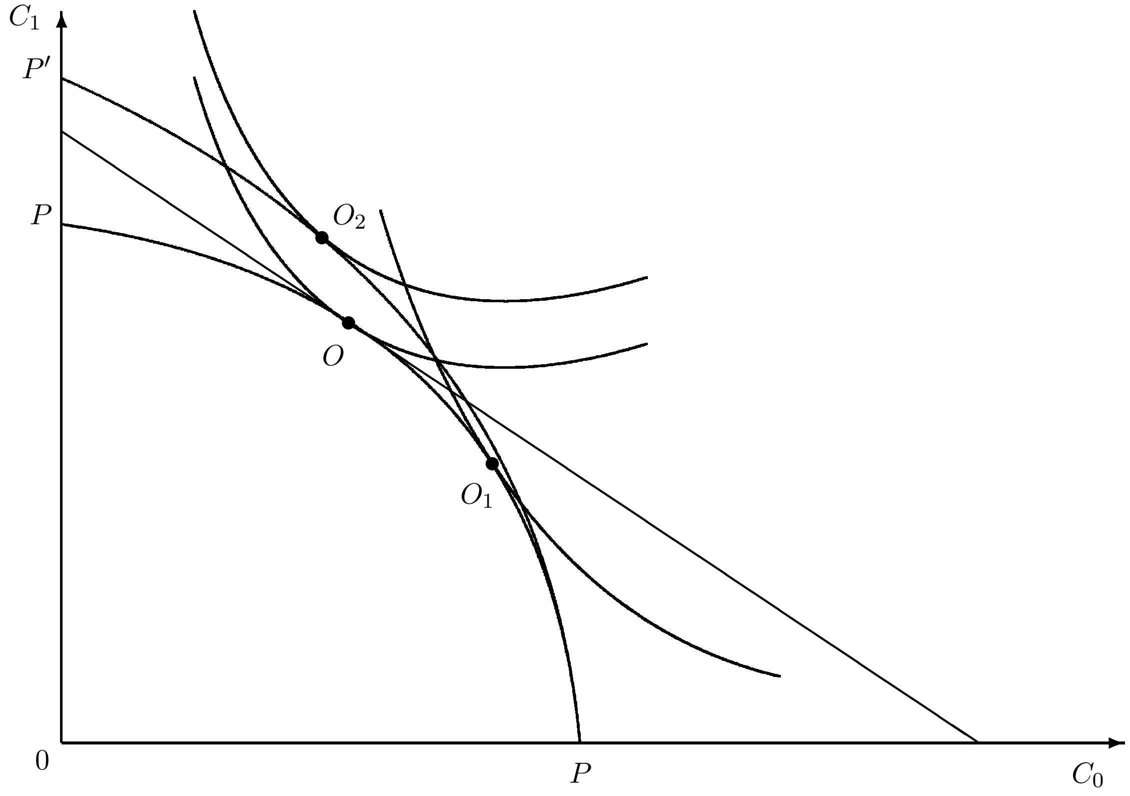
to scale. The tangent point, O, between the production frontier and the individual's indifference curve is the optimum, and the slope of the straight tangent line gives the equilibrium interest rate, r.

Figure 2 presents two comparative statics on the equilibrium interest rate. In the first experiment, we increase the marginal time preference of the individual, meaning a general steepening of the indifference curve. The optimum moves from point O to O_1, giving rise to an increase in the equilibrium interest rate. In the second experiment, we increase the marginal productivity of C_0 in generating C_1, meaning a steeper production frontier $P'P$. The optimum moves from point O to O_2, again resulting to an increase in the equilibrium interest rate.

Summarizing, Hirshleifer (1970, p. 192) writes:

> "The interest rate reflects the marginal rate of substitution between present and future consumption." "Geometrically, the conditions represent tangencies between market lines and indifference curves. The margin of choice involved is between dated consump-

Figure 2 : Comparative Statics of the Equilibrium Interest Rate



> tions, so that this is the time-preference margin."

In addition,

> "the interest rate reflects the marginal rate of substitution between present and future production." "Geometrically, these conditions correspond to tangencies between productive-opportunity loci and market lines."

As evident from the Fisherian economy, the consumption approach (the time-preference margin) is only one of two ways to derive the equilibrium interest rate. The investment approach (the time-productivity margin) was recognized by Fisher (1930) at the very beginning of modern economics, and was further elaborated by Hirshleifer (1970) as an integral part of asset pricing (see also Fama and Miller (1972)). In a treatise on capital theory, Hirshleifer (p. v) defines "capital theory" as representing "an extension of economic analysis into the domain of *time* and consequently *uncertainty* (the future being intrinsically uncertain) (original emphasis)." This "domain" is now what we call asset pricing. Clearly outside the scope of the covariances-based consumption framework, $r = f'(I_0)$ cannot be reduced to different functional forms of M. Also, to the extent that consumption and investment are two equally important building blocks of general equilibrium, $r = f'(I_0)$ is as primitive as $r = 1/E[M]$ in pinning down the equilibrium interest rate.

3.2 A Two-Period General Equilibrium Production Economy

Having shown that the investment approach was seeded in the earliest days of modern economics, we use a two-period general equilibrium economy to define what we mean by the investment approach and to clarify its relation with the standard consumption approach to asset pricing.

The economy, similar in spirit to the Long and Plosser (1983) economy, has three distinguishing characteristics that define new classical economics, a school of thought in modern macroeconomics that builds its analysis entirely on the neoclassical framework and emphasizes rigorous foundations based on microeconomics. The three characteristics are: (i) Agents have rational expectations; (ii) individuals maximize utility and firms maximize market value of equity; and (iii) agents act independently on the basis of full and relevant information.

There are two dates, 0 and 1. A representative household maximizes expected utility:

$$U(C_0) + \rho E_0[U(C_1)], \tag{7}$$

in which ρ is time preference, and C_0 and C_1 are aggregate consumption in dates 0 and 1, respectively. There are N firms, indexed by i, producing a single commodity to be consumed or invested in productive capital. Firm i starts with capital K_{i0} and produces in both dates. The firm exits at the end of date 1 with a liquidation value of zero. In particular, the depreciation rate of capital is 100%.

The operating cash flow of firm i is $\Pi_{it}K_{it}$, for $t = 0, 1$, in which Π_{it} is firm i's productivity subject to a vector of aggregate shocks affecting all firms and a vector of firm-specific shocks affecting only firm i. Both types of shocks are exogenous. The operating profits function exhibits constant returns to scale, meaning that Π_{it} is both marginal product of capital and average product of capital. Capital at the beginning of date 1 is $K_{i1} = I_{i0}$, in which I_{i0} is investment for date 0. Investing involves adjustment costs, $(a/2)(I_{i0}/K_{i0})^2 K_{i0}$, in which $a > 0$. In the two-period model, firms do not invest in date 1, $I_{i1} = 0$. The aggregate resource constraint in each date limits aggregate consumption, aggregate investment, and aggregate adjustment costs to aggregate resources:

$$C_0 + \sum_{i=1}^{N} I_{i0} + \sum_{i=1}^{N} \frac{a}{2} \left(\frac{I_{i0}}{K_{i0}} \right)^2 K_{i0} = \sum_{i=1}^{N} \Pi_{i0} K_{i0}, \tag{8}$$

$$C_1 = \sum_{i=1}^{N} \Pi_{i1} K_{i1}. \tag{9}$$

3.2.1 The Social Planner's Problem

The social planner chooses a distribution of capital to maximize the utility in equation (7) subject to the resource constraints (8) and (9). Capital is the only means to transfer consumption between dates and to diversify risk across firms. Optimal allocation of capital requires that:

$$1 + a \frac{I_{i0}}{K_{i0}} = E_0 \left[M_1 \Pi_{i1} \right], \tag{10}$$

in which $M_1 \equiv \rho U'(C_1)/U'(C_0)$. The left-hand side is the marginal cost of investing, and the right-hand side is the marginal benefit of investment, or marginal q. To generate an extra unit of capital at the beginning of date 1, the social planner must pay the price of capital (unity), and the marginal adjustment cost, $a(I_{i0}/K_{i0})$. The marginal benefit of this extra unit of capital over date 1 is the marginal product of capital, Π_{i1}. Discounting the date 1's marginal benefit using the representative consumer's intertemporal marginal rate of substitution yields the marginal q.

Craine (1989) shows how risk affects the optimal allocation of capital across firms in this economy. Rewrite the right-hand side of equation (10) as $E_0 [M_1] E_0 [\Pi_{i1}] + \mathrm{Cov}_0 [M_1, \Pi_{i1}]$, in which $\mathrm{Cov}_0 [M_1, \Pi_{i1}]$ is the covariance between M_1 and Π_{i1}. If firm i is risky, Π_{i1} will have a high covariance with the representative consumer's consumption growth from date 0 to date 1, and the (negative) covariance term, $\mathrm{Cov}_0 [M_1, \Pi_{i1}]$, will have a large magnitude. Optimal condition (10) then implies that the social planner will invest less in firm i's capital because of its high risk.

3.2.2 The Decentralized Economy

In the decentralized economy, the representative household's optimal consumption-portfolio choice problem is standard. Let P_{it} and D_{it} be the ex dividend equity value and dividend of firm i at date $t = 0, 1$, respectively. The consumption first-order condition says that:

$$P_{i0} = E_0[M_1(P_{i1} + D_{i1})] \qquad \Rightarrow \qquad E_0[M_1 r_{i1}^S] = 1, \tag{11}$$

in which the stock return is defined as $r_{i1}^S \equiv (P_{i1} + D_{i1})/P_{i0}$. Using the definition of covariance, we can rewrite $E_0[M_1 r_1^S] = 1$ in the beta-pricing form (see, e.g., Cochrane (2005, p. 16)):

$$E_0[r_{i1}^S] - r_f = \beta_i^M \lambda_M, \tag{12}$$

in which $r_f \equiv 1/E_0[M_1]$ is the risk-free rate, $\beta_i^M \equiv -\mathrm{Cov}(r_{i1}^S, M_1)/\mathrm{Var}(M_1)$ is the sensitivity of the stock return with respect to M_1, and $\lambda_M \equiv \mathrm{Var}(M_1)/E_0[M_1]$ is the price of risk. Under more restrictive assumptions such as quadratic utility or exponential utility with normal distributions of returns, the beta-pricing equation (12) can be reduced further to the Sharpe (1964) and Lintner (1965) CAPM (see, e.g., Cochrane (p. 152–155) for detailed derivations).

In the production side, firm i uses the operating cash flow at date 0 to pay investment, I_{i0}, and adjustment costs, $(a/2)(I_{i0}/K_{i0})^2 K_{i0}$. If the free cash flow, $D_{i0} = \Pi_{i0}K_{i0} - I_{i0} - (a/2)(I_{i0}/K_{i0})^2 K_{i0}$, is positive, firm i distributes it back to the household. A negative D_{i0} means external equity. At date 1, the firm uses capital K_{i1} to obtain the operating profits, $\Pi_{i1}K_{i1}$, which are distributed as dividends, D_{i1}. The ex dividend equity value, P_{i1}, is zero.

Taking M_1 as given, firm i chooses investment to maximize date 0's cum dividend equity value:

$$P_{i0} + D_{i0} \equiv \max_{\{I_{i0}\}} \left[\Pi_{i0}K_{i0} - I_{i0} - \frac{a}{2}\left(\frac{I_{i0}}{K_{i0}}\right)^2 K_{i0} + E_0\left[M_1 \Pi_{i1} K_{i1}\right] \right]. \tag{13}$$

The investment first-order condition is identical to equation (10), which characterizes the social planner's solution. Following Cochrane (1991), we define the investment return as the ratio of the date 1's marginal benefit of investment divided by the date 0's marginal cost of investment:

$$r_{i1}^{I} = \frac{\Pi_{i1}}{1 + a(I_{i0}/K_{i0})}. \tag{14}$$

Equation (10) implies that the expected discounted investment return equals one, $E_0[M_1 r_{i1}^I] = 1$.

Because $D_{i0} = \Pi_{i0}K_{i0} - I_{i0} - (a/2)(I_{i0}/K_{i0})^2 K_{i0}$, the cum dividend equity value in equation (13) means that the ex dividend equity value is given by:

$$P_{i0} = E_0[M_1 \Pi_{i1} K_{i1}], \tag{15}$$

and the stock return, r_{i1}^S, is given by:

$$r_{i1}^{S} = \frac{P_{i1} + D_{i1}}{P_{i0}} = \frac{\Pi_{i1}K_{i1}}{E_0[M_1 \Pi_{i1} K_{i1}]}. \tag{16}$$

Following Cochrane (1991), we derive the stock return purely in terms of characteristics. Dividing both the numerator and the denominator of r_{i1}^S by K_{i1} and using equation (10), we obtain:

$$r_{i1}^{S} = \frac{\Pi_{i1}}{1 + a(I_{i0}/K_{i0})} = r_{i1}^{I}. \tag{17}$$

Intuitively, firm i keeps investing until the marginal cost of investment at date 0, $1 + a(I_{i0}/K_{i0})$ (increasing in I_{i0}/K_{i0}), is equated to the marginal benefit of investment at date 1, Π_{i1}, discounted back to date 0 with the discount rate given by the stock return. Alternatively, the discount rate can be backed out from the ratio of the marginal benefit of investment at date 1 divided by the marginal cost of investment at date 0. In effect, equation (17) is the two-period producer-based counterpart to the two-period consumer-based Sharpe (1964) and Lintner (1965) CAPM. This dual relation is similar to the dual relation between $r = 1/\left[\rho \frac{U'(C_1)}{U'(C_0)}\right]$ and $r = f'(I_0)$ in the Fisherian economy.

Incorporating corporate taxes and debt financing into the baseline model, Liu, Whited, and Zhang (2009) show that the investment return equals the weighted average cost of capital (WACC):

$$r_{it+1}^{I} = w_{it}\, r_{it+1}^{Ba} + (1 - w_{it})\, r_{it+1}^{S}, \tag{18}$$

in which w_{it} is market leverage, and r^{Ba}_{it+1} is the after-tax corporate bond return. Combining equations (14) and (18) yields:

$$w_{it}\, r^{Ba}_{it+1} + (1 - w_{it})\, r^{S}_{it+1} = \frac{\Pi_{i1}}{1 + a(I_{i0}/K_{i0})}. \tag{19}$$

Equation (19) provides *the microeconomic foundation* for the standard WACC approach to capital budgeting (e.g., Berk and Demarzo (2011, Chapter 18)). The WACC approach instructs managers to invest in all projects with positive net present values, so that at the margin the last infinitesimal project has a zero net present value. In our economy, the present value of the marginal project is $\Pi_{i1}/[w_{it}\, r^{Ba}_{it+1} + (1 - w_{it})\, r^{S}_{it+1}]$, which uses the WACC as the discount rate. The cost of the marginal project is $1 + a(I_{i0}/K_{i0})$. Optimal investment means that $1 + a(I_{i0}/K_{i0}) = \Pi_{i1}/[w_{it}\, r^{Ba}_{it+1} + (1 - w_{it})\, r^{S}_{it+1}]$, i.e., the net present value for the marginal project is zero.

By connecting stock returns to anomaly variables, the investment return equation provides a powerful framework for the cross section of expected returns. In particular, equation (17) says that: (i) Given expected profitability, high investment-to-capital firms should earn lower expected returns than low investment-to-capital firms; and (ii) given investment-to-capital, firms with high expected profitability should earn higher expected returns than firms with low expected profitability.[2]

The investment first-order condition is initially given by equation (10). Because it depends on M_1, the analysis on the discount rate-investment relation is complicated by the empirical difficulties of M_1. However, we build on Cochrane (1991) to derive an equivalent representation of the optimality condition as in equation (19), with the natural interpretation as the WACC approach to capital budgeting. As such, equation (19) is not just an ancillary technical condition: It is *the* investment first-order condition. Although M_1 does not enter equation (19) explicitly, its impact

---

[2] Implementing equation (17) with factor regressions, Chen, Novy-Marx, and Zhang (2011) show that a new three-factor model consisting of the market factor, an investment factor, and a profitability factor is a good start to describing the cross section of returns. The new factor model reduces the magnitude of the abnormal returns to a wide range of anomalies-based trading strategies, often to insignificance. Liu, Whited, and Zhang (2009) implement the weighted average cost of capital equation (19) via structural estimation by testing whether the expected stock return equals the expected levered investment return via generalized method of moments. Empirically, the equation fits well the average returns across testing portfolios formed on earnings surprises, book-to-market, and capital investment.

is captured indirectly. A higher amount of risk will increase the WACC on average, all else equal, equation (19) then says that firms will respond by lowering investment. This risk analysis is exactly the same as that articulated in Craine (1989). However, to the extent that equation (19) does not depend on M_1, at least explicitly, and is thus immune to measurement errors in consumption data and specification errors in utility functions, equation (19) is more useful in practice.

4 Questioning the Doctrine

The investment approach questions the deep-ingrained doctrine that if an anomaly is consistent with investor “rationality,” it must be “explained” by a risk factor model.

We proceed as follows. First, differing from Fama and French (1993, 1996), we argue that characteristics-based factors are not necessarily risk factors. Instead, characteristics-based factor models are linear approximations to firm-level investment returns (Section 4.1). Second, differing from Daniel and Titman (1997), we argue that the evidence that characteristics dominate covariances in horse races does not necessarily mean mispricing. Instead, measurement errors in covariances are likely to blame (Sections 4.2 and 4.3). Third, we articulate the (lack of) causal relations among covariances, expected returns, and characteristics in general equilibrium (Section 4.4). Finally, we address several common misconceptions on the investment approach (Section 4.5).

4.1 Are Characteristics-Based Factors Risk Factors in ICAPM or APT?

Brock (1982) shows how to derive Merton’s (1973) ICAPM and Ross’s (1976) APT within the Arrow-Debreu general equilibrium economy. Firm i faces a vector of F aggregate technological uncertainties, denoted $\tilde{X}_t^f$, for $f = 1, \ldots, F$. Brock models technological heterogeneity across firms as firm-specific loadings, denoted L_{it}^f, on the aggregate factors, $\tilde{X}_t^f$. Although time-varying, the loadings are not stochastic. The marginal product of capital, Π_{it}, is given by:

$$\Pi_{it} \equiv \sum_{f=1}^{F} L_{it}^f \tilde{X}_t^f. \tag{20}$$

In the Brock (1982) economy, there are no capital adjustment costs. All the other aspects of his economy are similar to what we have in Section 3.2. Equation (17) then implies that the stock return admits an F-factor structure:

$$r_{i1}^S = \sum_{f=1}^{F} L_{i1}^f \tilde{X}_1^f, \tag{21}$$

which inherits the F-factor structure in aggregate technological uncertainties. Even with capital adjustment costs, the stock return admits a similar F-factor structure. In equation (17), the investment-to-capital ratio, I_{i0}/K_{i0}, is decided at date 0 before the realization of $\tilde{X}_1^f$ at the beginning of date 1. As such, a linear F-factor structure continues to hold:

$$r_{i1}^S = \sum_{f=1}^{F} \frac{L_{i1}^f}{1 + a(I_{i0}/K_{i0})} \tilde{X}_1^f. \tag{22}$$

As such, for a factor to be a *risk* factor, it must be a source of *aggregate* uncertainty, $\tilde{X}_1^f$, affecting *all* firms in the economy. Examples include shocks to total factor productivity, government policy shocks, and aggregate demand shocks driven by changes to preferences or animal spirits. From this economic perspective, characteristics-based factors are *not* ICAPM or APT risk factors. *Firm-specific* characteristics, on which these factors are based, have *no* immediate linkages with aggregate sources of uncertainty affecting the fundamentals of all the firms in the economy *simultaneously*.

The investment approach provides a natural interpretation for characteristics-based factors. In particular, we interpret the Chen, Novy-Marx, and Zhang (2011) three-factor model, consisting of the market factor, an investment factor, and a profitability factor, as *the factor-based linear approximation to firm-level investment returns* in equation (17). The market factor aims to anchor a stock's average return to the market's average, motivated from the consumption side of the general equilibrium (see equation (12)). The investment and profitability factors aim to capture a stock's average return variation related to investment and profitability, as predicted by equation (17).

The Fama-French model can also be interpreted as an investment-based model. Because the marginal cost of investment equals marginal q, the denominator of the investment return in equation (17) is market-to-book. As such, it is no wonder why the value factor and the investment factor play a similar role in describing cross-sectional returns (e.g., Xing (2008)). Market capitalization does not appear directly in the investment return equation, above and beyond being the numerator of market-to-book. In this sense, the size factor is redundant to the value factor in the context of the investment return framework. This redundancy is probably the reason why the average size factor return has been insignificant, and why the size factor has not been effective in recent samples.

4.2 Characteristics Dominating Covariances = Mispricing?

We take three steps to demonstrate why the evidence that characteristics dominate covariances in predicting returns does not necessarily mean mispricing. First, the evidence that characteristics predict stocks returns is consistent with the investment return equation. Investment and profitability from the investment return equation combine to predict, on *economic* ground, many empirical relations between anomaly variables and subsequent stock returns (e.g., Liu, Whited, and Zhang (2009) and Chen, Novy-Marx, and Zhang (2011)). Although these relations are anomalies to the consumption approach, these are potentially *regularities* to the investment approach.

Second, the covariances-based consumption model and the characteristics-based investment model of expected returns are theoretically equivalent, as in the Fisherian economy. The two classes of models are not mutually exclusive in general equilibrium. On the contrary, covariances and characteristics are the two sides of the same coin, delivering *identical* expected stock returns. Their relation is neatly complementary. To see this point, combining equations (12) and (17), we obtain:

$$r_f + \beta_i^M \lambda_M = E_0[r_{i1}^S] = \frac{E_0[\Pi_{i1}]}{1 + a(I_{i0}/K_{i0})}. \tag{23}$$

Solving for β_i^M provides an analytical link between covariances and characteristics:

$$\beta_i^M = \left[\frac{E_0[\Pi_{i1}]}{1 + a(I_{i0}/K_{i0})} - r_f\right] / \lambda_M. \tag{24}$$

The covariances-based model, $E_0[r_{i1}^S] = r_f + \beta_i^M \lambda_M$, derived from the consumption first-order condition, connects risk premiums to covariances. It says that covariances are sufficient statistics of expected returns: Once covariances are controlled for, characteristics should not affect the cross section of expected returns. This prediction is the basic premise underlying traditional asset pricing. In contrast, the characteristics-based model, $E_0[r_{i1}^S] = \frac{E_0[\Pi_{i1}]+1-\delta}{1+a(I_{i0}/K_{i0})}$, derived from the investment first-order condition, connects expected returns to characteristics. It says that characteristics are sufficient statistics of expected returns: *Once characteristics are controlled for, covariances should not affect the cross section of expected returns!* This seemingly outrageous statement from the investment approach is in effect the dual counterpart to the equally "outrageous" statement that covariances are sufficient statistics of expected returns from the consumption approach. In general equilibrium, equation (23) shows that both approaches give the correct answer, as in Fisher (1930).

Third, if covariances and characteristics are conceptually equivalent "explanatory" variables of expected returns, why do characteristics dominate covariances in asset pricing tests? Rather than mispricing per Daniel and Titman (1997), we interpret this evidence as saying that *characteristics are more precisely measured than covariances in the data.* Horse races between covariances and characteristics only test whether covariances or characteristics are more correlated with future returns. However, the correlations depend on the magnitude of measurement and specification errors in the "explanatory" variables. The magnitude is likely to be substantially larger for covariances than for characteristics. We demonstrate this point quantitatively in the next subsection.

4.3 Why Do Characteristics Dominate Covariances Empirically?

We use the Zhang (2005) economy to study the Daniel and Titman (1997) horse races between covariances and characteristics. The bottomline is that even though the model admits a covariance structure, characteristics dominate covariances from the model's simulations.

4.3.1 The Economy

The economy is a simplified version of the Zhang (2005) model. The production function is given by:

$$\Pi_{it} = X_t Z_{it} K_{it}^{\alpha} - f, \tag{25}$$

in which Π_{it} is firm i's operating profits, K_{it} is capital, $0 < \alpha < 1$ is the curvature parameter, and $f > 0$ is the fixed cost of production. The aggregate productivity, X_t, has a stationary Markov transition function. Denoting $x_t \equiv \log X_t$, we assume that the transition function follows:

$$x_{t+1} = \bar{x}(1 - \rho_x) + \rho_x x_t + \sigma_x \mu_{t+1}, \tag{26}$$

in which μ_{t+1} is an independent and identically distributed (i.i.d.) standard normal shock. The firm-specific productivity for firm i, Z_{it}, has a transition function given by:

$$z_{it+1} = \rho_z z_{it} + \sigma_z \nu_{it+1}, \tag{27}$$

in which $z_{it+1} \equiv \log Z_{it+1}$, and ν_{it+1} is an i.i.d. standard normal shock. We assume that ν_{it+1} and ν_{jt+1} are uncorrelated for any $i \neq j$, and that μ_{t+1} and ν_{it+1} are uncorrelated for any i.

We parameterize the stochastic discount factor, denoted M_{t+1}, as follows:

$$\begin{aligned} M_{t+1} &= \eta \exp\left[\gamma_t (x_t - x_{t+1})\right], & (28) \\ \gamma_t &= \gamma_0 + \gamma_1 (x_t - \overline{x}), & (29) \end{aligned}$$

in which $0 < \eta < 1$, $\gamma_0 > 0$, and $\gamma_1 < 0$ are constant parameters. Upon observing X_t and Z_{it}, firm i chooses optimal investment, I_{it}, to maximize its market value of equity. Capital accumulates as:

$$K_{it+1} = I_{it} + (1 - \delta) K_{it}, \tag{30}$$

in which δ is the rate of depreciation. Capital investment entails asymmetric adjustment costs:

$$\Phi(I_{it}, K_{it}) = \begin{cases} a^{+}K_{it} + \frac{c^{+}}{2}\left(\frac{I_{it}}{K_{it}}\right)^2 K_{it} & \text{for} \quad I_{it} > 0 \\ 0 & \text{for} \quad I_{it} = 0 \\ a^{-}K_{it} + \frac{c^{-}}{2}\left(\frac{I_{it}}{K_{it}}\right)^2 K_{it} & \text{for} \quad I_{it} < 0 \end{cases} \tag{31}$$

in which $a^{-} > a^{+} > 0$, and $c^{-} > c^{+} > 0$ capture the nonconvexities of adjustment costs.

The cum-dividend market value of equity, $V_{it} \equiv V(K_{it}, X_t, Z_{it})$, is given by:

$$V(K_{it}, X_t, Z_{it}) = \max_{\{I_{it}\}} \left[\Pi_{it} - I_{it} - \Phi(I_{it}, K_{it}) + E_t\left[M_{t+1}V(K_{it+1}, X_{t+1}, Z_{it+1})\right]\right], \tag{32}$$

subject to equation (30). Evaluating the value function at the optimum yields $V_{it} = D_{it} + E_t[M_{t+1}V_{it+1}]$, with $D_{it} \equiv \Pi_{it} - I_{it} - \Phi(I_{it}, K_{it})$. Equivalently, $E_t[M_{t+1}r^S_{it+1}] = 1$, in which $r^S_{it+1} = V_{it+1}/(V_{it} - D_{it})$ is the stock return. Rewriting in the beta-pricing form yields: $E_t[r^S_{it+1}] = r_{ft} + \beta^M_{it}\lambda_{Mt}$, in which $r_{ft} = 1/E_t[M_{t+1}]$ is the real interest rate, $\beta^M_i \equiv -\text{Cov}_t[r^S_{it+1}, M_{t+1}]/\text{Var}_t[M_{t+1}]$ is the true conditional beta, and $\lambda_{Mt} \equiv \text{Var}_t[M_{t+1}]/E_t[M_{t+1}]$ is the price of risk.

The model is calibrated in monthly frequency. We calibrate $\eta = 0.9945$, $\gamma_0 = 50$, and $\gamma_1 = -1,000$ to match the average real interest rate, 1.40% per annum; the annualized volatility of the real interest rate, 2.48%; and the average Sharpe ratio, 0.40 per annum. We set the persistence of aggregate productivity $\rho_x = 0.95^{\frac{1}{3}}$ and conditional volatility $\sigma_x = 0.007/3$, which correspond to the quarterly values of 0.95 and $\sigma_x = 0.007$, respectively, in Cooley and Prescott (1995). For the remaining parameters, we set $\bar{x} = -3.65, \rho_z = 0.97$, $\sigma_z = 0.15$, $\alpha = 0.70$, $\delta = 0.01$, $f = 0.00328$, and the adjustment cost parameters: $a^{+} = 0.01, a^{-} = 0.03$, $c^{+} = 20$, and $c^{-} = 200$.

We iterate on the value function to solve the model (see Zhang (2005, Appendix B)). We use the model's solution to simulate 100 artificial samples, each of which has 5,000 firms and 1,000 monthly periods. We start the simulations by setting the initial capital stocks of all firms at their long-run average level and by drawing their firm-specific productivity levels from the unconditional distribution of Z_{it}. We drop the first 400 months to neutralize the impact of the initial condition.

The remaining 600 months of simulated data are treated as those from the economy's stationary distribution. The sample size is largely comparable to the merged CRSP/Compustat dataset. We implement empirical procedures on each artificial sample and report cross-simulation averaged results as model moments. We then compare the model moments with the moments in the real data.

With the calibrated parameters, the model produces an average investment-to-capital ratio (across firms and over time) of 0.12 per annum, a time series volatility of investment-to-capital (averaged across firms) of 12% per annum, and an average market-to-book ratio (across firms and over time) of 1.57. These basic quantity moments seem reasonable. The model also generates a value premium of 6.67% per annum, which is measured as the average return of the high-minus-low book-to-market decile constructed from the standard Fama-French (1993) portfolio approach.

4.3.2 Covariances versus Characteristics in the Data

Our goal is to examine the properties of the Daniel and Titman's (1997) covariances versus characteristics test. We first replicate their results using an updated sample. We obtain monthly stock returns from the Center for Research in Security Prices (CRSP) and accounting information from the CRSP/Compustat Merged Annual Industrial Files. The sample is from July 1973 to June 2010: The starting point is the same as in Daniel and Titman. We exclude from the sample any firm-year observation for which book equity is either zero or negative, and require firms to have ordinary common equity as classified by CRSP. We omit firms whose primary SIC classification is between 4900 and 4999 (regulated firms) or between 6000 and 6999 (financial firms). Finally, we require a firm to have at least two years of data in Compustat before including it in the sample.

In June of year t, we sort all NYSE/Amex and Nasdaq firms into quintiles based on the NYSE breakpoints of book-to-market at the last fiscal year end. Book-to-market is calculated with the book equity at the end of the fiscal year ending in calendar year $t-1$ and the market equity at the end of December of year $t-1$. The firms remain in these quintiles from July of year t to June of year $t+1$. The individual firms in each quintile are further sorted into one of five subportfolios based on their preformation HML loadings. The HML loadings are estimated from the Fama-French (1993) three-factor regressions performed for the period from month -42 to -7 relative to the formation date (June of year t). We then calculate the value-weighted returns for the resulting 25 portfolios for each month, and report their average returns in excess of the risk-free rate.

From Panel A of Table 1, book-to-market dominates the HML loading empirically in predicting future returns. The last row of Panel A shows that after we control for the characteristic, the average return spread between the low- and the high-HML loading quintiles is 0.29% per month. In contrast, the last column of Panel A shows that after we control for the HML loading, the average return spread between the low and the high book-to-market quintiles is 0.65%. As such, consistent with Daniel and Titman (1997), it is the characteristic rather than the covariance structure of returns that appears to "explain" the cross-sectional variation of stock returns.

4.3.3 Covariances versus Characteristics in the Model

To see whether the Daniel and Titman's (1997) evidence is as "disturbing" as claimed, we implement the exactly same procedure on simulated data from the model economy in Section 4.3.1. In Panel B of Table 1, we conduct their test on each of 100 artificial samples, and report cross-sample averaged results. The pattern in the model is largely consistent with the pattern in the data. After we control for book-to-market, the HML loading only generates a tiny average return spread of 0.02% per month between the extreme quitiles. In contrast, after we control for the HML loading, book-to-market produces a higher average return spread, 0.37%, across the extreme quintiles.

Because the HML loadings are *estimated* on each artificial sample, these loadings are noisy proxies for the true risk, β_i^M, in the model. To evaluate the impact of measurement errors in the estimated loadings, we repeat the analysis from Panel B but with the HML loadings replaced by β_i^M. The true beta can be calculated accurately within the model (see Zhang (2005, Appendix B)). From the last row in Panel C, after we control for the book-to-market characteristic, the true beta produces an average return spread of 0.41% per month across the extreme quintiles. From the last

Table 1 : Mean Monthly Percentage Excess Returns of the 25 Portfolios Formed on Book-to-Market and HML Loadings (in the Data and in the Model) and the 25 Portfolios Formed on Book-to-Market and True Conditional Betas (in the Model)

Monthly stock returns are from the Center for Research in Security Prices (CRSP) and accounting information is from the CRSP/Compustat Merged Annual Industrial Files. The sample is from July 1973 to June 2010. Following Fama and French (1993), we define book common equity as the Compustat book value of stockholders' equity, plus balance-sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation, or par value (in that order) to estimate the value of preferred stock. We exclude from the sample any firm-year observation for which book value of equity is either zero or negative. We require firms with ordinary common equity as classified by CRSP, and also omit firms whose primary SIC classification is between 4900 and 4999 (regulated firms) or between 6000 and 6999 (financial firms). A firm must have at least two years of data in Compustat before being included in the sample. We first rank all NYSE firms by their book-to-market at the end of year $t-1$ and form 20%, 40%, 60% and 80% breakpoints based on these rankings. Starting in July of year t, we place all NYSE/Amex and Nasdaq firms into the five book-to-market groups based on these breakpoints. The firms remain in these portfolios from July of year t to June of year $t+1$. The individual firms in each of these five portfolios are further sorted into five subportfolios based on their HML loadings. The HML loadings are from the Fama-French three-factor regressions performed between 42 months and 6 months prior to the formation date. The data for the three Fama-French factors and the risk-free rate are from Kenneth French's Web site. We then calculate the value-weighted returns for the 25 book-to-market and HML loadings portfolios for each month, and report the average returns in excess of the risk-free rate in Panel A. In Panel B, we simulate 100 artificial panels from the model economy in Section 4.3. Each panel contains 5,000 firms and 1,000 monthly observations. We drop the first 400 months to neutralize the impact of initial conditions in the simulations. We then replicate the empirical procedures in Panel A on each of the artificial panels, and report the cross-sample averaged results in Panel B. In Panel C, we report the simulated results from the empirical procedures similar to those in Panel B, except that we replace HML loadings with true conditional betas, β_i^M, on each artificial panel.

	Panel A: Data, HML loadings						Panel B: Model, HML loadings						Panel C: Model, true betas					
	Low	2	3	4	High	Ave.	Low	2	3	4	High	Ave.	Low	2	3	4	High	Ave.
Low	0.32	0.25	0.37	0.44	0.53	0.38	0.80	0.79	0.79	.80	0.81	0.80	0.70	0.77	0.81	0.86	0.96	0.82
2	0.49	0.41	0.72	0.65	0.86	0.63	0.87	0.85	0.86	0.86	0.88	0.87	0.76	0.83	0.88	0.93	1.05	0.89
3	0.76	0.59	0.54	0.83	0.94	0.73	0.93	0.92	0.92	0.92	0.94	0.93	0.82	0.89	0.94	1.00	1.12	0.95
4	0.51	0.62	0.63	0.78	0.87	0.68	1.01	0.99	0.99	1.00	1.03	1.00	0.85	0.96	1.03	1.11	1.27	1.04
High	0.98	1.02	0.97	0.87	1.31	1.03	1.18	1.14	1.14	1.16	1.23	1.17	0.94	1.08	1.20	1.36	1.70	1.25
Ave.	0.61	0.58	0.65	0.71	0.90		0.96	0.94	0.94	0.95	0.98		0.81	0.91	0.97	1.05	1.22	

column, even after we control for the true beta, book-to-market still generates an average return spread of 0.43% across the extreme quintiles. The key message from the table is that measurement errors in the empirical proxies of covariances make characteristics dominate covariances in the Daniel and Titman (1997) tests, even in a model embedded with a dynamic covariance structure.[3]

4.3.4 Where Do Measurement Errors in Covariances Come From?

More generally, covariances have severe sources of measurement and specification errors. First, any M mis-specifications translate directly into specification errors in β_i^M. Second, even if we can correctly specify M, the time-variation in betas in a dynamic model can induce large measurement errors in regression-based proxies for β_i^M. In particular, the beta estimates from 36-month rolling-window regressions in Daniel and Titman (1997) are unconditional betas in the past three-year period. This time-lag between the estimated betas and portfolio formation date reduces the ability of the estimated betas to predict returns. The conditioning approach of Ferson and Harvey (1991) and Jagannathan and Wang (1996) uses up-to-date information in estimating conditional betas. However, in theory both aggregate and firm-specific variables should enter the beta specification nonlinearly. As a result, linear models contain large specification errors (e.g., Ghysels (1998)). Third, even industry costs of capital from unconditional factor models are extremely imprecise because of noisy estimates of factor loadings and risk premiums (e.g., Fama and French (1997)).

The point on measurement errors in covariances is not new. Miller and Scholes (1972) use randomly generated returns constructed to satisfy the CAPM, but find that the results from simulated asset pricing tests are virtually indistinguishable from those in the real data. Gomes, Kogan, and Zhang (2003) show that in a single-factor model size and book-to-market are correlated with the true beta and appear to predict stock returns. The cross-sectional relations between characteristics

---

[3] There are several reasons why book-to-market retains strong predictive power in the model's simulations, even after we control for the true beta. First, the true beta and book-to-market are positively correlated in the model, but the multivariate sort from Daniel and Titman (1997) fails to disentangle their effects on expected returns. Second, the price of risk, λ_{Mt}, is time-varying in the model, but the sort fails to capture this time-variation. Finally, Daniel and Titman implement the multivariate sort sequentially (not independently), first on characteristics then on covariances. This sequential order is likely to bias the results in favor of characteristics.

and returns can subsist even after one controls for empirical estimates of beta. Li, Livdan, and Zhang (2008) show that even if one can measure betas perfectly, linear cross-sectional regressions are misspecified in testing conditional asset pricing models because the price of risk is time-varying.

Because of these measurement errors, costs of capital from the consumption approach are extremely unreliable. Accounting academics, who face the daily task of teaching students to value real companies, have long been disillusioned. Lundholm and Sloan (2007, p. 193) lament:

> "None of the standard finance models provide estimates that describe the actual data very well. The discount rate that you use in your valuation has a large impact on the result, yet you will rarely feel very confident that the rate you have assumed in the right one. The best we can hope for is a good understanding of what the cost of capital represents and some ballpark range for what a reasonable estimate might be."

Penman (2010, p. 666) also expresses extreme pessimism:

> "Compound the error in beta and the error in the risk premium and you have a considerable problem. The CAPM, even if true, is quite imprecise when applied. Let's be honest with ourselves: No one knows what the market risk premium is. And adopting multifactor pricing models adds more risk premiums and betas to estimate. These models contain a strong element of smoke and mirrors."

4.4 Are There Causal Relations Among Covariances, Expected Returns, and Characteristics in General Equilibrium?

In this subsection, we clarify how the characteristics-based investment approach fits, conceptually, with the covariances-based consumption approach to asset pricing.

Asset pricing often works directly with $E[Mr^S] = 1$. However, both M and r^S are endogenous in general equilibrium. There are three possibilities how one can provide *technological underpinnings* beneath M and r^S: (i) the Lucas tree economy, in which technologies allow no substitution of consumption across time or states of nature; (ii) the linear technologies economy, in which investment can be used freely to smooth consumption; and (iii) the adjustment costs economy, in which one must pay adjustment costs of capital when using investment to smooth consumption. We study causality among covariances, expected returns, and characteristics in each economy. The bottomline is that the doctrine that covariances determine expected returns is valid only in the Lucas economy, but not in general equilibrium economies more broadly.

4.4.1 The Lucas Tree Economy

In this economy, the dividends of firms (trees) are exogenously fixed across dates and states of nature. As such, $E[r^S] - r_f = -r_f \text{Cov}(M, r^S)$, which is equivalent to $E[Mr^S] = 1$, implies that covariances *determine* expected returns. To see why, once the dividends of all the firms in the economy are given, M is pinned down by a utility function and aggregate consumption as the sum of all firms' dividends. The covariance of a firm's dividends (and, loosely, stock returns) with M is subsequently pinned down as well. As such, causality runs *from covariances to expected returns*. Once $\text{Cov}(M, r^S)$ is specified (almost) exogenously, we know $E[r^S]$. This causal relation is why Daniel and Titman (1997, p. 4) would write: "In equilibrium asset pricing models the covariance structure of returns *determines* expected returns (our emphasis)."

4.4.2 The Linear Technologies Economy

This linear technologies economy is the polar extreme to the Lucas tree economy. Famous examples include Long and Plosser (1983) and Cox, Ingersoll, and Ross (1985). In this economy, the dividends of firms arise endogenously as the outcome of firms' optimal production and investment decisions in maximizing the market value of equity. However, firms do not incur capital adjustment costs when investing. Using our notation in Section 3.2, equation (17) says that the stock return is given by $r^S_{i1} = \Pi_{i1} + 1 - \delta$, in which Π_{i1} is the stochastic productivity (with accounting profitability as a common proxy). Once Π_{i1} is specified exogenously, we know r^S_{i1}. As such, in this economy, the causality runs *from characteristics (profitability) to expected returns and then to covariances* via

$E[r^S] - r_f = -r_f \text{Cov}(M, r^S)$! This causality goes in the opposite direction to the basic premise of Daniel and Titman (1997) and Davis, Fama, and French (2000).

4.4.3 The Adjustment Costs Economy

The adjustment costs economy differs from the linear technologies economy in that firms incur adjustment costs when investing. The adjustment costs economy lies somewhere in between the Lucas tree economy and the linear technologies economy, and paints a more realistic picture of the world.

The Arrow-Debreu economy from Section 3.2 provides such an example. As noted, the consumption first-order condition $E_0[M_1 r^S_{i1}] = 1$ and the investment first-order condition $r^S_{i1} = \frac{\Pi_{i1}+1-\delta}{1+a(I_{i0}/K_{i0})}$ are two key equilibrium conditions. With adjustment costs, $a > 0$, the stock return is endogenous, and is no longer determined exogenously by the marginal product of capital, Π_{i1}. In general, *consumption (covariances), expected returns, and investment (characteristics) are all endogenous variables determined by a system of simultaneous optimality conditions in general equilibrium.* No causality whatsoever runs across these endogenous variables. We can only make inferences about structural *correlations* between covariances, expected returns, and characteristics from equilibrium conditions. Neither consumption nor investment says anything about causal forces driving expected returns. Neither covariances nor characteristics are more primitive in "explaining" expected returns.

Although the word "explain" has unfortunately crept into our writings on occasion, we do not claim that the investment approach "explains" asset pricing anomalies. In particular, we have never claimed (meant to claim) that characteristics *cause* expected returns to vary across firms. The investment approach is based on the investment first-order condition, which does not establish causality from investment to expected returns. The logic is as consistent with the view that investment "explains" expected returns as with the view that expected returns "explain" investment.

However, the investment approach is *no more and no less* important than the consumption approach (or a risk factor model) in "explaining" asset pricing anomalies. Suppose a utility function and some consumption data (or several risk factors as a linear specification of M) were found such that $E[Mr^S] = 1$ holds across portfolios sorted on an anomaly variable. A researcher will be tempted to claim that such an M model "explain" the anomaly in question. However, this would-be finding does *not* support such a claim. The consumption first-order condition says that consumers adjust consumption correctly in response to stock price movements. If equilibrium stock prices move arbitrarily with Mars synodic cycles, $E[Mr^S] = 1$ just aligns consumption with the stock prices correctly. Consumption is as endogenous to the consumption first-order condition as investment is to the investment first-order condition.

Alas, because of a half-century indoctrination of the CAPM, reinforced by the economic logic that is specific only to the Lucas tree economy, it is commonly believed that covariances are primitive "explanatory" variables of expected returns. The presentation flow of standard corporate finance texts also gives the false impression that covariances are "causal" forces driving expected returns. For example, Berk and DeMarzo (2011) teach students to estimate costs of equity and debt from the CAPM in Chapter 12, before describing how to use the estimated WACC to calculate the net present values of new projects in Chapter 18. The "causality" seems to flow from covariances to WACC and then to investment. As noted, this "causality" is an *illusion*, which does not hold in the adjustment costs economy. In fact, the exact reverse causality is true in the linear technologies economy.[4]

The investment approach thinks about asset pricing differently from the consumption approach, but in a complementary way. The consumption approach connects unobservable and hard-to-measure expected returns to equally unobservable and hard-to-measure covariances. The investment approach connects expected returns to observable and easier-to-measure characteristics. As noted, the standard practice in finance is to estimate the discount rate from the CAPM or multifactor models, and subsequently use the estimated discount rate in capital budgeting. The investment

---

[4] Cochrane (2005, p. 41) also writes: "We routinely think of betas and factor risk prices — components of $E(mx)$ — as *determining* expected returns. For example, we routinely say things like 'the expected return of a stock increased *because* the firm took on riskier projects, thereby increasing its beta.' But the whole consumption process, discount factor, and factor risk premia change when the production technology changes. Similarly, we are on thin ice if we say anything about the effects of policy interventions, new markets and so on. The equilibrium consumption or asset return process one has modeled statistically may change in response to such changes in structure. For such questions one really needs to start thinking in general equilibrium terms (original emphasis)."

approach turns finance on its head: Because the levered investment return equals the WACC, we can *back out* the cost of equity from the levered investment return estimated from observable characteristics. The big news is that covariances are only one half of asset pricing, with characteristics being the other half! The investment approach *completes* the consumption approach in general equilibrium!

4.5 Misconceptions

In this subsection, we address several common misconceptions about the investment approach. Critiques are not unexpected. For example, Kuhn (1962, p. 109) says:

> "In learning a paradigm the scientist acquires theory, methods, and standards together, usually in an inextricable mixture. Therefore, when paradigms change, there are usually significant shifts in the criteria determining the legitimacy both of problems and of proposed solutions."

Alas, Kuhn also warns that:

> "To the extent, as significant as it is incomplete, that two scientific schools disagree about what is a problem and what a solution, they will inevitably talk through each other when debating the relative merits of their respective paradigms."

4.5.1 Does the Investment Approach "Explain" Anomalies?

Critics argue that in modern asset pricing, the cross section of risk premiums are "determined" by exposures to risk factors times the factor risk premiums. A rational "explanation" for anomalies should account for why: (i) Anomaly variables are correlated with future returns, so sorting on these variables gives rise to portfolios with different subsequent returns on average; (ii) there seems to be a common factor related to a given anomaly variable, and this common return variation suggests that extreme portfolios have different exposures to unknown sources of systematic risk; and (iii) extreme portfolios have similar market betas and consumption betas, suggesting that the CAPM and the consumption CAPM fail to "explain" the average return spread across the extreme portfolios.

Before addressing these points, we first note that the investment approach is as primitive as the consumption approach in interpreting expected returns (see Section 4.4). As such, exposures to risk factors and factor risk premiums do not "determine" or "explain" expected returns. This causal language should not be used to characterize the consumption approach or the investment approach.

To address point (i), the investment return equation predicts the correlations between a wide array of anomaly variables and subsequent stock returns (e.g., Liu, Whited, and Zhang (2009) and Chen, Novy-Marx, and Zhang (2011)). To address point (ii), there is nothing mysterious about common variations in characteristics-based factors (see Section 4.1). The common variations are not some inexplicable sources of systematic risk. As components of firm-level investment returns, characteristics are *supposed* to predict subsequent returns in cross-sectional regressions.

Also, time series and cross-sectional regressions are largely equivalent ways of summarizing empirical correlations in the data. If a characteristic shows up significant in cross-sectional regressions, its factor mimicking portfolio is likely to show "explanatory" power in time series factor regressions. If a factor loading shows up significant in time series regressions, its underlying characteristic is likely significant in cross-sectional regressions. Factor loadings (covariances) are no more primitive than characteristics, and characteristics are no more primitive than factor loadings in "explaining" expected returns. And to the extent that characteristics are more precisely measured than covariances (Section 4.3), characteristics *should* be more useful in forecasting returns in practice.

Point (iii) emphasizes the importance of matching the stylized fact that extreme anomalies-based testing portfolios have similar market betas and consumption betas. This critique speaks to Zhang (2005, Table III), who reports a low market beta of 0.14 for the Fama-French's HML factor in the 1927–2001 sample, but a high market beta of 0.43 for the HML factor simulated from the model.[5] We view this critique as valid and potentially important.

Panel A of Table 2 shows why the critique is valid. The panel reports descriptive statistics of

---

[5] Belo and Lin (2011) raise a similar point that calibrated investment models à la Zhang (2005) fail to reproduce quantitatively the failure of the CAPM, and more so the failure of the Fama-French model in factor regressions.

ten book-to-market deciles from 1965 to 2010. The portfolio data are from Kenneth French's Web site. In this post-Compustat sample, the value premium — the average return of the high-minus-low decile — is 0.55% per month, which is more than 2.7 standard errors from zero. The CAPM alpha of the high-minus-low decile is 0.56% ($t = 2.36$). More important, the market beta for the zero-investment portfolio is zero: The extreme deciles have "similar" market betas around 1.06.

While the evidence in Panel A is important, we caution against taking it literally and ignoring the sampling variation in realized returns (e.g., Merton (1980)). Panel B of Table 2 repeats the same analysis as in Panel A but for the full sample from January 1927 to December 2010. With 38 years of more data, Panel B shows that the value premium remains at 0.53% per month, which is close to 0.55% from the shorter sample. However, the CAPM alpha for the high-minus-low decile is insignificant, 0.25% ($t = 1.20$), and its market beta is 0.45, which is more than three standard errors from zero (see also Ang and Chen (2007)). The adjusted R^2 of the CAPM regression of the high-minus-low decile goes up from zero in the post-Compustat sample to 0.14 in the full sample.

What does the investment model in Section 4.3.1 have to say about the properties of ten book-to-market deciles? In Panel C of Table 2, we repeat the analysis in Panels A and B but on artificial samples simulated from the model. On each artificial sample, we follow the exact timing convention of Fama and French (1993) and sort stocks into deciles in June of each year t on book-to-market, defined as K_{it}/P_{it} in the model, at the end of December of year $t-1$. We value-weight the portfolio returns from July of year t to June of year $t+1$, and rebalance the portfolios in June. We conduct the CAPM regressions on each artificial sample, and report the across-simulation averaged returns.

Panel C of Table 2 shows that the model matches the average value premium at 0.56% per month. The CAPM alpha of the high-minus-low decile is 0.12% ($t = 1.25$), which is not far from the 0.25% estimate from the 1927–2010 sample. The CAPM beta is 0.49, which is also close to 0.45 in the longer sample. However, the model *fails* in two aspects. First, although the CAPM alpha of 0.56% for the high-minus-low decile from the 1965–2010 sample lies within the 95% confidence inter-

Table 2 : Properties of the Book-to-Market Equity Deciles (in the Data and in the Model)

The returns data on the book-to-market deciles, the market factor, and the risk-free rate are from Kenneth French's Web site. For each decile and the high-minus-low portfolio, H−L, we report the mean percentage excess returns (Mean), stock return volatility in monthly percent (Std), as well as the CAPM regressions including α, β, their t-statistics adjusted for heteroscedasticity and autocorrelations, and adjusted R^2. We report these statistics using the real data across two samples, the post-Compustat sample from 1965 to 2010 (Panel A) and the full sample from 1927 to 2010 (Panel B). In Panel C, we simulate 100 artificial panels from the model economy in Section 4.3. Each panel contains 5,000 firms and 1,000 monthly observations. We drop the first 400 months to neutralize the impact of initial conditions in the simulations. We then follow the Fama and French (1993) portfolio approach in constructing ten book-to-market deciles on each of the artificial panels, perform the CAPM regressions, and report the cross-sample averaged results. Panel C also reports the 2.5 and 97.5 percentiles (%) of α and β across the 100 simulations.

	Low	2	3	4	5	6	7	8	9	High	H−L
Panel A: January 1965–December 2010											
Mean	0.33	0.44	0.48	0.48	0.46	0.55	0.60	0.65	0.74	0.88	0.55
Std	5.30	4.85	4.77	4.89	4.58	4.62	4.53	4.65	4.90	5.99	4.77
α	−0.13	0.01	0.06	0.06	0.08	0.16	0.23	0.27	0.35	0.43	0.56
t_α	−1.19	0.12	0.98	0.54	0.73	1.60	2.04	2.10	3.15	2.86	2.36
β	1.07	1.01	0.98	0.99	0.91	0.93	0.87	0.88	0.93	1.06	0.00
t_β	33.40	35.01	25.66	24.83	24.35	26.51	19.59	15.26	17.35	12.55	−0.03
adj. R^2	0.86	0.92	0.90	0.87	0.83	0.85	0.77	0.76	0.75	0.67	0.00
Panel B: January 1927–December 2010											
Mean	0.55	0.65	0.64	0.63	0.71	0.74	0.75	0.91	0.97	1.08	0.53
Std	5.79	5.54	5.36	6.12	5.70	6.24	6.71	7.04	7.64	9.46	6.69
α	−0.07	0.05	0.05	−0.03	0.10	0.08	0.05	0.19	0.20	0.18	0.25
t_α	−1.00	0.93	1.07	−0.42	1.28	0.88	0.56	1.82	1.86	1.10	1.20
β	1.00	0.98	0.94	1.06	0.98	1.07	1.12	1.16	1.24	1.45	0.45
t_β	37.54	35.05	29.09	18.90	21.08	15.02	12.26	10.57	14.05	11.86	3.08
adj. R^2	0.90	0.93	0.92	0.90	0.88	0.88	0.83	0.81	0.79	0.71	0.14
Panel C: Model simulations											
Mean	0.77	0.81	0.86	0.86	0.93	0.92	0.99	1.02	1.11	1.32	0.56
Std	6.79	7.16	7.58	7.62	8.18	8.15	8.60	8.95	9.50	10.89	4.37
α	−0.01	−0.01	−0.01	−0.01	0.00	0.00	0.01	0.01	0.04	0.11	0.12
t_α	−0.53	−0.44	−0.38	−0.35	−0.16	−0.25	0.35	0.19	0.86	1.32	1.25
α, 2.5%	−0.07	−0.08	−0.05	−0.07	−0.04	−0.03	−0.03	−0.03	−0.03	−0.03	−0.05
α, 97.5%	0.02	0.03	0.02	0.03	0.05	0.03	0.06	0.09	0.23	0.56	0.65
β	0.86	0.90	0.95	0.96	1.03	1.02	1.08	1.12	1.19	1.35	0.49
t_β	−0.53	−0.44	−0.38	−0.35	−0.16	−0.25	0.35	0.19	0.86	1.32	1.25
β, 2.5%	0.82	0.87	0.93	0.93	1.00	1.00	1.05	1.06	1.11	1.16	0.27
β, 97.5%	0.89	0.93	0.97	0.98	1.06	1.06	1.12	1.19	1.31	1.58	0.77
adj. R^2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.98	0.85

val of alpha from the model, $[-0.05\%, 0.65\%]$, the market beta of zero does not. The 2.5% percentile for the market beta is 0.27 in the model, and is still substantially higher than zero. Second, the adjusted R^2 from the CAPM regression of the high-minus-low decile is too high in the model, 85%. In contrast, the adjusted R^2 is 14% in the 1927–2010 sample, and is zero in the 1965–2010 sample.

We consider the failure of the CAPM in the post-Compustat sample to be an important stylized fact. As such, we suggest that future work should incorporate additional aggregate shocks to replicate the failure of the CAPM in simulations (e.g., Kogan and Papanikolaou (2011)). The critique is valid in the context of Zhang (2005) precisely because this class of models parameterizes M with a dynamic single factor structure. The failure of the CAPM means that M is more complex in the data. In addition, multiple firm-specific shocks should be incorporated to lower the adjusted R^2 from the CAPM regressions of zero-investment portfolios in simulations.

However, this critique does not apply, at least directly, to the investment return framework in equation (17). The investment return equation does not take a stand on M! The framework is *general* as it is immune to specification errors in M and measurement errors in consumption data.

This generality cannot be overemphsized. Although Cochrane (1991) was published a decade prior to the first wave of investment-based studies on the cross section of returns, these studies opt to construct fully specified and explicitly solved models (e.g., Gomes, Kogan, and Zhang (2003) and Zhang (2005)). Alas, the economic mechanisms are often special to their specific assumptions, yet these models are impossible to implement empirically. Their computational complexity makes direct estimation infeasible. An ancillary yet wrong specification can easily render a model rejected, even the model delivers economically interesting and empirically relevant insights. To make progress, it is critical to break down an intractable problem into several tractable subproblems. While clear about the limitations from the absence of M, the investment approach provides an empirically tractable way to tackle many issues on cross-sectional returns. No one has ever said M is not important, but M does not have to be imposed mechanically on every paper.

4.5.2 Does the Investment Approach Say Anything about Return Predictability?

Critics also argue that return predictability is identical to time-varying risk premiums in a rational model. Because the investment approach does not model risk premiums, it has nothing to say about return predictability. As such, the investment approach does not do asset pricing: It only tests the neoclassical investment theory in first differences rather than in levels.

This critique is nonsense. Because the risk-free rate is not very predictable, stock return predictability is largely identical to time-varying risk premiums. At the aggregate level, Cochrane (1991) is the classical work that shows aggregate investment forecasts aggregate stock market returns with a negative slope. In the cross section, because *the risk-free rate does not vary across firms*, the cross-sectional variation of expected returns *is* the cross-sectional variation of risk premiums. The high-minus-low expected return *is* the high-minus-low risk premium. Modeling the cross section of expected returns via the investment approach *is* modeling the cross section of risk premiums. Recall the lesson from Fisher (1930)! Dismissing the investment approach as neoclassical theory of investment in first differences is as silly as dismissing the consumption approach as permanent income theory of consumption, not asset pricing.[6]

4.5.3 What Does the Investment Approach Say about Rationality or Irrationality?

The investment approach is motivated as an alternative to investor "irrationality" as an interpretation of anomalies.[7] Critics argue that because investors are abstracted away, the investment

---

[6] Although we maintain that the investment approach is an effective way of interpreting the cross section of expected returns, it is *incomplete* for asset pricing. In particular, it cannot be used to price derivatives. An investment approach to pricing derivatives would need a purely production-based M (see Cochrane (1993) and Belo (2010)).

[7] Asset pricing anomalies are often interpreted as investor "irrationality." De Bondt and Thaler (1985, p. 793) write: "Research in experimental psychology suggests that, in violation of Bayes' rule, most people tend to 'overreact' to unexpected and dramatic news events. This study of market efficiency investigates whether such behavior affects stock prices. The empirical evidence, based on CRSP monthly return data, is consistent with the overreaction hypothesis. Substantial weak form market inefficiencies are discovered." Jegadeesh and Titman (1993, p. 90) write: "The market underreacts to information about the short-term prospects of firms but overreacts to information about their long-term prospects," and that "investor expectations are systematically biased." Ritter (1991, p. 3) interprets the long-run performance of initial public offerings as "consistent with an IPO market in which (1) investors are periodically overoptimistic about the earnings potential of young growth companies, and (2) firms take advantage of these 'window of opportunity'." Sloan (1996, p. 289) write: "[S]tock prices are found to act as if investors 'fixate' on earnings, failing to reflect fully information contained in the accrual and cash flow components of current earnings until that information impacts future earnings." Cooper, Gulen, and Schill (2008, p. 1648) write: "In functionally efficient markets, invest-

approach has nothing to say about whether anomalies are driven by rational or irrational forces.

We disagree. First, by spelling out the technological underpinnings of asset prices, the investment approach has some implications for the "rationality" of asset prices. In a linear technologies economy, any investor irrationality must only impact on quantities via the optimal investment of rational firms, with no effect on asset prices. If we lived in an endowment economy, quantities cannot be adjusted, and investor irrationality must impact fully on asset prices. The real world is an adjustment costs economy that lies somewhere in between. Irrationality could deliver a short term effect on prices, but rational firms eventually come in, overcome adjustment costs, and flood any "fire" of bubble with the "water" of investment, so as to extinguish any long term impact on asset prices.

Second, more directly, the investment approach shows that firms' investment decisions are aligned correctly with the cost of capital. Firms invest more when their costs of capital are low, and vice versa. To the extent that this alignment manifests itself as many empirical relations between characteristics and average returns, these relations in the anomalies literature directly say *nothing* about investor rationality or irrationality. A low cost of capital could result from the sentiment of irrationally optimistic investors or the low market prices of risk demanded by rational investors. The investment first-order condition then correctly connects the low cost of capital with, for example, high asset growth, low profitability, low short-term prior returns, and high accruals.

Third, we do *not* claim that connecting the cost of capital with characteristics "explains" anomalies, or proves the rationality of asset prices. To do so, one must meet the high hurdle of showing that both the consumption first-order condition *and* the investment first-order condition hold in the data. We have not made any progress in improving the performance of the consumption model. In fact, it is precisely the poor performance of the consumption approach that forces us to explore the investment approach as an alternative to understanding anomalies without invoking mispricing.

More important, our point is that *connecting the cost of capital with characteristics per se does*

---

ment opportunities are priced such that capital can be systematically allocated to the most productive uses." "In contrast, bias in the capitalization of new investments leads to a host of potential investment policy distortions." Cooper et al. interpret the asset growth effect as saying "such potential distortions are present and economically meaningful."

not prove "irrationality" either, in contrast to the standard practice in the anomalies literature (see footnote 7). Behaviorists are likely to respond by saying that $E[Mr^S] = 1$ fails to "explain" the connection. But the failure can be due to specification errors of M and measurement errors in consumption data. Also, the failure speaks only to the consumption approach, and should not be treated as a failure of the entire "rationality" paradigm. By Fama's (1998) bad-model logic, if a model of market equilibrium fails with a lot of measurement difficulties, we go to a different model. In this sense, we say that the investment approach fills a gap left by the consumption approach.

4.5.4 Is the Investment-Based Expected Return Test A Weak Consistency Test?

Cochrane (1991) notes that equation (17), taken literally, means that the investment return equals the stock return for every stock, every period, and every state of the world. Because no choice of parameters can satisfy this 100% R^2 equation, it is formally rejected at any level of significance. One has to use some common sense to decide how to take this equation to the data, while being artful about when the equation illuminates key aspects of the data and when it is being pushed too far. Liu, Whited, and Zhang (2009), for example, test the ex-ante restriction that the *expected* (levered) investment return equals the *expected* stock return, and interpret this restriction as a weaker condition than the ex-post restriction in equation (17).

Critics argue that the ex-ante restriction is not literally a model of expected return "determination," but rather an ancillary implication of the neoclassical investment theory. We disagree. First, the restriction is derived from the investment first-order condition, which is as primitive as the consumption first-order condition in general equilibrium. As equilibrium conditions, first-order conditions are not "ancillary." Second, testing the ex-ante restriction captures the *essence* of the economic question: Why do testing portfolios formed on anomaly variables earn different subsequent stock returns *on average*? The cross section of returns means the cross section of *expected* returns!

Third, unlike $E[Mr^S] = 1$ that is only about expected returns, the investment return equation speaks to an infinite number of moments for stock returns. It is an advantage for a parsimonious equation to produce more refutable predictions. In fact, the ex-post nature of the investment return helps interpret the puzzling pattern of earnings announcement returns. For example, Jegadeesh and Titman (1993, Table IX) document that the average three-day returns (from day -2 to 0) around quarterly earnings announcement dates represent about 25% of momentum for the first six-month holding period, and the announcement date returns also display reversal in long horizons.[8]

Because $E[Mr^S] = 1$ has nothing to say about ex-post returns, the announcement returns pattern has been interpreted as investors' expectational errors. However, the investment return equation in fact predicts the ex-post realizations of returns around earnings announcements. The logic is as follows. Without adjustment costs, the ex-post investment return equation (17) reduces to $\Pi_{i1} + 1 - \delta$. The expected return should be realized *only* around earnings announcement dates, when news about firm i's profitability is released to the markets. In a dynamic world with adjustment costs, in addition to ex-post profitability, the ex-post investment return is also affected by the ex-post growth rate of investment-to-capital. As such, only a portion of the expected return is realized around earnings announcement dates (see, e.g., Wu, Zhang, and Zhang (2010)).

Critics also argue that the expected investment return should be used as a means to forecast stock returns ex-ante. In cross-sectional regressions of subsequent stock returns, the expected investment return should have a positive slope, and even drive out common predictors. In portfolio sorts, we should see a large average return spread across portfolios formed on the expected investment return. Although theoretically sound, this idea is not practical. It amounts to pushing the 100% R^2 equation too far. There are measurement errors in components of the investment return such as capital, and specification errors in the functional forms of the production and capital adjustment technologies. These errors can destroy the predictive power of the expected investment return for future stock returns. A more productive strategy would be to pick *components* of the

---

[8] In addition, Sloan (1996, p. 312) documents that "over 40 percent of the predictable stock returns [related to accruals] are concentrated around the subsequent quarterly earnings announcements, even though the announcement period contains less than five percent of the total trading days." La Porta, Lakonishok, Shleifer, and Vishny (1997, p. 859) study "stock price reactions around earnings announcements for value and glamour stocks over a 5-year period after portfolio formation" and find that "a significant portion of the return difference between value and glamour stocks is attributable to earnings surprises that are systematically more positive for value stocks."

investment return that are relatively immune to measurement and specification errors, and use them instead to forecast stock returns. This strategy is exactly what Chen, Novy-Marx, and Zhang (2011) pursue by using investment-to-capital and profitability to forecast stock returns.

5 Conclusion

What risk factors "explain" asset pricing anomalies? The question has been at the center of modern asset pricing research. We question the question. That no solution has emerged from decades of search should at least alert us of the possibility that the question itself is a *wrong* question to ask. In particular, Kuhn (1962, p. 36–37) argues:

> "[T]he really pressing problems, e.g., a cure for cancer for the design of a lasting peace, are often not puzzles at all, largely because they may not have any solution. Consider the jigsaw puzzle whose pieces are selected at random from each of two different puzzle boxes. Since that problem is likely to defy (though it might not) even the most ingenious of men, it cannot serve as a test of skill in solution. In any usual sense, it is not a puzzle at all. *Though intrinsic value is no criterion for a puzzle, the assured existence of a solution is* (our emphasis)."

To us, the search for risk factors is pointless. The identity of risk factors is not a clear testable implication of equilibrium theory: First-order conditions are! The search is also hopeless: Measurement errors forever doom covariances in horse races against characteristics. Because expected returns can be inferred directly from characteristics, why do we keep fixating on risk factors? We are stuck with the risk factors way of thinking purely as a result of inertia from the CAPM. This atheoretical *mechanical* way of thinking has been asking the *wrong* question, dividing the field into two warring schools based on a *false* premise with *no* economic basis, and leading our science into an *abyss* with insurmountable measurement difficulties.

While sympathetic to Daniel and Titman (1997) in using characteristics to forecast returns, we suggest that their covariances versus characteristics test should be interpreted with extreme caution. Behavioral theories such as Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Hong and Stein (1998) typically rule out any rational variation of expected returns by assuming risk neutrality. As such, the behavioral null underlying the Daniel and Titman test is that covariances do not matter at all, and only characteristics matter in "explaining" expected returns. Their evidence does not imply mispricing automatically, for two reasons: (i) There are likely severe measurement errors in estimated covariances; and (ii) the standard investment model predicts that characteristics *should* be connected with expected returns.

More generally, we argue that the anomalies literature is built on the *wrong* premise: If an empirical characteristic-return relation is consistent with investor rationality, the relation must be "explained" by a risk-based factor model. The investment approach turns this doctrine upside down: (i) We interpret characteristics-based factor models not as risk factors in the ICAPM or APT, but as linear approximations to firm-level investment returns; (ii) we interpret the evidence that characteristics dominate covariances in asset pricing tests not as mispricing, but as measurement errors in estimated covariances; and (iii) we interpret the investment return equation not as the neoclassical investment theory in first differences, but as an indispensable complement to the consumption approach to asset pricing in general equilibrium, especially for the cross section.

If searching for risk factors is fruitless, where do we go from here? The investment approach is a new basis for asset pricing research: The anomalies literature should be turned into the *regularities* literature using characteristics-based expected return models from the investment approach.

First, one can adapt the Liu, Whited, and Zhang (2009) model to examine a broader set of anomalies. It is also possible, but challenging, to extend their model to firm level estimation. The main obstacle is to extend the Euler equation approach underlying the investment return to firm level data that often involve nonconvex capital adjustment technologies. Second, the investment-based three-factor model of Chen, Novy-Marx, and Zhang (2011) provides an alternative to the Fama-French model. Third, Xue (2011) proposes new characteristics-based metrics by matching on investment-to-capital and profitability to evaluate mutual fund performance. Daniel, Grinblatt, Titman, and Wermers (1997) use characteristics-based benchmarks by matching on size, book-to-market, and short-term prior returns. Alas, factor models are more common, probably because one can appeal, albeit informally, to the ICAPM or APT for "motivation," while characteristics appear ad hoc. However, the investment approach, by predicting the relations between characteristics and expected returns, has provided the economic motivation for matching on characteristics directly.

References

Ang, Andrew and Joseph Chen, 2007, CAPM over the long run: 1926–2001, *Journal of Empirical Finance* 14, 1–40.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barberis, Nicholas, and Richard Thaler, 2003, A survey of behavioral finance, in *Handbook of the Economics of Finance*, George M. Constantinides, Milton Harris, and René M. Stulz, eds., Elsevier, North Holland, 1053–1123.

Belo, Frederico, 2010, Production-based measures of risk for asset pricing, *Journal of Monetary Economics* 57, 146–163.

Belo, Frederico, and Xiaoji Lin, 2011, The investment growth spread, forthcoming, *Review of Financial Studies.*

Berk, Jonathan, and Peter DeMarzo, 2011, *Corporate Finance*, 2nd edition, Prentice Hall.

Brock, William A., 1982, Asset prices in a production economy, in: John L. McCall, ed., *Economics of Information and Uncertainty,* University of Chicago Press, Chicago Illinois.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2011, An alternative three-factor model, working paper, The Ohio State University.

Cochrane, John H., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–237.

Cochrane, John H., 1993, Rethinking production under uncertainty, working paper, University of Chicago.

Cochrane, John H., 2005, *Asset Pricing*, Princeton University Press, Princeton New Jersey.

Cooley, Thomas F., and Edward C. Prescott, 1995, Economic growth and business cycles, in Thomas F. Cooley, ed., *Frontiers of Business Cycle Research*, Princeton University Press, Princeton New Jersey.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1652

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, An intertemporal general equilibrium model of asset prices, *Econometrica* 53, 363–384.

Craine, Roger, 1989, Risky business: The allocation of capital, *Journal of Monetary Economics* 23, 201–218.

Daniel, Kent D., Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristics-based benchmarks, *Journal of Finance* 52, 1035–1058.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

Daniel, Kent D., David Hirshleifer, and Siew Hong Teoh, 2002, Investor psychology in capital markets: Evidence and policy implications, *Journal of Monetary Economics* 49, 139–209.

Daniel, Kent D., and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Daniel, Kent D., Sheridan Titman, and K. C. John Wei, 2001, Explaining the cross-section of stock returns in Japan: Factors or characteristics, *Journal of Finance* 56, 743–766.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

De Bondt, Werner F. M., and Richard Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793–805.

Dechow, Patricia M., Natalya V. Khimich, and Richard G. Sloan, 2011, The accrual anomaly, working paper, University of California, Berkeley.

Fama, Eugene F., 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283–306.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–466.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanation of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of capital, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F. and Merton H. Miller, 1972, *The Theory of Finance*, Holt, Rinehart, and Winston.

Ferson, Wayne E., and Campbell R. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385–415.

Fisher, Irving, 1930, *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*, 1st ed., New York: The Macmillan Co.

Gebhardt, William R., Soeren Hvidkjaer, and Bhaskaran Swaminathan, 2005, The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of Financial Economics* 75, 85–114.

Gomes, Joao F., Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

Ghysels, Eric, 1998, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance* 53, 549–573.

Hirshleifer, David, Kewei Hou, and Siew Hong Teoh, 2011, The accrual anomaly: Risk or mispricing? forthcoming, *Management Science.*

Hirshleifer, Jack, 1970, *Investment, Interest, and Capital*, Prentice-Hall, Englewood Cliffs, New Jersey.

Hong, Harrison and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hou, Kewei, G. Andrew Karolyi, Bong-Chan Kho, 2011, What factors drive global stock returns? forthcoming, *Review of Financial Studies.*

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Kogan, Leonid, and Dimitris Papanikolaou, 2011, Investment shocks, firm characteristics, and the cross-section of expected returns, working paper,

Kuhn, Thomas S., 1962, *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago and London.

La Porta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny, 1997, Good news for value stocks: Further evidence on market efficiency, *Journal of Finance* 52, 859–874.

Li, Erica X. N., Dmitry Livdan, and Lu Zhang, 2009, Anomalies, *Review of Financial Studies* 22, 4301–4334.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

Liu, Laura Xiaolei, Toni M. Whited, and Lu Zhang, 2009, Investment-based expected stock returns, *Journal of Political Economy* 117, 1105–1139.

Long, John B. Jr., and Charles I. Plosser, 1983, Real business cycles, *Journal of Political Economy* 91, 39–69.

Lundholm, Russell J., and Richard G. Sloan, 2007, *Equity Valuation and Analysis*, 2nd edition, McGraw-Hill Irwin.

Lucas, Robert E., Jr., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429–1445.

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

Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.

Miller, Merton H., and Myron S. Scholes, 1972, Rate of return in relation to risk: A reexamination of some recent findings, in *Studies in the Theory of Capital Markets*, Michael C. Jensen ed., New York: Praeger.

Penman, Stephen H., 2010, *Financial Statement Analysis and Security Valuation*, 4th edition, McGraw-Hill Irwin.

Richardson, Scott, İrem Tuna, and Peter Wysocki, 2010, Accounting anomalies and fundamental analysis: A review of recent research advances, *Journal of Accounting and Economics* 50, 410–454.

Ritter, Jay R., 1991, The long-run performance of initial public offerings, *Journal of Finance* 46, 3–27.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Rubinstein, Mark, 2006, *A History of the Theory of Investments*, John Wiley and Sons, Inc.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289–315.

Xing, Yuhang, 2008, Interpreting the value effect through the Q-theory: An empirical investigation, *Review of Financial Studies* 21, 1767–1795.

Xue, Chen, 2011, The investment approach to evaluating mutual fund performance, working paper, University of Michigan.

Wu, Jin (Ginger), Lu Zhang, and X. Frank Zhang, 2010, The q-theory approach to understanding the accrual anomaly, *Journal of Accounting Research* 48, 177–223.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.