Anomalies

Lu Zhang*

October 2004†

Abstract

I construct a neoclassical, Q-theoretical foundation for time-varying expected returns in connection with corporate policies. Under certain conditions, stock return equals investment return, which is directly tied with characteristics. This single equation is shown analytically to be qualitatively consistent with many anomalies, including the relations of future stock returns with market-to-book, investment and disinvestment rates, seasoned equity offerings, tender offers and stock repurchases, dividend omissions and initiations, expected profitability, profitability, and to certain extent, earnings announcement. The Q-framework also provides a new asset pricing test.

*William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, NY 14627. Tel: (585)275-3491, fax: (585)273-1140, and email: zhanglu@simon.rochester.edu.

†I thank Rui Albuquerque, Yigit Atilgan, Mike Barclay, Robert Bloomfield, Murray Carlson, Joao Gomes, Leonid Kogan, John Long, Roni Michaely, Bill Schwert, Jerry Warner, David Weinbaum, Joanna Wu, and other seminar participants at Johnson School of Management at Cornell University, Simon School of Business and Department of Economics at University of Rochester for helpful comments. The usual disclaimer applies.

1 Introduction

A large body of empirical literature in financial economics has documented relations of stock returns with firm characteristics and corporate events, relations that are called anomalies because they are hard to explain using current asset pricing models (e.g., Fama (1998) and Schwert (2003)). Many believe that these anomalies are strong evidence against efficient markets and rational expectations (e.g., Shleifer (2000) and Barberis and Thaler (2003)).

I construct a neoclassical, Q-theoretical foundation for time-varying expected returns in connection with corporate policies. If the operating-profit and the adjustment-cost functions have the same degree of homogeneity, stock return equals investment return, which is directly tied with characteristics and corporate policies via the first principles of optimal investment.

By basically taking and signing partial derivatives of investment returns, I demonstrate analytically that the Q-framework is qualitatively consistent with many anomalies that are often interpreted as over- and underreaction under inefficient markets, anomalies including:[1]

- The investment anomaly: The investment rate is negatively correlated, but the disinvestment rate is positively correlated with future stock returns.
- The value anomaly: Average returns correlate negatively with market-to-book, and the magnitude of this correlation decreases with the market value.
- The payout anomaly: When firms tender for their stocks or announce share repurchases and dividend initiations, they earn positive long-term abnormal returns. And the magnitude of the abnormal returns is stronger in value firms than in growth firms.
- The seasoned-equity-offering (SEO) anomaly: Firms conducting SEOs earn lower average returns in the next three to five years than nonissuing firms. And the magnitude of this underperformance is stronger in small firms than in big firms.
- The expected-profitability anomaly: Expected profitability correlates positively with expected returns. And this correlation decreases with the market value.

[1] Appendix A briefly reviews the specific empirical papers documenting these anomalies. Behavioral theories that use over- and underreaction to explain these anomalies include Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Hong and Stein (1999).

- The profitability anomaly: Given market-to-cash flows or market-to-book, more profitable firms earn higher average returns. This relation is stronger in small firms.
- The earnings-momentum anomaly (post-earnings-announcement drift): Firms with high earnings surprise earn higher average returns than firms with low earnings surprise. And this anomaly is stronger in small firms.

In a nutshell, I demonstrate that, much like aggregate expected returns that vary over business cycles, expected returns in the cross section vary with corporate policies or events. And this is achieved in a prototypical, neoclassical model with rational expectations.

Intuitively, investment return from time t to $t+1$ equals the ratio of the marginal profit of investment at $t+1$ divided by the marginal cost of investment at t. This equation suggests two economic mechanisms through which the anomalies can be understood.

The first four anomalies can be explained by the Q-mechanism. Specifically, the Q-theory is a theory of investment demand. And the downward-sloping investment-demand function implies a negative relation between cost of capital or expected return and investment rate. Further, investment rate increases with the marginal q, the expected present value of future marginal profits of capital. The marginal q is in turn proportional to market-to-book. The negative slope of the investment-demand function then implies a negative relation between expected return and market-to-book. The payout anomaly follows because firms' cash-flow constraint (that equates the sources with the uses of funds) implies a negative relation between the payout and investment rates. And the SEO anomaly follows because the cash-flow constraint implies a positive relation between the financing and investment rates.

With decreasing return to scale or strictly convex adjustment costs, the relation between expected return and market-to-book is convex — the second-order partial derivative of expected return with respect to market-to-book is strictly positive.[2] This convexity manifests itself, by the chain rule of partial derivatives, as the stronger value anomaly in small firms, the stronger SEO anomaly in small firms, and the stronger payout anomaly in value firms.

[2] In the example of quadratic adjustment costs, the investment-demand function is also convex.

In contrast, the last three listed anomalies can be explained by the marginal product of capital (MPK) at time $t+1$ in the numerator of investment return through the MPK-mechanism. Specifically, MPK is proportional to profitability, a fact that implies a positive relation between expected profitability and expected return. This positive relation in turn explains the profitability anomaly because profitability is a strong, positive predictor of future profitability. And the post-earnings-announcement drift can be explained if earnings surprise and profitability contain similar information on future profitability.

Intriguingly, my Q-explanations of anomalies do not involve risk, at least directly, even though my model is entirely rational. The reason is that I derive expected returns from firms' optimality conditions, instead of consumers'. As a result, stochastic discount factor (SDF) and its covariances with returns do not directly enter the expected-return determination. And firm characteristics are sufficient statistics for expected returns. Therefore, the current debate on covariances versus characteristics in the empirical literature (e.g., Daniel and Titman (1997) and Davis, Fama, and French (2000)) is not a well-defined question.

I also propose the Q-representation of expected returns as a new empirical asset pricing framework. Although internally consistent with the beta- and the SDF-framework in theory, the Q-representation is likely to have comparative advantages over the two standard methods in practice. The reason is that the Q-representation avoids the difficult tasks of estimating covariances and of identifying the right form of SDF.[3]

The insight that stock and investment returns are equal first appears in Cochrane (1991). Cochrane (1991, 1996) is also among the first to study asset prices from firms' perspective. Restoy and Rockinger (1994) formally establish this equivalence under linear homogeneity. An early version of Gomes, Yaron, and Zhang (2004) extends the result under debt financing. And I extend the result under homogeneity of the same degree for the operating-profit and adjustment-cost functions. I differ further from these previous papers, which focus on aggregate investment returns, in that I aim to understand anomalies in the cross section.

The Q-theory is originated by Tobin (1969). Hayashi (1982) establishes the equivalence

[3]Fama and French (1997) show that estimated factor loadings in beta-pricing models are extremely imprecise even at the industry level.

between marginal q and average Q under linear homogeneity. Abel and Eberly (1994) extend this result into a stochastic setting with partial irreversibility and fixed costs proportional to capital. They also show that marginal q is proportional to average Q when the operating-profits and the adjustment-cost functions are homogeneous of the same degree, a result I use extensively. The Q-theory has been used mostly to understand the behavior of investment. But I open the door for its large-scale applications to the cross section of returns.

My work shares its long-term goal with the growing literature that aims to understand the real determinants of the cross section of returns (e.g., Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003)).[4] I contribute by expanding greatly the scope of explained anomalies and by unifying many anomalies under a single, analytical framework. I also propose a new empirical framework in which many ideas of this theoretical literature can be tested. Comparisons with specific papers are presented throughout Section 3.

The rest of the paper is organized as follows. Section 2 sets up the model and establishes the equivalence between stock and investment returns. Section 3 uses this equivalence to explain anomalies. Section 4 discusses empirical implications. Finally, Section 5 concludes.

2 The Model of the Firm

This section presents the basic elements of the Q-theory. My exposition is heavily influenced by Abel and Eberly (1994) and an early version of Gomes, Yaron, and Zhang (2004). Section 2.1 describes the basic environment. Section 2.2 characterizes the behavior of firm value-maximization, and establishes the equivalence between stock and investment returns.

2.1 The Environment

Consider a firm that uses capital and a vector of costlessly adjustable inputs, such as labor, to produce a perishable output. The firm chooses the levels of these inputs each period to maximize its operating profit, defined as its revenue minus the expenditures on these inputs.

[4]Other examples include Berk (1995), Johnson (2002), Berk, Green, and Naik (2004), Carlson, Fisher, and Giammarino (2004a, 2004b), Cooper (2004), Gomes, Yaron, and Zhang (2004), Kogan (2004), Sapriza and Zhang (2004), Whited and Wu (2004), and Zhang (2004).

Taking the operating profit as given, the firm then chooses optimal investment to maximize its market value. Capital investment involves costs of adjustment.

The Operating-Profit Function

Let $\Pi_t = \Pi(K_t, X_t)$ denote the maximized operating profit at time t, where K_t is the capital stock at time t and X_t is a vector of random variables representing exogenous shocks to the operating profit, such as aggregate and firm-specific shocks to production technology, shocks to the prices of costlessly adjusted inputs, or industry- and firm-specific shocks to the demand of the output produced by the firm.

Assumption 1 *The operating profit function is homogeneous of degree α with $\alpha \leq 1$:*

$$\Pi(K_t, X_t) = \Lambda(X_t) K_t^{\alpha} \quad \text{where} \quad \Lambda(X_t) > 0 \tag{1}$$

If $\alpha = 1$, the operating-profit function displays linear homogeneity in K_t. This applies to a competitive firm being a price-taker in output and factor markets.[5] And $\alpha < 1$ means market power (e.g., Cooper and Ejarque (2001)).

From Assumption 1,

$$\alpha \Pi(K_t, X_t) = \Pi_1(K_t, X_t) K_t \tag{2}$$

Marginal product of capital is strictly positive, $\Pi_1(K_t, X_t) > 0$, where subscript i denotes the first-order partial derivative with respect to the i^{th} argument. Multiple subscripts denote high-order partial derivatives. $\Pi_1(K_t, X_t)$ decreases with capital, reflecting decreasing return to scale, $\Pi_{11}(K_t, X_t) \leq 0$, where the inequality is strict when $\alpha < 1$. Finally, $\Pi_{111}(K_t, X_t) \geq 0$.

[5] This can be seen from the static maximization problem of the firm that chooses the vector of costlessly adjustable inputs. Let $\mathbf{L}_t$ denote this vector and $F(K_t, \mathbf{L}_t, X_t)$ denote the revenue function that is linearly homogenous in K_t and $\mathbf{L}_t$. If the firm is a price-taker, its operating profit can be written as:

$$\Pi(K_t, X_t) = \max_{\mathbf{L}_t} \{F(K_t, \mathbf{L}_t, X_t) - \mathbf{W}_t' \mathbf{L}_t\} = \max_{\mathbf{L}_t/K_t} \{[F(1, \mathbf{L}_t/K_t, X_t) - \mathbf{W}_t'(\mathbf{L}_t/K_t)] K_t\} = \Lambda(X_t) K_t$$

where $\mathbf{W}_t$ is the vector of market prices of the costlessly adjustable inputs, the second equality follows from the linear homogeneity of $F(K, \mathbf{L}, X)$ in K and $\mathbf{L}$, and the third equality follows by defining $\Lambda(X_t) \equiv \max_{\mathbf{L}_t/K_t} \{[F(1, \mathbf{L}_t/K_t, X_t) - \mathbf{W}_t'(\mathbf{L}_t/K_t)]\}$. The first-order condition with respect to $\mathbf{L}_t$ says that $F_2(K_t, \mathbf{L}_t, X_t) = \mathbf{W}_t$. The linear homogeneity of $F(K_t, \mathbf{L}_t, X_t)$ in K_t and $\mathbf{L}_t$ then implies that $F(1, \mathbf{L}_t/K_t, X_t) - \mathbf{W}_t'(\mathbf{L}_t/K_t) = F_1(1, \mathbf{L}_t/K_t, X_t)$ which is clearly positive. Therefore, $\Pi_1(K_t, X_t) = \Lambda(X_t) > 0$. If $F_3(K_t, \mathbf{L}_t, X_t)$ is positive, then $\Lambda'(X_t) > 0$.

More importantly, equation (2) implies that marginal product of capital, Π_1, is also proportional to operating profit-to-capital, $\Pi(K_t, X_t)/K_t$. This ratio corresponds roughly to accounting profitability (earnings-to-book) plus depreciation rate.[6]

The Augmented Adjustment-Cost Function

Capital accumulates according to:

$$K_{t+1} = I_t + (1 - \delta)K_t \tag{3}$$

Thus end-of-period capital equals real investment plus beginning-of-period capital net of depreciation. And capital depreciates at a fixed proportional rate of δ.

When the firm undertakes investment, it incurs costs because of: (i) purchase/sale costs, (ii) convex costs of physical adjustment, and (iv) weakly convex costs of raising capital when the sum of the first two components is higher than the operating profit.

(i) Purchase/sales costs are incurred when the firm buys or sells uninstalled capital. When the firm disinvests, this cost is negative. For analytical convenience, I assume that the relative purchase price and relative sale price of capital are both equal to unity. This differs from Abel and Eberly (1994), who assume that purchase price is higher than sale price to capture costly reversibility because of, for example, firm-specificity of capital and adverse selection in the market for used capital. In this case, the purchase/sale cost function is not differentiable at $I_t = 0$. My assumption maintains this differentiability. Costly reversibility can still be captured by letting the convex costs of disinvestment be uniformly higher than those of investment with equal magnitudes (e.g., Hall (2001) and Zhang (2004)).

(ii) Convex costs of physical adjustment are nonnegative costs that are zero when $I_t = 0$. These costs are continuous, strictly convex in I_t, non-increasing in capital K_t, and differentiable with respect to I_t and K_t everywhere. The second-order partial derivative of the convex-cost function with respect to K_t is nonnegative. It is straightforward to verify

[6] The operating profit in the model corresponds approximately to earnings plus capital depreciation in the data. This assumes that accruals are only used to mitigate the accounting timing and matching problems that deviate operating cash flow from earnings in practice (e.g., Dechow (1994)). These accounting problems are abstracted from the model.

that the standard quadratic, convex adjustment-cost function satisfies all these assumptions.

(iii) Costs of raising capital are incurred when the financial deficit, denoted O_t, is strictly positive. I define O_t as the higher value between zero and the sum of the purchase/sale costs and convex costs of adjustment minus the operating profit. I assume that the financing-cost function is continuous, weakly convex in O_t (and hence in I_t) and decreasing in K_t. Its first-order partial derivative with respect to O_t (and hence with respect to I_t) is zero when $O_t = 0$. The financing-cost function is differentiable with respect to O_t (and hence with respect to I_t) and K_t everywhere. And the second-order partial derivative of the function with respect to K_t is nonnegative. Previous studies of financing costs (e.g., Gomes (2001) and Hennessy and Whited (2004)) assume that the costs are proportional to the amount of funds raised. And quadratic costs can be defined as $(b/2)(O_t/K_t)^2 K_t$ with $b > 0$. Both the proportional and the quadratic financing-cost functions satisfy the aforementioned assumptions.

The flip side of financial deficit is free cash flow, denoted C_t. I define C_t as the higher value between zero and the operating profit minus the sum of the purchase/sale costs and the convex costs of adjustment. I assume that whenever C_t is strictly positive, the firm pays it back to its shareholders either in the form of dividends or stock repurchases. The model is silent on the behavior of cash hoarding or on the form of payout. Further, I assume that the firm does not pay any extra costs when paying cash out of the firm. Therefore, the firm either raises capital or distributes payout, but never at the same time.

The total cost of investment represents the sum of purchase/sale costs, convex costs of physical adjustment, and costs of raising capital. I denote the total cost as $\Phi(I_t, K_t)$, and refer to it as the augmented adjustment-cost function. To summarize,

Assumption 2 *The augmented adjustment-cost function* $\Phi(I_t, K_t)$ *satisfies that:*

$$\Phi_2(I_t, K_t) \leq 0; \quad \Phi_{22}(I_t, K_t) \geq 0; \quad \text{and} \quad \Phi_{11}(I_t, K_t) > 0;$$

The most important technical assumption is stated explicitly below:

Assumption 3 *The augmented adjustment-cost function is homogeneous of the same* *degree, α, in I_t and K_t, as the operating-profit function is in K_t. In other words,*

$$\Phi(I_t, K_t) = G\left(\frac{I_t}{K_t}\right) K_t^{\alpha} \tag{4}$$

Coupled with Assumption 2, Assumption 3 implies that $G''(\cdot) > 0$ and that

$$\alpha \Phi(I_t, K_t) = \Phi_1(I_t, K_t) I_t + \Phi_2(I_t, K_t) K_t \tag{5}$$

Assumption 3 is crucial in establishing the equivalence between stock and investment returns (see the proof of Proposition 2 in Appendix B). But how restrictive is Assumption 3? Abel and Eberly (1994) discuss its content for the case of linear homogeneity. I follow their exposition except for the financing-cost function. The linear homogeneity of $\Phi(I_t, K_t)$ means that each of its three components is linearly homogenous. (i) A doubling of I_t doubles the purchase/sale costs that are linear in I_t, and are independent of K_t. (ii) The investment literature typically assumes that physical adjustment costs are linearly homogenous (e.g., Hayashi (1982), Abel and Blanchard (1983), and Abel and Eberly (1994)). And (iii) the proportional and quadratic financing-cost functions are linearly homogeneous in I_t and K_t.

Relative to the specification in Abel and Eberly (1994), my augmented adjustment-cost function adds the convex costs of financing, but ignores the wedge between purchase and sale prices of capital and fixed costs of adjustment. The fixed costs of raising capital are not included either. Incorporating these features will compromise the differentiability of $\Phi(I_t, K_t)$ with respect to I_t at the two points where $I_t = 0$ and $O_t = 0$. The theory below works almost everywhere but at these two points where investment return is ill-defined because Φ_1 does not exist (see equation (15) below). Although not implemented here, it is possible to define two different investment returns at these two points using the left- and the right-side partial derivatives of Φ with respect to I_t.

More important, including the wedge between the purchase and sale prices of capital and fixed costs of investment and raising capital leaves the crucial Assumption 3 unaltered. As argued in Abel and Eberly (1994), the purchase/sale costs are proportional to I_t. And the fixed costs are linearly homogenous in K_t, if they reflect the costs of interrupting production, and are hence proportional to the operating profit and to capital.

2.2 Value Maximization

I now characterize firm's value-maximization behavior. The dynamic problem is:

$$V(K_t, X_t) = \max_{\{I_{t+j}, K_{t+1+j}\}_{j=0}^{\infty}} \mathrm{E}_t \left[\sum_{j=0}^{\infty} M_{t,t+j} \left(\Pi(K_{t+j}, X_{t+j}) - \Phi(I_{t+j}, K_{t+j}) \right) \right] \qquad (6)$$

where $V(K_t, X_t)$ is the cum-dividend market value — when $j=0$, $\Pi(K_t, X_t) - \Phi(I_t, K_t)$ is included in $V(K_t, X_t)$. $M_{t,t+j} > 0$ is the stochastic discount factor from time t to $t+j$. $M_{t,t} = 1$ and $M_{t,t+i} M_{t+i+1,t+j} = M_{t,t+j}$ for some integer i between 0 and j. For notational simplicity, I use M_{t+j} to denote $M_{t,t+j}$ whenever the starting date is t.

Marginal q, Tobin's Average Q, and Market-to-Book

Lemma 1 *Under Assumptions 1 and 3, the value function is also homogenous of degree α:*

$$\alpha V(K_t, X_t) = V_1(K_t, X_t) K_t$$

Define Tobin's average Q as $\widehat{Q}_t \equiv V(K_t, X_t)/K_t$, then $V_1(K_t, X_t) = \alpha \widehat{Q}_t$.

Proof. See Appendix B. ■

Let q_t be the present-value multiplier associated with capital accumulation equation (3). The Lagrange formulation of the firm value, $V(K_t, X_t)$, is then:

$$\max_{\{I_{t+j}, K_{t+1+j}\}_{j=0}^{\infty}} \mathrm{E}_t \left[\sum_{j=0}^{\infty} M_{t+j} \left(\Pi(K_{t+j}, X_{t+j}) - \Phi(I_{t+j}, K_{t+j}) - q_{t+j}[K_{t+j+1} - (1-\delta) K_{t+j} - I_{t+j}] \right) \right] \qquad (7)$$

The first-order conditions with respect to I_t and K_{t+1} are, respectively,

$$q_t = \Phi_1(I_t, K_t) \qquad (8)$$

$$q_t = \mathrm{E}_t[M_{t+1}[\Pi_1(K_{t+1}, X_{t+1}) - \Phi_2(I_{t+1}, K_{t+1}) + (1-\delta) q_{t+1}]] \qquad (9)$$

Solving equation (9) recursively yields an economic interpretation for the marginal q:

Lemma 2 *The marginal q is the expected present value of marginal profits of capital:*

$$q_t = \mathrm{E}_t \left[\sum_{j=1}^{\infty} M_{t+j} (1-\delta)^{j-1} (\Pi_1(K_{t+j}, X_{t+j}) - \Phi_2(I_{t+j}, K_{t+j})) \right] \qquad (10)$$

Proof. See Appendix B. ■

Proposition 1 ***(The Link between Marginal*** q ***and Market-to-Book)*** *Define the ex-dividend firm value,* P_t*, as:*

$$P_t \equiv P(K_t, K_{t+1}, X_t) = V(K_t, X_t) - \Pi(K_t, X_t) + \Phi(I_t, K_t) \tag{11}$$

And define the market-to-book equity as $Q_t \equiv P_t/K_{t+1}$ *then under Assumptions 1 and 3,*

$$q_t = \alpha Q_t \tag{12}$$

Proof. See Appendix B. ■

In continuous time formulation of the Q-theory (e.g., Hayashi (1982) and Abel and Eberly (1994)), the marginal q_t is proportional to Tobin's average Q_t, i.e., $q_t = \alpha \widehat{Q}_t$. But in discrete time, $V_1(K_t, X_t)$ is not exactly the marginal q_t. The time-to-build convention reflected in the capital accumulation equation (3) implies that one unit of investment today only becomes effective next period. As a result, q_t and $\widehat{Q}_t$ are linked through:

$$q_t = \alpha \mathrm{E}_t[M_{t+1} \widehat{Q}_{t+1}] \tag{13}$$

To see this, taking partial derivatives of both sides of equation (7) with respect to K_t yields $V_1(K_t, X_t) = \Pi_1(K_t, X_t) - \Phi_2(I_t, K_t) + q_t(1-\delta)$. Combining this equation with Lemma 1 and equation (9) yields equation (13).

Several useful properties of the function $\Phi(I_t, K_t)$ evaluated at the optimum can be established using equation (8) and the link between marginal q and average $\widehat{Q}_t$.

Lemma 3 *Under Assumptions 1 and 2, the augmented adjustment-cost function* $\Phi(I_t, K_t)$*, when evaluated at the optimum, satisfies:*

$$\Phi_1(I_t, K_t) > 0; \quad \Phi_{12}(I_t, K_t) \leq 0; \quad \text{and} \quad \Phi_{122}(I_t, K_t) \geq 0$$

Proof. See Appendix B for the proof of the last two inequalities. The first inequality can be shown as follows. From Assumptions 1 and 2, $\Pi_1 > 0$ and $\Phi_2 \leq 0$, equation (10) then implies that $q_t > 0$. But from equation (8), Φ_1 equals q_t at the optimum. Therefore, although

Φ_1 in general can be positive, negative, or zero when $I_t \leq 0$, it is strictly positive at the optimum. Equivalently, $G'(\cdot)$ is strictly positive at the optimum. ■

Investment and Stock Returns

Combining the first-order conditions in equations (8) and (9) yields:

$$\mathrm{E}_t[M_{t+1} r^I_{t+1}] = 1 \tag{14}$$

where r^I_{t+1} denotes the investment return:

$$r^I_{t+1} = \frac{\Pi_1(K_{t+1}, X_{t+1}) - \Phi_2(I_{t+1}, K_{t+1}) + (1-\delta)\Phi_1(I_{t+1}, K_{t+1})}{\Phi_1(I_t, K_t)} \tag{15}$$

The investment-return equation (15) is very intuitive. r^I_{t+1} can be interpreted as the ratio of the marginal benefit of investment at time $t+1$ divided by the marginal cost of investment at time t. The denominator, $\Phi_1(I_t, K_t)$, is the marginal cost of investment. By optimality, it equals the marginal q_t, the expected present value of marginal profits of investment. In the numerator of equation (15), $\Pi_1(K_{t+1}, X_{t+1})$ is the extra operating profit from the extra capital at $t+1$; $-\Phi_2(I_{t+1}, K_{t+1})$ captures the effect of extra capital on the augmented adjustment cost; and $(1-\delta)\Phi_1(I_{t+1}, K_{t+1})$ is the expected present value of marginal profits evaluated at time $t+1$, net of depreciation.

Proposition 2 ***(The Equivalence between Stock and Investment Returns)*** *Define stock return as:*

$$r^S_{t+1} \equiv \frac{P_{t+1} + \Pi(K_{t+1}, X_{t+1}) - \Phi(I_{t+1}, K_{t+1})}{P_t} \tag{16}$$

Then $\mathrm{E}_t[M_{t+1} r^S_{t+1}] = 1$*. Under Assumptions 1 and 3, stock return equals investment return:*

$$r^S_{t+1} = r^I_{t+1} \tag{17}$$

Proof. See Appendix B. ■

Given this equivalence, I will use the common notation r_{t+1} to denote both returns.

3 Confronting Anomalies

The equivalence between stock and investment returns is an extremely powerful result. It provides a rigorous, analytical link between expected returns and firm characteristics, a link that can serve as an economic foundation for understanding anomalies.

Developing this foundation is the heart of this paper. I first discuss in Section 3.1 the methodology of the Q-determination of expected returns, and its relation to the standard risk-based determination. Section 3.2 fixes the basic intuition using two canonical examples. And Section 3.3 extends the intuition into the more general Q-theoretical framework.

3.1 Methodology

My analytical methods are very simple. They basically amount to taking and signing partial derivatives of the expected investment return in equation (15) with respect to various anomaly-related variables. Using partial derivatives is reasonable because to establish a new anomaly, empiricists often control for other known anomalies, a practice corresponding naturally to partial derivatives.[7] Cochrane (1991, 1996) uses similar techniques to explain the return-investment relations. Similar methods are commonly used in empirical literature to develop testable hypotheses from valuation models (e.g., Fama and French (2004)).

As a more fundamental departure from traditional asset pricing methodology that determines expected returns through risk, I derive expected returns from firms' first-order conditions, as opposed to consumers'. As a result, expected returns are directly tied with firm characteristics. This approach to explaining anomalies is not new in itself. Previous examples include Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003).[8]

But I go one step further. I use the expected investment-return equation to analyze how expected return varies with characteristics and corporate policies. Intriguingly, the

[7]For example, Chan, Jegadeesh, and Lakonishok (1996) and Haugen and Baker (1996) control for valuation ratios when they document the earnings momentum and profitability anomalies, respectively.

[8]Rubinstein (2001, p. 23) highlights the importance of analyzing corporate decisions in solving anomalies: "For the most part, financial economists take the stochastic process of stock prices, the value of the firms, or dividend payments as primitive. But to explain some anomalies, we may need to look deeper into the guts of corporate decision making to derive what the processes are."

stochastic discount factor, M_{t+1}, and its covariances with returns (i.e., risk) do not enter the expected-return determination, at least directly. And firm characteristics are sufficient statistics for expected returns. I thus need not specify M_{t+1} as done in all previous studies.

However, this practice only means that the effect of M_{t+1} is indirect, not irrelevant. For example, if M_{t+1} were a constant, M, then equation (14) implies that the expected return $\mathrm{E}_t[r_{t+1}]=\frac{1}{M}$, a constant uncorrelated with any firm-level variables. For another example, if the correlation between M_{t+1} and X_{t+1} is zero, i.e., firms' operating profits are unaffected by aggregate shocks, then equation (14) implies that $\mathrm{E}_t[r_{t+1}] = r_{ft}$, where $r_{ft} \equiv \frac{1}{\mathrm{E}_t[M_{t+1}]}$ is risk-free rate. And there is no cross-sectional variation in expected returns. In this case, my analysis below in effect provides time-series correlations between the risk-free rate and firm characteristics. Since I study expected return directly, I need not restrict the correlation between M_{t+1} and X_{t+1}, the correlation that determines expected excess return.

My characteristic-based approach is consistent with the traditional risk-based approach. From Proposition 2, $\mathrm{E}_t[M_{t+1}r^S_{t+1}] = 1$. Following Cochrane (2001, p. 19), I can rewrite this equation as the beta-representation, $\mathrm{E}_t[r^S_{t+1}] = r_{ft} + \beta_t\lambda_{Mt}$, where $\beta_t \equiv \frac{-\mathrm{Cov}_t[r^S_{t+1}, M_{t+1}]}{\mathrm{Var}_t[M_{t+1}]}$ is the amount of risk, and $\lambda_{Mt} \equiv \frac{\mathrm{Var}_t[M_{t+1}]}{\mathrm{E}_t[M_{t+1}]}$ is the price of risk. Now Proposition 2 also says that $\mathrm{E}_t[r^S_{t+1}]=\mathrm{E}_t[r^I_{t+1}]$ where the right-hand-side only depends on characteristics from equation (15). Further, $\mathrm{E}_t[r^I_{t+1}]=\mathrm{E}_t[r^S_{t+1}]=r_{ft}+\beta_t\lambda_{Mt}$, implies that $\beta_t = \frac{\mathrm{E}_t[r^I_{t+1}]-r_{ft}}{\lambda_{Mt}}$. This ties covariances with characteristics. But apart from this mechanical link, risk only plays a secondary role in my characteristic-based determination of expected returns.

3.2 Intuition in Two Canonical Examples

I construct two canonical examples to illustrate the basic intuition underlying the anomalies explanations. Both examples have constant return to scale, $\alpha=1$. In the first example, the only costs of investment are linear purchase/sale costs, i.e., $\Phi(I_t, K_t)=I_t$. And in the second example, there are also quadratic costs of physical adjustment, i.e.,

$$\Phi(I_t, K_t) = I_t + \frac{a}{2}\left(\frac{I_t}{K_t}\right)^2 K_t \quad \text{where} \quad a > 0 \tag{18}$$

Linear Purchase/Sale Costs

This example can explain the earnings-related anomalies. Intuitively, the marginal product of capital (i.e., MPK) at time $t+1$ is in the numerator of investment return. But MPK is closely related to profitability, so expected return increases with expected profitability.

When $\Phi(I_t, K_t) = I_t$, equation (15) implies that:

$$\mathrm{E}_t[r_{t+1}] = \mathrm{E}_t[\Pi_1(K_{t+1}, X_{t+1})] + (1-\delta) \tag{19}$$

i.e., expected net return is expected marginal product of capital minus depreciation rate. Let $N_t \equiv \Pi_t - \delta K_t$ denote earnings (see footnote 6). Equations (2) and (19) imply that:

$$\mathrm{E}_t[r_{t+1}] = \mathrm{E}_t\left[\frac{\Pi_{t+1}}{K_{t+1}}\right] + (1-\delta) = \mathrm{E}_t\left[\frac{N_{t+1}}{K_{t+1}}\right] + 1 \tag{20}$$

i.e., expected return *is* expected profitability!

The example is also consistent with the profitability anomaly. Intuitively, profitability is highly persistent; therefore, high profitability gives rise to high expected profitability, which in turn leads to high expected returns.

The following assumption captures the persistence of profitability:

Assumption 4 *The operating profit-to-capital ratio (or equivalently profitability) follows:*

$$\frac{\Pi_{t+1}}{K_{t+1}} = \overline{\pi}(1-\rho_\pi) + \rho_\pi \left(\frac{\Pi_t}{K_t}\right) + \varepsilon^\pi_{t+1} \tag{21}$$

where $\overline{\pi} > 0$ *and* $0 < \rho_\pi < 1$ *are the long-run average and the persistence of operating profit-to-capital, respectively. And* ε^π_{t+1} *is a normal random variate with a zero mean.*

Since the operating profit-to-capital ratio equals profitability plus a constant depreciation rate, Assumption 4 basically says that profitability is persistent. The specific autoregressive form is unimportant. There is much evidence corroborative of Assumption 4. For example, Fama and French (2004) report that the current profitability has the strongest forecast power in predicting profitability one to three years ahead.

Combining equations (20) and (21) yields:

$$\mathrm{E}_t[r_{t+1}] = (\overline{\pi} - \delta)(1 - \rho_\pi) + \rho_\pi \left(\frac{N_t}{K_t}\right) + 1 \tag{22}$$

i.e., expected return is an increasing, linear function of profitability.

The same mechanism driving the expected-profitability and profitability anomalies is also useful for explaining the post-earnings-announcement drift that has bewildered financial economists for thirty-five years. Intuitively, earnings surprise and profitability are both scaled earnings, and should contain similar information on future profitability.[9] If earnings surprise captures a principal component of expected profitability as profitability does, then earnings surprise should correlate positively with expected returns.

Although useful for explaining the sign of the earnings-related anomalies, the simple example with $\Phi(I_t, K_t) = I_t$ has many limitations. First, the inverse relation between the magnitude of the earnings-related anomalies and the market value cannot be explained. From equations (20) and (22), the partial derivatives of expected return with respect to expected profitability and profitability are both constant, independent of the market value. Second, the example cannot explain the value anomaly because $\Phi(I_t, K_t) = I_t$ implies that $Q_t = q_t = \Phi_1(I_t, K_t) = 1$. Firms do not differ in market-to-book. Third, substituting $K_{t+1} = \left(\frac{I_t}{K_t} + (1-\delta)\right) K_t$ into equation (20) and differentiating both sides yield $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} = \mathrm{E}_t[\Pi_{11}(K_{t+1}, X_{t+1})]K_t = 0$, where the last equality follows from constant return to scale. This says that expected return is independent of investment rate, and hence independent of payout and financing rates.

Quadratic Adjustment Costs

I now show that all the limitations in the first example can be eliminated by introducing adjustment costs into the model. To fix intuition, I use a parametric example with quadratic

[9] To be precise, earnings surprise is commonly measured as Standardized Unexpected Earnings (SUE) (e.g., Chan, Jegadeesh, and Lakonishok (1996)). The SUE for stock i in month t is defined as $\mathrm{SUE}_{it} \equiv \frac{e_{iq} - e_{iq-4}}{\sigma_{it}}$, where e_{iq} is quarterly earnings per share most recently announced as of month t for stock i, e_{iq-4} is earnings per share four quarters ago, and σ_{it} is the standard deviation of unexpected earnings, $e_{iq} - e_{iq-4}$, over the preceding eight quarters.

adjustment costs, i.e., $\Phi(I_t, K_t)$ follows equation (18). Then equation (15) implies that:

$$r_{t+1} = \frac{\Pi_1(K_{t+1}, X_{t+1}) + (a/2)(I_{t+1}/K_{t+1})^2 + (1-\delta)[1 + a(I_{t+1}/K_{t+1})]}{1 + a(I_t/K_t)} \tag{23}$$

The Earnings Anomalies Since $\Pi_1(K_{t+1}, X_{t+1}) = \frac{\Pi_{t+1}}{K_{t+1}} = \frac{N_{t+1}}{K_{t+1}} + \delta$, taking conditional expectations and differentiating both sides of equation (23) with respect to expected profitability yield $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial \mathrm{E}_t[N_{t+1}/K_{t+1}]} = \frac{1}{1+a(I_t/K_t)} > 0$. The inequality results from its denominator, $\Phi_1(I_t, K_t) = q_t$, which is positive by Lemma 3. Therefore, controlling for market-to-book (effectively the denominator of investment return), expected return increases with expected profitability. And because the marginal q_t equals market-to-book, $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial \mathrm{E}_t[N_{t+1}/K_{t+1}]} = \frac{K_{t+1}}{P_t}$, which is inversely related with the market value, P_t. This explains why the earnings-related anomalies are stronger in small firms.

The Investment Anomaly Intuitively, the Q-theory is a theory of investment demand. And the downward-sloping investment-demand function naturally implies a negative relation between investment rate and cost of capital (i.e., expected return). In essence, investment rate increases with net present value of capital (e.g., Brealey and Myers (2003, Chapter 2)). But the net present value is inversely related to cost of capital. Higher cost of capital implies lower expected net present value, which in turn implies lower investment rate, and vice versa.

I now formally demonstrate the negative slope of the investment-demand function using equation (23). Let $U^q_{t+1} > 0$ denote the numerator of the investment return. Taking conditional expectations and differentiating both sides with respect to $\frac{I_t}{K_t}$ yield: $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} = -\frac{a\mathrm{E}_t[U^q_{t+1}]}{[1+a(I_t/K_t)]^2} + \frac{1}{1+a(I_t/K_t)} \frac{\partial \mathrm{E}_t[U^q_{t+1}]}{\partial (I_t/K_t)}$. To show $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} < 0$, it then suffices to show $\frac{\partial \mathrm{E}_t[U^q_{t+1}]}{\partial (I_t/K_t)} < 0$. But rewriting I_{t+1} and K_{t+1} in $\mathrm{E}_t[U^q_{t+1}]$ as $K_{t+2} - (1-\delta)\left(\frac{I_t}{K_t} + (1-\delta)\right)K_t$ and $\left(\frac{I_t}{K_t} + (1-\delta)\right)K_t$, respectively, and differentiating yield $\frac{\partial \mathrm{E}_t[U^q_{t+1}]}{\partial (I_t/K_t)} = -\frac{aK_t(\mathrm{E}_t[K_{t+2}])^2}{K^3_{t+1}} < 0$.[10]

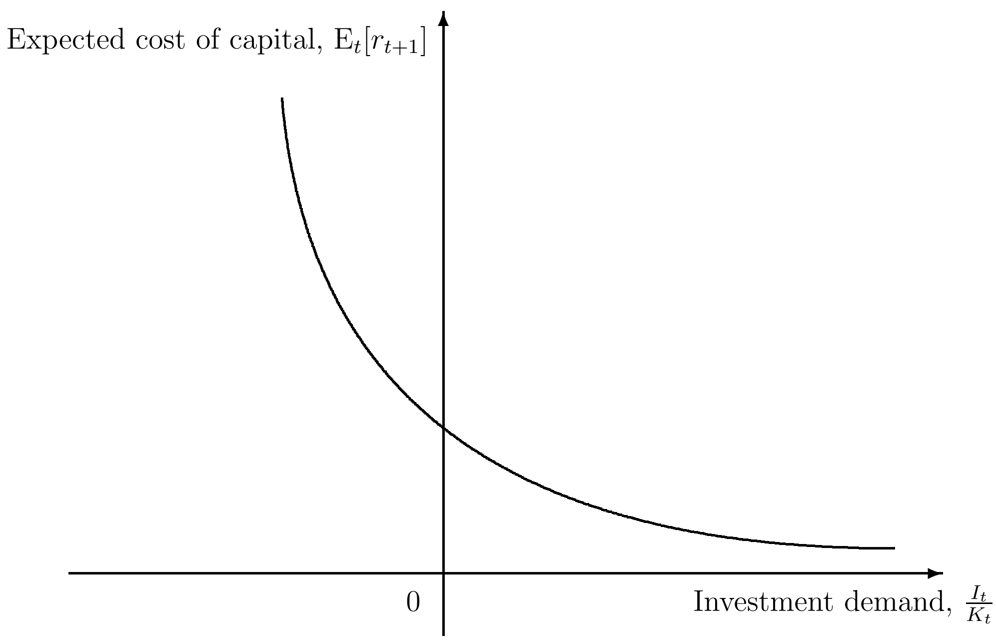
Figure 1 plots the investment-demand function in this example with quadratic adjustment costs. The function is downward-sloping, implying that expected return decreases with positive investment, but increases with the magnitude of disinvestment.

[10] I have used Leibniz integral rule to change the order of integration and differentiation.

From Figure 1, the investment-demand function is also convex. To see this, differentiating $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}$ once more with respect to $\frac{I_t}{K_t}$ yields $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)^2} = -\frac{a}{[1+a(I_t/K_t)]^2} \frac{\partial \mathrm{E}_t[U^q_{t+1}]}{\partial(I_t/K_t)} + \frac{2a^2 \mathrm{E}_t[U^q_{t+1}]}{[1+a(I_t/K_t)]^3} + \frac{1}{1+a(I_t/K_t)} \frac{\partial^2 \mathrm{E}_t[U^q_{t+1}]}{\partial(I_t/K_t)^2} - \frac{\partial \mathrm{E}_t[U^q_{t+1}]}{\partial(I_t/K_t)} \frac{a}{[1+a(I_t/K_t)]^2} > 0$, where the inequality follows from $\frac{\partial \mathrm{E}_t[U^q_{t+1}]}{\partial(I_t/K_t)} < 0$ and $\frac{\partial^2 \mathrm{E}_t[U^q_{t+1}]}{\partial(I_t/K_t)^2} = \frac{3aK_t^2(\mathrm{E}_t[K_{t+2}])^2}{K^4_{t+1}} > 0$.

Figure 1. The Downward-Sloping and Convex Investment-Demand Function

This figure plots the downward-sloping and convex investment-demand function. The horizontal axis is investment demand measured by the investment-to-capital ratio, $\frac{I_t}{K_t}$. And the vertical axis is expected cost of capital, $\mathrm{E}_t[r_{t+1}]$.



Titman, Wei, and Xie (2003) show that the investment anomaly is stronger in firms with high operating income-to-asset ratios. This stylized fact can be captured in the example. Using equation (21) to express $\frac{\Pi_{t+1}}{K_{t+1}}$ in terms of $\frac{\Pi_t}{K_t}$ and differentiating $-\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}$ with respect to $\frac{\Pi_t}{K_t}$ yields $\partial \left| \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)} \right| / \partial \left(\frac{\Pi_t}{K_t} \right) = \frac{a\rho_\pi}{[1+a(I_t/K_t)]^2} > 0$.

The Value Anomaly The downward-sloping and convex investment-demand function manifests itself as many anomalies other than the investment anomaly. The value anomaly can be explained using the investment-demand function. From the optimality condition (8), $1+a\frac{I_t}{K_t} = q_t = Q_t$, so $\frac{\partial(I_t/K_t)}{\partial Q_t} = \frac{1}{a}$. The chain rule of partial derivatives then implies that $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} = \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)} \frac{\partial(I_t/K_t)}{\partial Q_t} < 0$: growth firms with high market-to-book earn lower average returns than value firms with low market-to-book.

And the value anomaly is stronger in small firms. To see this, again by the chain rule, $\partial \left|\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\right| / \partial P_t = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t \partial P_t} = -\frac{1}{a}\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)^2}\frac{\partial(I_t/K_t)}{\partial P_t}$. To show the left-hand-side is negative, the convexity in the investment-demand function implies that it suffices to show $\frac{\partial(I_t/K_t)}{\partial P_t} > 0$. But from $1+a\frac{I_t}{K_t} = q_t = Q_t = \frac{P_t}{K_{t+1}}$, $P_t = \left[1+a\frac{I_t}{K_t}\right]\left[\frac{I_t}{K_t} + (1-\delta)\right] K_t$. Differentiating both sides with respect to $\frac{I_t}{K_t}$ yields $\frac{\partial P_t}{\partial(I_t/K_t)} = q_t K_t + a K_{t+1} > 0$.[11]

The Payout and SEO-Underperformance Anomalies The payout and SEO anomalies can also be explained using the investment-demand function. Intuitively, firms' cash-flow constraint says that the sources and the uses of funds must be equal. With quadratic adjustment costs, when free cash flow $C_t > 0$, this constraint is $\frac{C_t}{K_t} = \frac{\Pi_t}{K_t} - \frac{I_t}{K_t} - \frac{a}{2}\left(\frac{I_t}{K_t}\right)^2$, and when outside equity $O_t > 0$, this constraint is $\frac{O_t}{K_t} = \frac{I_t}{K_t} + \frac{a}{2}\left(\frac{I_t}{K_t}\right)^2 - \frac{\Pi_t}{K_t}$. As a result, $\frac{\partial(C_t/K_t)}{\partial(I_t/K_t)} = -\left(1+a\frac{I_t}{K_t}\right) = -q_t < 0$, and $\frac{\partial(O_t/K_t)}{\partial(I_t/K_t)} = q_t > 0$. In other words, controlling for profitability, optimal payout and investment rates are negatively correlated, and optimal equity-financing and investment rates are positively correlated. By the chain rule, the negative slope of the investment-demand function then manifests itself as the positive expected return-payout relation (i.e., $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(C_t/K_t)} = \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\frac{\partial(I_t/K_t)}{\partial(C_t/K_t)} > 0$) and the negative expected return-financing relation (i.e., $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(O_t/K_t)} = \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\frac{\partial(I_t/K_t)}{\partial(O_t/K_t)} < 0$).

And the convexity of the investment-demand function manifests itself as the stronger payout anomaly in value firms (i.e., $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(C_t/K_t)\partial Q_t} = \frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)^2}\frac{\partial(I_t/K_t)}{\partial Q_t}\frac{\partial(I_t/K_t)}{\partial(C_t/K_t)} < 0$) and the

[11] $\frac{\partial P_t}{\partial(I_t/K_t)} > 0$ and $\frac{\partial Q_t}{\partial(I_t/K_t)} > 0$ both imply that growth firms invest more and grow faster. $\frac{\partial P_t}{\partial(I_t/K_t)} > 0$ does not contradict the evidence that small firms invest more and grow faster than big firms (e.g., Evans (1987) and Hall (1987)). The evidence is documented with the logarithm of employment as the measure of firm size. This measure corresponds to $\log(K_t)$ in the model. The model is consistent with the evidence because $P_t = \left[1+a\frac{I_t}{K_t}\right]\left[\frac{I_t}{K_t} + (1-\delta)\right] K_t$ implies that $K_t = P_t\left(\left[1+a\frac{I_t}{K_t}\right]\left[\frac{I_t}{K_t} + (1-\delta)\right]\right)^{-1}$, which in turn implies that $\frac{\partial K_t}{\partial(I_t/K_t)} < 0$.

stronger SEO-underperformance in small firms (i.e., $\partial\left|\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(O_t/K_t)}\right|/\partial P_t = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(O_t/K_t)\partial P_t} = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)^2}\frac{\partial(I_t/K_t)}{\partial P_t}\frac{\partial(I_t/K_t)}{\partial(O_t/K_t)} - \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\frac{\partial^2(I_t/K_t)}{\partial P_t \partial(O_t/K_t)} < 0$). The last inequality follows because $\frac{\partial(O_t/K_t)}{\partial(I_t/K_t)} = q_t = Q_t = \frac{P_t}{K_{t+1}}$ implies that $\frac{\partial^2(I_t/K_t)}{\partial(O_t/K_t)\partial P_t} = -\frac{K_{t+1}}{P_t^2} < 0$.

3.3 The General Model

I now extend the results in the last subsection into the general Q-theoretical framework. Except for a few technical details, all the intuition obtained in the quadratic-adjustment-cost example holds in the general setup.

To simplify exposition, I first prove the following lemma.

Lemma 4 *Define the numerator of investment return:*

$$U_{t+1} \equiv \Pi_1(K_{t+1}, X_{t+1}) - \Phi_2(I_{t+1}, K_{t+1}) + (1-\delta)\Phi_1(I_{t+1}, K_{t+1}) \tag{24}$$

Under Assumptions 1–3, $U_{t+1} > 0$, *and the (gross) returns are positive,* $r_{t+1} = \frac{U_{t+1}}{\Phi_1(I_t, K_t)} > 0$.

Proof. $\Pi_1 > 0$ and $\Phi_2 \leq 0$ follow from Assumptions 1 and 2, respectively. And $\Phi_1 = q_t > 0$ follows from Lemma 3. ■

The Investment Anomaly

Proposition 3 *Under Assumptions 1–3, expected returns decrease with investment rate:* $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)} < 0$. *If Assumption 4 also holds, the magnitude of the investment anomaly increases with operating profit-to-asset:* $\partial\left|\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\right|/\partial\left(\frac{\Pi_t}{K_t}\right) > 0$.

Proof. See Appendix B. ■

Cochrane (1991, 1996) uses a parametric model with quadratic adjustment costs to derive the negative relation between expected return and investment rate. He also shows that the relation between expected return and expected investment rate and the relation between expected return and expected investment growth are both positive. I extend his argument into the general Q-framework, and use it as a foundation to explain other anomalies.

To derive the positive relations of expected returns with expected investment rate and expected investment growth, note that equation (15) implies that $\frac{\partial r_{t+1}}{\partial(I_{t+1}/K_{t+1})} =$

$\frac{K_t^{\alpha-1}}{q_t}\left[(1-\alpha)G'\left(\frac{I_{t+1}}{K_{t+1}}\right)\right]+\frac{K_t^{\alpha-1}}{q_t}\left[G''\left(\frac{I_{t+1}}{K_{t+1}}\right)\left(\frac{I_{t+1}}{K_{t+1}}+(1-\delta)\right)\right]$. The first term in the right-hand side is nonnegative because $1-\alpha \geq 0$. And the second term is strictly positive because $\frac{I_{t+1}}{K_{t+1}}+(1-\delta)=\frac{K_{t+2}}{K_{t+1}}>0$ and $G''>0$. Given that expected return decreases with $\frac{I_t}{K_t}$ but increases with $\frac{I_{t+1}}{K_{t+1}}$, it should also increase with $\frac{I_{t+1}}{K_{t+1}}/\frac{I_t}{K_t}$. But as a ratio of two stock variables, $\frac{K_t}{K_{t+1}}$ is likely to be dominated by the ratio of two flow variables, $\frac{I_{t+1}}{I_t}$. Therefore, expected return should increase with expected investment growth.

Most models cited in footnote 4 can explain the investment anomaly, including the real options models of Berk, Green, and Naik (1999) and Carlson, Fisher, and Giammarino (2004a). I contribute by unifying their explanations with the investment-return equation, by using this equation to explain other anomalies, and by illustrating the interaction of the return-investment relation with the operating income-to-asset ratio.

The Value Anomaly

In the general model, the investment rate is an increasing function of the marginal q. The optimality condition (8) implies that $G'\left(\frac{I_t}{K_t}\right)K_t^{\alpha-1}=q_t$ or $\frac{I_t}{K_t}=G'^{-1}(q_t K_t^{1-\alpha})$, where $G'^{-1}(\cdot)$ is the inverse function of $G'(\cdot)$. And since $G''>0$, both $G'(\cdot)$ and $G'^{-1}(\cdot)$ are increasing functions. Because the marginal q is in turn proportional to market-to-book, the negative slope of the investment-demand function then gives rise to the negative relation between expected return and market-to-book.

To explain the stronger value anomaly in small firms, I need one more assumption:

Assumption 5 *The augmented adjustment-cost function* $\Phi(I_t, K_t)$ *satisfies:*

$$\Phi_{111}(I_t, K_t) \geq 0; \quad \Phi_{112}(I_t, K_t) \leq 0; \quad \text{and} \quad \Phi_{222}(I_t, K_t) \leq 0$$

It is easy to verify that the standard specifications such as quadratic adjustment-cost function and the proportional and quadratic financing-cost functions satisfy this assumption.

Proposition 4 *Under Assumptions 1–3, expected returns correlate negatively with market-to-book,* $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}<0$. *If Assumption 5 also holds, the magnitude of this correlation decreases in the market value,* $\partial\left|\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\right|/\partial P_t<0$.

Proof. See Appendix B. ■

Some new testable hypotheses are collected below.

Proposition 5 *Under Assumptions 1–3 and 5: (i) the relation between expected return and market-to-book is convex,* $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} > 0$*; (ii) the investment anomaly is stronger in small firms,* $\partial \left| \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} \right| / \partial P_t < 0$*; and (iii) the investment anomaly is stronger in value firms.,* $\partial \left| \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} \right| / \partial Q_t < 0$*. If* $2 \left[G'' \left(\frac{I_t}{K_t} \right) \right]^2 \geq G''' \left(\frac{I_t}{K_t} \right) G' \left(\frac{I_t}{K_t} \right)$ *also holds, then (iv) the investment-demand function is convex, i.e.,* $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)^2} > 0$*.*

Proof. See Appendix B. ■

Proposition 5 says that the convexity in the investment-demand function holds in the general model only if $2 \left[G'' \left(\frac{I_t}{K_t} \right) \right]^2 \geq G''' \left(\frac{I_t}{K_t} \right) G' \left(\frac{I_t}{K_t} \right)$. Although compatible with quadratic adjustment costs, this condition is more strict than the other assumptions.

Several recent studies have proposed rational explanations for the value anomaly using investment-based models. Berk, Green, and Naik (1999) construct a real options model in which endogenous changes in assets in place and in growth options impart predictability in returns. Also using real options models, Carlson, Fisher, and Giammarino (2004a) emphasize the role of operating leverage, and Cooper (2004) emphasize the role of fixed costs in driving the value anomaly. Gomes, Kogan, and Zhang (2003) use a dynamic general equilibrium production economy, Kogan (2004) uses a two-sector general equilibrium model, and Zhang (2004) uses a Q-model of optimal investment to link expected returns to firm characteristics.

My model is most related to Zhang's (2004). By making Assumptions 1–3, I now obtain some analytical results. The scope of anomalies addressed is also much wider. Zhang offers an explicitly solved model in which the degree of homogeneity of the operating-profit function is not equal to that of the adjustment-cost function. And his simulation results are consistent with my analytical results. Importantly, I am not aware of previous explanations of the inverse relation between the value anomaly and the market value.

The Payout Anomaly

In the general model, firms' cash-flow constraint becomes $\frac{C_t}{K_t} = \frac{\Pi_t}{K_t} - G\left(\frac{I_t}{K_t}\right) K_t^{\alpha-1}$ when payout $C_t > 0$. Controlling for profitability, optimal payout and investment rates are again negatively correlated. This negative correlation, coupled with the downward-sloping investment-demand function, implies a positive expected return-payout relation. And by the chain rule, the convexity of the relation between expected return and market-to-book gives rise to the stronger payout anomaly in value firms.

Proposition 6 *Denote payout or free cash flow as:*

$$C_t \equiv C(K_t, K_{t+1}, X_t) = (\Pi(K_t, X_t) - \Phi(I_t, K_t))\, \mathbf{1}_{\{\Pi(K_t,X_t)-\Phi(I_t,K_t)>0\}} \tag{25}$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function that takes the value of one if the event described in $\{\cdot\}$ is true and zero otherwise. Under Assumptions 1–3, expected returns increase weakly with the payout rate, $\frac{C_t}{K_t}$, i.e., $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (C_t/K_t)} \geq 0$, where the inequality is strict when C_t is strictly positive. If Assumption 5 also holds, the payout anomaly is stronger in value firms than that in growth firms, $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial (C_t/K_t) \partial Q_t} \leq 0$, where the inequality is strict when $C_t > 0$.

Proof. See Appendix B. ■

I am not aware of other rational explanations for the payout anomaly.

The SEO-Underperformance Anomaly

Firms' cash-flow constraint is $\frac{O_t}{K_t} = G\left(\frac{I_t}{K_t}\right) K_t^{\alpha-1} - \frac{\Pi_t}{K_t}$ when outside equity $O_t > 0$. Controlling for profitability, optimal financing and investment rates are positively correlated. This correlation, coupled with the downward-sloping investment-demand function, implies a negative expected return-financing relation. By the chain rule, the convex relation between expected return and market-to-book implies the stronger SEO anomaly in small firms.

Proposition 7 *Denote the outside equity, O_t, as:*

$$O_t \equiv O(K_t, K_{t+1}, X_t) = (\Phi(I_t, K_t) - \Pi(K_t, X_t))\, \mathbf{1}_{\{\Phi(I_t,K_t)-\Pi(K_t,X_t)>0\}} \tag{26}$$

Under Assumptions 1 and 3, expected returns decrease weakly with the rate of external or outside equity, $\frac{O_t}{K_t}$*, i.e.,* $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (O_t/K_t)} \leq 0$*, where the inequality is strict when* O_t *is strictly positive. If Assumption 5 also holds, the magnitude of this correlation is stronger in small firms,* $\partial \left| \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (O_t/K_t)} \right| / \partial P_t \leq 0$*, where the inequality is strict when* $O_t > 0$*.*

Proof. See Appendix B. ■

Some corroborative evidence is provided by Loughran and Ritter (1997) and Richardson and Sloan (2003). Loughran and Ritter shows that issuing firms have much higher investment rates measured as capital expenditure plus R&D divided by assets than nonissuing firms for nine years around the issuing date. And issuing firms are disproportionately high-growth firms. That issuing firms invest more than nonissuing firms is reflected in equation (B24).

Richardson and Sloan (2003) find that the negative relation between external finance and expected returns varies systematically with the use of the proceeds. When the proceeds are invested in net operating assets as opposed to being stored as cash, the negative relation is stronger. In contrast, the negative relation is much weaker when the proceeds are used for refinancing or retained as cash. This evidence suggests the important role of capital investment in driving the SEO anomaly. This is exactly my theoretical approach.

However, Ritter (2003) interprets the SEO underperformance as suggesting a window-of-opportunity theory of financing decisions. Richardson and Sloan (2003) interpret their results as "strong and pervasive (p. 34)" evidence of inefficient markets. But I think otherwise.

And the stronger value anomaly in small firms and the stronger SEO anomaly in small firms are basically the same phenomenon, driven by the convex relation between expected return and market-to-book. This prediction is consistent with the evidence that the SEO underperformance reduces greatly once both size and book-to-market are controlled for (e.g., Brav, Geczy, and Gompers (2000) and Eckbo, Masulis, and Norli (2000)).

Several recent studies have proposed rational explanations of the SEO anomaly. Eckbo, Masulis, and Norli (2000) argue that issuing firms are less risky because their leverage ratios are lowered. There is no leverage in my model, and the economic mechanism works through optimal investment. Schultz (2003) argues that using event studies is likely to find negative abnormal performance ex post, even if there is no abnormal performance ex ante.[12] But the calendar time evidence is immune to this problem. Importantly, the Q-explanation applies to calendar-time as well as event-time underperformance.

Carlson, Fisher, and Giammarino (2004b) argue that that prior to issuance, the firm has both assets in place and an option to expand; this composition is a levered, risky position. If the exercise of the option is financed by equity issuance, then risk must drop afterwards. This real options explanation and my Q-explanation are related because they both work through optimal investment. But I do not assume growth options to be riskier than assets in place, although it is likely to be true in good times when the option to expand is important.

The Expected-Profitability Anomaly

Proposition 8 *Under Assumptions 1 and 3, expected returns correlate positively with expected profitability,* $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial \mathrm{E}_t[N_{t+1}/K_{t+1}]} > 0$*, and the magnitude of the correlation decreases with the market value,* $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial \mathrm{E}_t[N_{t+1}/K_{t+1}]\partial P_t} < 0$.

Proof. From equation (2), $\Pi_1(K_t, X_t) = \alpha\left(\frac{\Pi_t}{K_t}\right) = \alpha\left(\frac{N_t}{K_t} + \delta\right)$. From equation (15),

$$\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial \mathrm{E}_t[N_{t+1}/K_{t+1}]} = \frac{\alpha}{\Phi_1(I_t, K_t)} = \frac{\alpha}{q_t} > 0$$

Also recall $q_t = \alpha \frac{P_t}{K_{t+1}}$ from equation (12),

$$\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial \mathrm{E}_t[N_{t+1}/K_{t+1}]\partial P_t} = -\frac{K_{t+1}}{P_t^2} < 0$$

■

The Profitability Anomaly

Proposition 9 *Under Assumptions 1–4, expected returns correlate positively with current-period profitability, and the magnitude of this correlation decreases with the market value,*

[12] Schutlz (2003) uses his argument to explain the underperformance of initial public offerings (IPOs). The same logic applies to SEOs. If early in a sample period, SEOs underperform, there will be few SEOs in the future because investors are less interested in them. The average performance will be weighted more towards the early SEOs that underperformed. If early SEOs outperform, there will be many more SEOs in the future. The early performance will be weighted less in the average performance. Weighting each period equally as in calendar-time regressions solves this problem.

i.e., $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (N_t/K_t)} > 0$ *and* $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial (N_t/K_t)\partial P_t} < 0$.

Proof. From equation (15) and Assumption 4,

$$\begin{aligned}\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (N_t/K_t)} &= \frac{\alpha \rho_\pi}{\Phi_1(I_t, K_t)} > 0 \\ \frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial (N_t/K_t)\partial P_t} &= \partial\left(\frac{\rho_\pi K_{t+1}}{P_t}\right)/\partial P_t = -\frac{\rho_\pi K_{t+1}}{P_t^2} < 0\end{aligned} \tag{27}$$

■

A new testable hypothesis emerges from equation (27): The magnitude of the earnings-related anomalies should increase with the persistence of profitability.

The Post-Earnings-Announcement Drift

Proposition 9 is useful for explaining the post-earnings-announcement drift if earnings surprise and profitability contain similar information on future profitability. And the second inequality in Proposition 9 is intriguing because the magnitude of the post-earnings-announcement drift is inversely related to the market value (e.g., Bernard (1993)).

However, I caution the readers that, unlike most of the other anomalies, earnings surprise does not appear directly in investment return. Using Proposition 9 to explain earnings momentum relies on an implicit assumption that earning surprise, like profitability, is a strong, positive predictor of future profitability, a premise that should be tested thoroughly.

I am not aware of other rational explanations of the earnings-related anomalies. Two papers offer explanations for a related anomaly, price momentum that buying recent winners and selling recent losers yield positive abnormal returns (e.g., Jegadeesh and Titman (1993)). In Berk, Green, and Naik (1999), the composition and systematic risk of the firm's assets are persistent, leading to positive autocorrelations of expected returns. In Johnson (2002), recent winners have temporarily higher expected growth than recent losers. Assuming that stocks with higher expected growth earn higher average returns, Johnson shows that his model can generate price momentum. I complement his work because Proposition 8 shows that his key assumption arises naturally from the first principles of optimal investment.

4 Empirical Implications

This section discusses empirical implications of my theoretical results. Section 4.1 proposes the Q-representation of expected returns as a new empirical asset pricing framework. Sections 4.2, 4.3, and 4.4 compare the Q-framework with the standard beta- and SDF-framework as well as the Ohlson (1995) valuation model popular in accounting research, respectively. I argue that, although internally consistent with the existing frameworks in theory, the Q-framework is likely to have comparative advantages in practice.

4.1 A New Asset Pricing Test

By Proposition 2, if the operating-profit and the augmented adjustment-cost functions have the same degree of homogeneity, stock return equals investment return. Ex ante, this implies that expected stock returns equal expected investment returns.

This ex ante restriction can be tested using the following moment conditions:

$$\mathrm{E}\left[\left(r^S_{t+1} - \frac{\Pi_1(K_{t+1}, X_{t+1}) - \Phi_2(I_{t+1}, K_{t+1}) + (1-\delta)\Phi_1(I_{t+1}, K_{t+1}))}{\Phi_1(I_t, K_t)}\right) \otimes \mathbf{Z}_t\right] = 0 \qquad (28)$$

where $\mathbf{Z}_t$ is a vector of instrumental variables known at time t; such as anomaly-related variables. r^S_{t+1} is the stock returns of portfolios sorted by the anomaly-related variables. Φ can be properly parameterized, as in the investment literature (e.g., Hubbard (1998) and Erickson and Whited (2000)). With the estimated parameters, expected stock returns can be constructed from economic fundamentals through the expected investment returns.[13]

4.2 The Q-Framework versus the Beta-Framework

The Q-framework helps dissipate one myth in the beta-framework, a popular belief that only covariances matter in rational markets (e.g., Daniel and Titman (1997)).

The beta-framework is most popular in empirical finance. In event studies, cumulative abnormal returns are computed as the difference between realized returns and expected returns from, for example, the CAPM. Tests are performed to see if cumulative abnormal

[13] Cochrane (1991) implicitly tests the moment condition (28) by comparing the properties of stock and investment returns, but both at the aggregate level.

returns are on average zero (e.g., Ball and Brown (1968) and Fama, Fisher, Jensen, and Roll (1969)). In cross-sectional tests, stock returns are regressed on beta and firm characteristics (e.g., Fama and French (1992)). In time-series tests, mimicking portfolios are formed based on characteristics of interest. Their average returns are tested to see if they equal zero, and zero-intercept tests are performed by regressing the portfolio returns on a set of benchmark factor returns (e.g., Fama and French (1993)).

The null hypotheses in these tests are derived from the CAPM and its various extensions, either static or conditional, single-factor or multi-factor models. Similar to the CAPM, all these extensions say that only covariances should explain expected returns. Anomalies emerge because characteristics often dominate covariances in explaining returns. That only covariances matter is the basic point of Daniel and Titman (1997), and is taken as "one general feature of the rational approach" in Barberis and Thaler (2003, p. 1091).

Not true. My results clearly show that characteristics can affect expected returns, often in the directions reported in the anomalies literature. But my model is entirely rational. Therefore, the debate on covariances vs. characteristics is not a well-defined question. And rejecting the CAPM and its close cousins is not equivalent to rejecting efficient markets (e.g., Fama (1970)) or rational expectations (e.g., Muth (1961) and Lucas (1972)).

Further, the Q-representation is perhaps more useful in practice than the beta-representation. The reason is that the right-hand-side of the Q-representation contains only firm characteristics. And measuring characteristics basically amounts to loading and cleaning data from Compustat, much easier than measuring covariances.

Measuring covariances is tricky. First, consumption-based asset pricing has not settled on the right form of the SDF, with which returns are supposed to covary. Second, all dynamic models imply that covariances are time-varying. But despite recent theoretical efforts, no easy-to-implement econometric specifications have been derived. And estimates of time-varying covariances in practice often use convenient, but ad hoc specifications yielding results sensitive to alternative methods. Finally, even if we assume the priced common factors are known and covariances are constant, estimates of expected returns from beta-pricing models are extremely imprecise even at the industry level (e.g., Fama and French (1997)).

The difficulty of measuring covariances is also illustrated in Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003). Covariances are sufficient statistics of expected returns in these models. In their simulations, true covariances indeed dominate characteristics in cross-sectional regressions. But when estimated covariances are used, they are dominated by characteristics. However, true covariances are unobservable in practice.

In sum, it is perhaps time to explore alternative empirical models of expected returns without having to estimate covariances, for example, along the lines of the Q-representation.

4.3 The Q-Framework versus the SDF-Framework

The Q-framework complements the stochastic discount factor framework that is the workhorse in consumption-based asset pricing (e.g., Campbell (2003)).

Tests under the SDF-framework are usually done by using $\mathrm{E}_t[M_{t+1} r^S_{t+1}] = 1$ as moment conditions in GMM, where M_{t+1} is the SDF that can be parameterized using data on aggregate consumption (e.g., Hansen and Singleton (1982)). This framework has had various degrees of success in understanding anomalies (e.g., Lettau and Ludvigson (2001), Lustig and Nieuwerburgh (2004), and Piazzesi, Schneider, and Tuzel (2004)).

But the SDF-framework still leaves plenty of room open. Most important, anomalies are empirical relations between expected returns and firm characteristics. But characteristics do not enter the moment conditions directly. They are buried in r^S_{t+1}, which is portfolio returns sorted on characteristics. Further, even if the moment conditions survive over-identification tests, it is not clear what economic mechanisms drive the results. For example, the empirical success of Lettau and Ludvigson (2001) and Lustig and Nieuwerburgh (2004) relies on the returns of value stocks covarying more with the price of risk in bad times than the returns of growth stocks. But why this occurs can perhaps be better understood by modeling expected returns and firm characteristics together, for example, in the Q-theoretical framework.

For another example, the important contributions of Bansal, Dittmar, and Lundblad (2004) and Campbell and Vuolteenaho (2004) show that the value anomaly can be explained because value stocks have higher cash-flow betas with higher price of risk than growth stocks. But the underlying economic mechanism is unknown because firm dynamics are not modeled.

Characteristics do enter the SDF framework directly in Cochrane (1996), Gomes, Yaron, and Zhang (2004), and Whited and Wu (2004). In Cochrane and Gomes et al., the SDF is a linear function of aggregate investment return constructed using aggregate fundamentals. But firm characteristics are absent. White and Wu test $\mathrm{E}_t[M_{t+1} r^I_{t+1}] = 1$ where M_{t+1} is a linear combination of the Fama-French (1993) factors. Firm fundamentals show up in constructing firm-level investment returns, r^I_{t+1}. But stock returns are now absent.

The most attractive feature of the Q-representation of expected return is that it does not involve SDF directly. This implies that production-based asset pricing can in principle be developed independently from consumption-based asset pricing, without being hindered by difficulties specific to the latter literature.

In all, by linking expected returns directly to characteristics in a rigorous, yet easy-to-implement framework, the Q-theory can cover the grounds missing from the SDF framework. In terms of the big picture, Fama (1991, p. 1610) calls for a coherent story that "(1) relates the cross-section properties of expected returns to the variation of expected returns through time, and (2) relates the behavior of expected returns to the real economy in a rather detailed way". Consumption-based asset pricing is naturally fit for the first goal, but production-based asset pricing is perhaps better equipped to achieve the second. And the coherent story envisioned by Fama can be provided by the conceptual framework of general equilibrium.

4.4 The Q-framework versus the Ohlson Framework

The Q-framework is related to Ohlson's (1995) residual income valuation model that is extremely popular in capital markets research in accounting (e.g., Frankel and Lee (1998), Dechow, Hutton, and Sloan (1999), and Kothari (2001)).

The valuation model says that:

$$\frac{P_t}{B_t} = 1 + \frac{\sum_{j=1}^{\infty} \mathrm{E}[Y_{t+j} - rB_{t+j-1}]/(1+r)^j}{B_t} \tag{29}$$

where B_t is book equity at time t, Y_{t+j} is earnings at $t+j$, $Y_{t+j} - rB_{t+j-1}$ is the residual income, defined as the difference between earnings and the opportunity cost of capital, and r is the discount rate for the expected residual income or the long-term expected stock return.

The model has several predictions.[14] First, controlling for expected residual earnings and expected book equity relative to current book equity, a higher book-to-market equity implies a higher expected return. Second, given book-to-market, firms with higher expected residual income relative to current book equity have higher expected returns. Third, controlling for book-to-market and the expected growth in book equity or investment growth, more profitable firms or firms with higher expected earnings relative to current book equity have higher expected returns. Finally, given book-to-market and expected profitability, firms with higher expected growth in book equity have lower expected returns.

These predictions are largely consistent with the predictions of the Q-model, implying that the Q-model is potentially useful in guiding empirical capital markets research.

But there is one notable difference. The Ohlson model says the high expected investment growth leads to low expected returns, but the Q-model says otherwise. Liu, Warner, and Zhang (2003) find that firms with higher expected investment growth earn higher average returns than firms with lower expected investment growth, but the average-return difference is only marginally significant. Further tests can sort out these two competing hypotheses.

Several recent papers use valuation models to estimate expected returns (e.g., Claus and Thomas (2001) and Gebhardt, Lee, and Swaminathan (2001)). They find that the estimated equity premium is only about 2–3%, much lower than the historical average. Kothari (2001) argues that their long-term growth forecasts, especially the terminal perpetuity growth rates, seems too low. In contrast, estimating expected returns from the Q-model only requires inputs of the one-period-ahead profitability and investment rate, which should be easier to forecast than their long-term counterparts.

More important, valuation models are tautological accounting models. In contrast, derived from the first principles, the Q-model is an economic model that can be used to generate economically interesting hypotheses linking expected returns to the real economy in a rather detailed way.

[14] My discussion on the predictions from the Ohlson model follows that of Fama and French (2004).

5 Conclusion

A voluminous literature on capital markets anomalies in financial economics has mounted an enormous challenge to efficient markets with rational expectations. These anomalies are empirical relations of future stock returns with firm characteristics and corporate policies and events, relations not predicted by current asset pricing models.

Using a neoclassical model with rational expectations, I demonstrate analytically that, much like aggregate expected returns that vary with business cycles, expected returns in the cross-section vary with corporate policies or events, often in the directions documented in the anomalies literature. These anomalies include the relations of future stock returns with market-to-book, investment and disinvestment, seasoned equity offerings, tender offers and stock repurchases, dividend omissions and initiations, expected profitability, profitability, and to certain extent, earnings announcement. And I also propose a new empirical framework for asset pricing tests and expected return estimation, a framework that avoids the difficult task of estimating covariances and long-term growth rates.

My analytical results are derived under a crucial assumption that the operating-profit and the adjustment-cost functions are homogeneous of the same degree, $\alpha \leq 1$. It is natural to ask how restrictive this assumption is. While this is ultimately an empirical question, the special case of $\alpha = 1$ is standard in the investment literature (e.g., Erickson and Whited (2000) and Hubbard (1998)). And several numerically solved models (e.g., Zhang (2004)), where the operating-profit and the adjustment-cost functions have different degrees of homogeneity, yield similar qualitative results as the analytical results obtained here.

Much more work remains to be done in production-based asset pricing. Theoretically, the Q-framework can be extended to model the decisions to go public, merges and acquisitions, spinoffs, debt-financing, and other corporate decisions that are ignored here for simplicity. We can then evaluate whether the Q-framework is consistent with evidence of long-term stock price drift associated with these corporate events. Further, numerically solved models are valuable for producing richer economic insights (for example, cyclical variation of equity financing) than those available analytically.

Empirically and perhaps more importantly, the Q-representation of expected returns can be estimated to see to what extent the economic mechanisms identified in this paper can explain the anomalies quantitatively. The representation can also be implemented to estimate, say, industry cost of capital. More direct tests are also possible. We can measure earnings momentum after controlling for the pricing effects of expected profitability and profitability. Further, the theory implies that the investment, value, payout, and SEO anomalies are all driven by the downward-sloping and convex investment-demand function. It is hence interesting to measure how much payout and SEO anomalies subsist after controlling for the pricing effects of capital investment.

References

Abel, Andrew B., and Olivier J. Blanchard, 1983, An intertemporal model of saving and investment, *Econometrica* 51 (3), 675–692.

Abel, Andrew B., and Janice C. Eberly, 1994, A unified model of investment under uncertainty, *American Economic Review* 84 (5), 1369–1384.

Anderson, Christopher W., and Luis Garcia-Feijóo, 2004, Empirical evidence on capital investment, growth options, and security returns, working paper, University of Kansas.

Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159–178.

Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad, 2004, Consumption, dividends, and the cross-section of equity returns, forthcoming, *Journal of Finance*.

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 George Constantinides, Milton Harris, and Rene Stulz, eds.: *Handbook of the Economics of Finance* (North-Holland, Amsterdam).

Berk, Jonathan B., 1995, A critique of size related anomalies, *Review of Financial Studies* 8, 275–286.

Berk, Jonathan B, Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance*, 54, 1153–1607.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 2004, Valuation and return dynamics of new ventures, *Review of Financial Studies* 17 (1), 1–35.

Bernard, Victor L., 1993, Stock price reactions to earnings announcements: A summary of recent anomalous evidence and possible explanations, in *Advances in Behavioral Finance*, edited by Richard H. Thaler, Russell Sage Foundation, New York.

Bernard, Victor L., and Jacob K. Thomas, 1989, Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research, Supplement* 27, 1–48.

Bernard, Victor L., and Jacob K. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305–340.

Brav, Alon, Christopher Geczy and Paul A. Gompers, 2000, Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics* 56, 209–249.

Brealey, Richard A., and Stewart C. Myers, 2003, Principles of corporate finance, Irwin McGraw-Hill.

Campbell, John Y., 2003, Consumption-based asset pricing, in George Constantinides, Milton Harris, and Rene Stulz, eds.: *Handbook of the Economics of Finance* (North-Holland, Amsterdam).

Campbell, John Y., and Tuomo Vuolteenaho, 2004, Beta beta, good beta, forthcoming, *American Economic Review.*

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004a, Corporate investment and asset price dynamics: Implications for the cross-section of returns, forthcoming, *Journal of Finance.*

Carlson, Murray, Adlai, Fisher, and Ron Giammarino, 2004b, Corporate investment and asset price dynamics: Implications for SEO event studies and long-run performance, working paper, University of British Columbia.

Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *Journal of Finance*, LI (5), 1681–1713.

Claus, James, and Jacob Thomas, 2001, Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets, *Journal of Finance* 56, 1629–1666.

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

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572–621.

Cochrane, John H., 2001, Asset pricing, Princeton University Press, Princeton, New Jersey.

Cohen, Randolph B., Paul A. Gompers, and Tuomo Vuolteenaho, 2002, Who underreacts to cash-flow news? evidence from trading between individuals and institutions, *Journal of Financial Economics* 66, 409–462.

Cooper, Ilan, 2004, Asset pricing implications of non-convex adjustment costs and irreversibility of investment, working paper, Norwegian School of Management.

Cooper, Russell, and Joao Ejarque, 2001, Exhuming Q: Market power vs. capital market imperfections, working paper, NBER.

Cusatis, Patrick J., James A. Miles, and J. Randall Woolridge, 1993, Restructuring through spinoffs, *Journal of Financial Economics* 33, 293–311.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and capital asset pricing, *Journal of Finance*, 53, 1839–1885.

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

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance*, LV (1), 389–406.

Dechow, Patricia M., 1994, Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals, *Journal of Accounting and Economics* 18, 3–42.

Dechow, Patricia M., Amy P. Hutton, and Richard G. Sloan, 1999, An empirical assessment of the residual income valuation model, *Journal of Accounting and Economics* 26 (1), 1–34.

Eckbo, B. Espen, Ronald W. Masulis and Øyvind Norli, 2000, Seasoned public offerings: Resolution of the "new issues puzzle", *Journal of Financial Economics* 56, 251–291.

Erickson, Timothy, and Toni M. Whited, 2000, Measurement error and the relationship between investment and q, *Journal of Political Economy* 108 (5), 1027–1057.

Evans, David S., 1987, The relationship between firm growth, size, and age: Estimates for 100 manufacturing firms, *Journal of Industrial Economics* 35 (4), 567–581.

Fama, Eugene F., 1970, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 383–417.

Fama, Eugene F., 1991, Efficient capital markets: II, *Journal of Finance* XLVI, 1575–1617.

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

Fama, Eugene F., Lawrence Fisher, Michael C. Jensen, and Richard Roll, 1969, The adjustment of stock prices to new information, *International Economic Review* 10 (1), 1–21.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* XLVII (2), 427–465.

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, 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131–155.

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

Fama, Eugene F., and Kenneth R. French, 2004, Profitability, growth, and average returns, working paper, University of Chicago and Dartmouth College.

Frankel, Richard and Charles M.C. Lee, 1998, Accounting valuation, market expectations, and cross-sectional stock returns, *Journal of Accounting and Economics* 25 (3), 283–319.

Gebhardt, William R., Charles M. C. Lee, and Bhaskaram Swaminathan, 2001, Toward an implied cost of capital, *Journal of Accounting Research* 39, 135–176.

Gomes, Joao F., 2001, Financing investment, *American Economic Review* 90, 1263-1285.

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

Gomes, Joao F., Amir Yaron, and Lu Zhang, 2004, Asset pricing implications of firms' financing constraints, working paper, University of Pennsylvania.

Graham, Benjamin, and David L. Dodd, 1934, Security analysis, New York: McGraw Hill.

Hall, Bronwyn H., 1987, The relationship between firm size and firm growth in the US manufacturing sector, *Journal of Industrial Economics* 35 (4), 583–606.

Hall, Robert E., 2001, The stock market and capital accumulation, *American Economic Review* 91, 1185–1202.

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1288.

Haugen, Robert A., and Nardin L. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401–439.

Hayashi, Fumio, 1982, Tobin's marginal and average q: A neoclassical interpretation, *Econometrica* 50 (1), 213–224.

Hennessy, Christopher A., and Toni M. Whited, 2004, Debt dynamics, forthcoming, *Journal of Finance*.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance*, LIV (6), 2143–2184.

Hubbard, Glenn R., 1998, Capital-market imperfections and investment, *Journal of Economic Literature* XXXVI, 193–225.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181–208.

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

Johnson, Timothy C., 2002, Rational momentum effects, *Journal of Finance*, LVII (2), 585–608.

Kogan, Leonid, 2004, Asset prices and real investment, forthcoming, *Journal of Financial Economics*.

Kothari, S.P., 2001, Capital markets research in accounting, *Journal of Accounting and Economics* 31, 105–231.

Lakonishok, Josef, and Theo Vermaelen, 1990, Anomalous price behavior around repurchase tender offers, *Journal of Finance* 45, 455–477.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49 (5), 1541–1578.

Lettau, Martin, and Sydney C. Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy*, 109, 1238–1287.

Lettau, Martin, and Sydney C. Ludvigson, 2004, Expected returns and expected dividend growth, forthcoming, *Journal of Financial Economics*.

Liu, Laura X.L., Jerold B. Warner, and Lu Zhang, 2003, Economic fundamentals, risk, and momentum profits, working paper, University of Rochester.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50 (1), 23–51.

Lucas, Robert E. Jr., 1972, Expectations and the neutrality of money, *Journal of Economic Theory* 4, 103–124.

Lustig, Hanno N., and Stijn Van Nieuwerburgh, 2004, Housing collateral, consumption insurance, and risk premia: an empirical perspective, forthcoming, *Journal of Finance*.

Michaely, Roni, Richard H. Thaler, and Kent L. Womack, 1995, Price reactions to dividend initiations and omissions: overreaction or drift? *Journal of Finance* 50 (2), 573–608.

Miles, James A., and James D. Rosenfeld, 1983, The effect of voluntary spin-off announcements on shareholder wealth, *Journal of Finance* 38 (5), 1597–1606.

Muth, John F., 1961, Rational expectations and the theory of price movements, *Econometrica* 29, 315–335.

Ohlson, James A., 1995, Earnings, book values, and dividends in security valuation, *Contemporary Accounting Research* 18, 109–131.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel, 2004, Housing, consumption, and asset pricing, working paper, University of Chicago.

Piotroski, Joseph D., 2000, Value investing: The use of historical financial statement information to separate winners from losers, *Journal of Accounting Research* 38, 1–51.

Richardson, Scott A., and Richard G. Sloan, 2003, External financing and future stock returns, working paper, The Wharton School, University of Pennsylvania.

Ritter, Jay R., 2003, Investment banking and securities issuance, in George Constantinides, Milton Harris, and Rene Stulz, eds.: *Handbook of the Economics of Finance* (North-Holland, Amsterdam).

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–11.

Restoy, Fernando, and G. Michael Rockinger, 1994, On stock market returns and returns on investment, *Journal of Finance* 49 (2), 543–556.

Rubinstein, Mark, 2001, Rational markets: Yes or no? The affirmative case, *Financial Analyst Journal* May/June, 15–29.

Sapriza, Horacio, and Lu Zhang, 2004, A neoclassical model of financially constrained stock returns, working paper, University of Rochester.

Schultz, Paul, 2003, Pseudo market timing and the long-run underperformance of IPOs, *Journal of Finance* 58 (2), 483–517.

Schwert, G. William, 2003, Anomalies and market efficiency, in George Constantinides, Milton Harris, and Rene Stulz, eds.: *Handbook of the Economics of Finance* (North-Holland, Amsterdam).

Shleifer, Andrei, 2000, *Inefficient Markets: An Introduction to Behavioral Finance,* Clarendon Lectures in Economics (Oxford University Press, Oxford).

Spiess, D. Katherine, and John Afflect-Graves, 1995, Underperformance in long-run stock returns following seasoned equity offerings, *Journal of Financial Economics* 38, 243–267.

Stokey, Nancy L., Robert E. Lucas, Jr., and Edward C. Prescott, 1989, Recursive methods in economic dynamics, Harvard University Press, Cambridge, MA.

Titman, Sheridan, K.C. John Wei, and Feixue Xie, 2003, Capital investments and stock returns, forthcoming, *Journal of Financial and Quantitative Analysis.*

Tobin, James, 1969, A general equilibrium approach to monetary theory, *Journal of Money, Credit, and Banking* 1 (1), 15–29.

Whited, Toni M., and Guojun Wu, 2004, Financial constraints risk, working paper, University of Wisconsin-Madison and University of Michigan.

Vuolteenaho, Tuomo, 2002, What drives firm-level stock returns? *Journal of Finance* LVII (1), 233–264.

Xing, Yuhang, 2004, Firm investments and expected equity returns, working paper, Rice University.

Zhang, Lu, 2004, The value premium, forthcoming, *Journal of Finance.*

A The Anomalies

This appendix briefly reviews the empirical literature on the anomalies that are the targets of my theoretical explanations in this paper.

1. The Investment Anomaly Disinvesting firms earn higher average returns (e.g., Miles and Rosenfeld (1983), Cusatis, Miles, and Woolridge (1993)), but investing firm earn lower average returns in the future (e.g., Richardson and Sloan (2003), Titman, Wei, and Xie (2003), Anderson and Garcia-Feijóo (2004), and Xing (2004)). Titman et al. also shows that the anomaly is stronger for firms with higher operating income-to-asset ratios.

And the literature has proposed different interpretations on this evidence. Cusatis, Miles, and Woolridge (1993) attribute their evidence to market underreaction. Richardson and Sloan (2003) and Titman, Wei, and Xie (2003) interpret their evidence as investors underreacting to empire-building implications of capital investment. Anderson and Garcia-Feijóo (2003) interpret the evidence as consistent with the real options models of Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003), while Xing (2004) interprets the evidence using an optimal investment model similar to that in Zhang (2004).

2. The Value Anomaly Value stocks with lower market-to-book equity earn higher returns on average than growth stocks with higher market-to-book equity (e.g., Graham and Dodd (1934), Rosenberg, Reid, and Lanstein (1985), Fama and French (1992), and Lakonishok, Shleifer, and Vishny (1994)). Fama and French (1993) also document that the value anomaly is stronger in small firms than in big firms where firm size is measured as the ex-dividend market value.

3. The Payout Anomaly Anomalous long-term positive abnormal returns apply to firms paying cash-flow out to shareholders, and are often interpreted as underreaction. Lakonishok and Vermaelen (1990) find positive long-term abnormal returns when firms tender for their stocks. Ikenberry, Lakonishok, and Vermaelen (1995) find that the average abnormal four-year return after the announcements of open market share repurchases is significantly positive. And the average abnormal return is much higher for value firms, but is negative although insignificant for growth firms. Finally, Michaely, Thaler, and Womack (1995) find that stock prices underreact to the negative information in dividend omissions and the positive information in initiations.

4. The SEO-Underperformance Anomaly Anomalous long-term negative abnormal returns apply to firms raising capital from external markets, and are often interpreted as overreaction. Loughran and Ritter (1995) and Spiess and Affleck-Graves (1995) document that firms conducting seasoned equity offerings (SEO) earn much lower returns over the next three to five years than nonissuing firms with similar characteristics. Brav, Geczy, and Gompers (2000) and Eckbo, Masulis, and Norli (2000) find that the underperformance is more pronounced for small firms. A frequent conclusion in this literature is that firms time their external financing decisions to exploit the mispricing of their securities in capital markets because of investor overreaction (e.g., Ritter (2003)).

5. The Expected-Profitability Anomaly Stock prices "underreact" to new information about future cash flow; i.e., shocks to expected cash flows are positively correlated with shocks to expected returns (e.g., Frankel and Lee (1998), Dechow, Hutton, and Sloan (1999), Cohen, Gompers, and Vuolteenaho (2002), Vuolteenaho (2002), and Fama and French (2004)). Lettau and Ludvigson (2004) find similar evidence at the aggregate level. Cohen et al. and Vuolteenaho also find that the magnitude of this correlation is inversely related with the market value.

6. The Profitability Anomaly The profitability anomaly says that, given market price relative to cash flows or book equity, more profitable firms earn higher average returns (e.g., Haugen and Baker (1996) and Piotroski (2000)). Piotroski also shows that this relation is stronger in small firms.

7. The Post-Earnings-Announcement Drift Ball and Brown (1968) and Bernard and Thomas (1989, 1990) document that stock price drifts in the direction of earnings surprise, defined as the scaled change in earnings. Bernard (1993) shows that the magnitude of the drift is inversely related to the market value. And Chan, Jegadeesh, and Lakonishok (1996) find similar evidence using time series and cross-sectional regressions. This anomaly is often interpreted as underreaction to earnings news.

B Technical Proofs

Proof of Lemma 1 By the Principle of Optimality (e.g., Theorem 9.2 of Stokey, Lucas, and Prescott (1989)), the firm's value function (6) can be rewritten recursively as:

$$V(K_t, X_t) = \max_{K_{t+1}} \Pi(K_t, X_t) - \Psi(K_t, K_{t+1}) + \mathrm{E}_t[M_{t+1} V(K_{t+1}, X_{t+1})] \tag{B1}$$

where

$$\Psi(K_t, K_{t+1}) \equiv \Phi\left(K_{t+1} - (1-\delta)K_t, K_t\right) \tag{B2}$$

The envelope condition is:

$$V_1(K_t, X_t) = \Pi_1(K_t, X_t) - \Psi_1(K_t, K_{t+1}) \tag{B3}$$

and the first-order condition is:

$$-\Psi_2(K_t, K_{t+1}) + \mathrm{E}_t[M_{t+1} V_1(K_{t+1}, X_{t+1})] = 0 \tag{B4}$$

Next, from equation (B2),

$$\begin{aligned}\Psi_1(K_t, K_{t+1})K_t + \Psi_2(K_t, K_{t+1})K_{t+1} &= G'\left(\frac{K_{t+1}}{K_t} - (1-\delta)\right) K_{t+1} K_t^{\alpha-1} + \alpha G\left(\frac{K_{t+1}}{K_t} - (1-\delta)\right) K_t^{\alpha} \\ &+ G'\left(\frac{K_{t+1}}{K_t} - (1-\delta)\right) K_t^{\alpha-1} K_{t+1} = \alpha \Psi(K_t, K_{t+1})\end{aligned} \tag{B5}$$

Now plugging equation (B3) into equation (B4) yields the stochastic Euler equation:

$$-\Psi_2(K_t, K_{t+1}) + \mathrm{E}_t[M_{t+1}\left(\Pi_1(K_{t+1}, X_{t+1}) - \Psi_1(K_{t+1}, K_{t+2})\right)] = 0 \qquad \text{(B6)}$$

And expanding the value function (B1) recursively and using equations (2) and (B5) to obtain:

$$\begin{aligned}
\alpha V(K_t, X_t) &= \Pi_1(K_t, X_t)K_t - \Psi_1(K_t, K_{t+1})K_t - \Psi_2(K_t, K_{t+1})K_{t+1} \\
&\quad + \mathrm{E}_t\left[M_{t+1}\left[\Pi_1(K_{t+1}, X_{t+1})K_{t+1} - \Psi_1(K_{t+1}, K_{t+2})K_{t+1} - \Psi_2(K_{t+1}, K_{t+2})K_{t+2}\right.\right. \\
&\quad \left.\left. + \mathrm{E}_{t+1}[M_{t+1,t+2}V(K_{t+2}, X_{t+2})]\right]\right] \\
&= \cdots = \Pi_1(K_t, X_t)K_t - \Psi_1(K_t, K_{t+1})K_t \\
&= V_1(K_t, X_t)K_t \qquad \text{(B7)}
\end{aligned}$$

where the third equality follows by recursive substitution and the stochastic Euler equation (B6). The last equality follows from the envelope condition (B3). ■

Proof of Lemma 2 Solving equation (9) forward yields

$$\begin{aligned}
q_t &= \mathrm{E}_t\left[M_{t+1}(\Pi_1(K_{t+1}, X_{t+1}) - \Phi_2(I_{t+1}, K_{t+1}))\right] + \mathrm{E}_t\left[M_{t+1}(1-\delta)q_{t+1}\right] \\
&= \mathrm{E}_t\left[M_{t+1}(\Pi_1(K_{t+1}, X_{t+1}) - \Phi_2(I_{t+1}, K_{t+1}))\right] + \\
&\quad \mathrm{E}_t\left[M_{t+1}(1-\delta)\mathrm{E}_{t+1}\left[M_{t+2}(\Pi_1(K_{t+2}, X_{t+2}) - \Phi_2(I_{t+2}, K_{t+2}) + (1-\delta)q_{t+2})\right]\right] \\
&= \mathrm{E}_t\left[M_{t+1}(\Pi_1(K_{t+1}, X_{t+1}) - \Phi_2(I_{t+1}, K_{t+1})) + M_{t+2}(1-\delta)(\Pi_1(K_{t+2}, X_{t+2}) - \Phi_2(I_{t+2}, K_{t+2}))\right] \\
&\quad + \mathrm{E}_t\left[M_{t+2}(1-\delta)^2 q_{t+2}\right] \\
&= \cdots = \mathrm{E}_t\left[\sum_{j=1}^{\infty} M_{t+j}(1-\delta)^{j-1}(\Pi_1(K_{t+j}, X_{t+j}) - \Phi_2(I_{t+j}, K_{t+j}))\right]
\end{aligned}$$

where the last two equalities follow from recursive substitution. ■

Proof of Proposition 1 From Proposition 1, $V_1(K_t, X_t)K_t = \alpha V(K_t, X_t)$. Both sides can be rewritten as:

$$\Pi_1(K_t, X_t)K_t - \Phi_2(I_t, K_t)K_t + q_t(1-\delta)K_t = \alpha P_t + \alpha\Pi(K_t, X_t) - \alpha\Phi(I_t, K_t)$$

Simplifying using homogeneity of $\Pi(K_t, X_t)$ and $\Phi(I_t, K_t)$ yields

$$q_t(1-\delta)K_t = \alpha P_t - \Phi_1(I_t, K_t)I_t$$

Equation (12) now follows because $q_t = \Phi_1(I_t, K_t)$ from equation (8). ■

Proof of Lemma 3 The first inequality is shown in the text. Now from Lemma 1 and equations (8) and (13),

$$\Phi_{12}(I_t, K_t) = \frac{\partial q_t}{\partial K_t} = \mathrm{E}_t\left[M_{t+1}\frac{\partial V_1(K_{t+1}, X_{t+1})}{K_t}\right] = (1-\delta)\mathrm{E}_t[M_{t+1}V_{11}(K_{t+1}, X_{t+1})] \quad \text{(B8)}$$

But differentiating both sides of $\alpha V(K_t, X_t) = V_1(K_t, X_t)K_t$ yields

$$V_{11}(K_t, X_t) = \frac{(\alpha-1)}{K_t}V_1(K_t, X_t) = \frac{(\alpha-1)}{K_t}\alpha\frac{V(K_t, X_t)}{K_t} = \frac{(\alpha-1)}{K_t}\alpha\widehat{Q}_t \leq 0 \quad \text{(B9)}$$

This says that the value function is weakly concave in capital. Now plugging equation (B9) into (B8) yields:

$$\Phi_{12}(I_t, K_t) = (1-\delta)(\alpha-1)\frac{1}{K_{t+1}}\mathrm{E}_t[M_{t+1}\alpha\widehat{Q}_{t+1}] = (1-\delta)(\alpha-1)\frac{q_t}{K_{t+1}} \leq 0$$

And differentiating both sides with respect to K_t yields:

$$\Phi_{122}(I_t, K_t) = (\alpha-1)(1-\delta)\frac{1}{K_{t+1}}\frac{\partial q_t}{\partial K_t} + (\alpha-1)(1-\delta)^2 q_t\left(-\frac{1}{K_{t+1}^2}\right) \geq 0$$

■

Proof of Proposition 2 First express stock return in equation (16) in terms of cum-dividend firm value as

$$r_{t+1}^S = \frac{V(K_{t+1}, X_{t+1})}{V(K_t, X_t) - \Pi(K_t, X_t) + \Psi(K_t, K_{t+1})}$$

The recursive value function (B1) evaluated at the optimum then yields $\mathrm{E}_t[M_{t+1}r_{t+1}^S] = 1$.

Combining equations (B3) and (B4) yields an alternative investment return, r_{t+1}^I:

$$r_{t+1}^I = \frac{V_1(K_{t+1}, X_{t+1})}{\Psi_2(K_t, K_{t+1})} = \frac{\Pi_1(K_{t+1}, X_{t+1}) - \Psi_1(K_{t+1}, K_{t+2})}{\Psi_2(K_t, K_{t+1})} \quad \text{(B10)}$$

which is equal to equation (15) since equation (B2) implies $\Psi_2 = \Phi_1$ and $\Psi_1 = \Phi_2 - \Phi_1(1-\delta)$.

Now,

$$
\begin{aligned}
r^I_{t+1} &= \frac{V_1(K_{t+1}, X_{t+1})}{\Psi_2(K_t, K_{t+1})} = \frac{V_1(K_{t+1}, X_{t+1})K_{t+1}}{\alpha\Psi(K_t, K_{t+1}) - \Psi_1(K_t, K_{t+1})K_t} \\
&= \frac{V_1(K_{t+1}, X_{t+1})K_{t+1}}{V_1(K_t, X_t)K_t - \Pi_1(K_t, X_t)K_t + \alpha\Psi(K_t, K_{t+1})} \\
&= \frac{\alpha V(K_{t+1}, X_{t+1})}{\alpha V(K_t, X_t) - \alpha\Pi(K_t, X_t) + \alpha\Psi(K_t, K_{t+1})} = r^S_{t+1}
\end{aligned}
$$

where the first equality follows from equation (B10), the second follows from equation (B5), the third equality follows from the envelope condition (B3), and the fourth equality follows from Lemma 1 and equation (2). ■

Proof of Proposition 3 From equations (4) and (15),

$$
\begin{aligned}
r_{t+1} &= \frac{U_{t+1}}{\Phi_1(I_t, K_t)} = \frac{U_{t+1}}{G'(I_t/K_t)K_t^{\alpha-1}} \qquad \text{(B11)} \\
\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)} &= -\frac{\mathrm{E}_t[U_{t+1}]}{K_t^{\alpha-1}[G'(I_t/K_t)]^2} G''\left(\frac{I_t}{K_t}\right) + \frac{1}{\Phi_1(I_t, K_t)} \frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial(I_t/K_t)} \qquad \text{(B12)}
\end{aligned}
$$

where the first term in equation (B12) is less than zero because Assumption 3 implies $G''(\cdot) > 0$. It then suffices to show that $\frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial(I_t/K_t)} < 0$. But plugging equation (3) into (24) yields:

$$
\begin{aligned}
\mathrm{E}_t[U_{t+1}] = \mathrm{E}_t\Bigg[&\Pi_1\left(\left[\frac{I_t}{K_t} + (1-\delta)\right]K_t, X_{t+1}\right) \\
&- \Phi_2\left(K_{t+2} - (1-\delta)\left[\frac{I_t}{K_t} + (1-\delta)\right]K_t, \left[\frac{I_t}{K_t} + (1-\delta)\right]K_t\right) \\
&+ (1-\delta)\Phi_1\left(K_{t+2} - (1-\delta)\left[\frac{I_t}{K_t} + (1-\delta)\right]K_t, \left[\frac{I_t}{K_t} + (1-\delta)\right]K_t\right)\Bigg] \qquad \text{(B13)}
\end{aligned}
$$

Differentiating both sides with respect to (I_t/K_t) yields:

$$
\begin{aligned}
\frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial(I_t/K_t)} = \mathrm{E}_t[&\Pi_{11}(K_{t+1}, X_{t+1})K_t + 2(1-\delta)\Phi_{12}(I_{t+1}, K_{t+1})K_t - \Phi_{22}(I_{t+1}, K_{t+1})K_t \\
&- (1-\delta)^2\Phi_{11}(I_{t+1}, K_{t+1})K_t] < 0 \qquad \text{(B14)}
\end{aligned}
$$

where the inequality follows from Assumptions 1 and 2 and Lemma 3. Finally, plugging equations (2), (24), (B14), and (21) into (B12) and using $\Pi_{11}(K_{t+1}, X_{t+1}) = \alpha(\alpha-1)\frac{\Pi_{t+1}}{K_{t+1}}\frac{1}{K_{t+1}}$:

$$\partial \left| \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} \right| / \partial \left(\frac{\Pi_t}{K_t} \right) = \frac{\alpha \rho_\pi G''(I_t/K_t)}{\Phi_1(I_t, K_t) G'(I_t/K_t)} + \alpha(1-\alpha)\rho_\pi \frac{K_t}{K_{t+1}} > 0 \qquad \text{(B15)}$$

■

Proof of Proposition 4 From the investment-return equation (15),

$$\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} = -\frac{\mathrm{E}_t[U_{t+1}]}{\alpha Q_t^2} + \frac{1}{\alpha Q_t} \frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial Q_t} \qquad \text{(B16)}$$

By Lemma 4, to show $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} < 0$, it suffices to show that $\frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial Q_t} < 0$. But $I_t/K_t = G'^{-1}(q_t K_t^{1-\alpha})$, from equation (8). Writing q_t further as αQ_t, plugging I_t/K_t into equation (B13), and using the Inverse Function Theorem yield $\partial(I_t/K_t)/\partial Q_t = \alpha K_t^{1-\alpha}/G''(G'^{-1}(\alpha Q_t K_t^{1-\alpha}))$, where $G'^{-1}(\cdot)$ is the inverse function of G'. Now by the chain rule and equation (B14),

$$\begin{aligned} \frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial Q_t} &= \frac{\alpha K^{1-\alpha} K_t}{G''\left(G'^{-1}(\alpha Q_t K_t^{1-\alpha})\right)} \mathrm{E}_t[\Pi_{11}(K_{t+1}, X_{t+1}) + 2(1-\delta)\Phi_{12}(I_{t+1}, K_{t+1}) \\ &\quad - \Phi_{22}(I_{t+1}, K_{t+1}) - (1-\delta)^2 \Phi_{11}(I_{t+1}, K_{t+1})] < 0 \qquad \text{(B17)} \end{aligned}$$

where the inequality follows because $\Pi_{11} \leq 0, \Phi_{12} \leq 0$, $\Phi_{22} \geq 0$, and $\Phi_{11} > 0$.

To establish the second inequality in the proposition, it suffices to show:

$$\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} > 0 \qquad \text{(B18)}$$

The reason is that, by the chain rule of partial derivatives,

$$\partial \left| \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} \right| / \partial P_t = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t \partial P_t} = -\partial \left(\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} \frac{\partial Q_t}{\partial P_t} \right) / \partial Q_t = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} \frac{\partial Q_t}{\partial P_t} = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} \frac{1}{K_{t+1}}$$

Now, from equation (B16),

$$\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} = \frac{2\mathrm{E}_t[U_{t+1}]}{\alpha Q_t^3} - \frac{2}{\alpha Q_t^2} \frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial Q_t} + \frac{1}{\alpha Q_t} \frac{\partial^2 \mathrm{E}_t[U_{t+1}]}{\partial Q_t^2}$$

To show equation (B18), it suffices to show $\frac{\partial^2 \mathrm{E}_t[U_{t+1}]}{\partial Q_t^2} \geq 0$ because $\frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial Q_t} < 0$.

For notational convenience, denote the term in the conditional expectation in equation (B17) as W_{t+1} that is negative. Now substituting $K_{t+1}=\left[G'^{-1}(\alpha Q_t K_t^{1-\alpha})+(1-\delta)\right]K_t$ and $I_{t+1}=K_{t+2}-(1-\delta)\left[G'^{-1}(\alpha Q_t K_t^{1-\alpha})+(1-\delta)\right]K_t$ into equation (B17) and differentiating the equation with respect to Q_t yield:

$$\frac{\partial^2 \mathrm{E}_t[U_{t+1}]}{\partial Q_t^2} = -\alpha^2 K_t^{2(1-\alpha)} K_t \frac{G'''(I_t/K_t)}{[G''(I_t/K_t)]^3}\mathrm{E}_t[W_{t+1}] + \frac{\alpha K^{1-\alpha}K_t}{G''(\alpha Q_t K_t^{1-\alpha})}\frac{\partial \mathrm{E}_t[W_{t+1}]}{\partial Q_t} \tag{B19}$$

To show $\frac{\partial^2 \mathrm{E}_t[U_{t+1}]}{\partial Q_t^2} \geq 0$, it suffices to show that $\frac{\partial \mathrm{E}_t[W_{t+1}]}{\partial Q_t} \geq 0$ because the first term in equation (B19) is nonnegative (Lemma 3 and Assumption 5 imply that $W_{t+1}<0$ and $G'''(\cdot)\geq 0$ because $\Phi_{111}\geq 0$). But,

$$\begin{aligned}\frac{\partial \mathrm{E}_t[W_{t+1}]}{\partial Q_t} &= \frac{\alpha K_t^{1-\alpha}K_t}{G''(I_t/K_t)}\mathrm{E}_t[\Pi_{111}(K_{t+1},X_{t+1}) - 3(1-\delta)^2\Phi_{112}(I_{t+1},K_{t+1}) \\ &\quad + 3(1-\delta)\Phi_{122}(I_{t+1},K_{t+1}) - \Phi_{222}(I_{t+1},K_{t+1}) + (1-\delta)^3\Phi_{111}(I_{t+1},K_{t+1})] \geq 0\end{aligned}$$

where the inequality follows from Assumption 5 and Lemma 3. ■

Proof of Proposition 5 $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} > 0$ follows from equation (B18). Equation (8) and Proposition 1 imply that $Q_t = G'(I_t/K_t)K_t^{\alpha-1}/\alpha$, which in turn implies that

$$\frac{\partial Q_t}{\partial(I_t/K_t)} = \frac{1}{\alpha}G''\left(\frac{I_t}{K_t}\right)K_t^{\alpha-1} > 0 \tag{B20}$$

$$\frac{\partial^2 Q_t}{\partial(I_t/K_t)\partial P_t} = \frac{\partial(1/K_{t+1})}{\partial(I_t/K_t)} = -\frac{K_t}{K_{t+1}^2} < 0 \tag{B21}$$

Now,

$$\begin{aligned}\partial\left|\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\right|/\partial P_t &= -\partial\left(\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\right)/\partial P_t = -\partial\left(\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial Q_t}{\partial(I_t/K_t)}\right)/\partial P_t \\ &= -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t \partial P_t}\frac{\partial Q_t}{\partial(I_t/K_t)} - \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial^2 Q_t}{\partial(I_t/K_t)\partial P_t}\end{aligned}$$

Because of $\partial \mathrm{E}_t[r_{t+1}]/\partial Q_t < 0$ and equations (B20) and (B21), to show the second inequality in the proposition, it suffices to show $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t \partial P_t} > 0$. But,

$$\begin{aligned}\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t \partial P_t} &= \partial\left(\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial Q_t}{\partial P_t}\right)/\partial Q_t = \frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2}\frac{1}{K_{t+1}} + \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial^2 Q_t}{\partial Q_t \partial P_t} \\ &= \frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2}\frac{1}{K_{t+1}} > 0\end{aligned} \tag{B22}$$

And,

$$\partial \left| \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} \right| / \partial Q_t = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t) \partial Q_t} = -\partial \left(\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} \frac{\partial Q_t}{\partial (I_t/K_t)} \right) / \partial Q_t$$
$$-\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} \frac{\partial Q_t}{\partial (I_t/K_t)} + \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} \frac{\partial^2 Q_t}{\partial (I_t/K_t) \partial Q_t} = -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2} \frac{\partial Q_t}{\partial (I_t/K_t)} < 0$$

where the last equality follows because $\frac{\partial^2 Q_t}{\partial (I_t/K_t) \partial Q_t} = \partial \left(\frac{\partial Q_t}{\partial Q_t} \right) / \partial (I_t/K_t) = 0$.

Finally, taking partial derivative of equation (B12) with respect to $\frac{I_t}{K_t}$ yields:

$$\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)^2} = \frac{2[G''(I_t/K_t)]^2 - G'''(I_t/K_t) G'(I_t/K_t)}{K_t^{\alpha-1} [G'(I_t/K_t)]^2} \mathrm{E}_t[U_{t+1}]$$
$$- \frac{2G''(I_t/K_t)}{K_t^{\alpha-1} [G'(I_t/K_t)]^2} \frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial (I_t/K_t)} + \frac{1}{G'(I_t/K_t) K_t^{\alpha-1}} \frac{\partial^2 \mathrm{E}_t[U_{t+1}]}{\partial (I_t/K_t)^2}$$

The second term is positive because $\frac{\partial \mathrm{E}_t[U_{t+1}]}{\partial (I_t/K_t)} < 0$ from equation (B14), and the third term is nonnegative because $\frac{\partial^2 \mathrm{E}_t[U_{t+1}]}{\partial (I_t/K_t)^2} = K_t \frac{\partial \mathrm{E}_t[W_{t+1}]}{\partial (I_t/K_t)} = K_t \frac{\partial \mathrm{E}_t[W_{t+1}]}{\partial Q_t} \frac{\partial Q_t}{\partial (I_t/K_t)} \geq 0$. And the first term is nonnegative when $2 \left[G'' \left(\frac{I_t}{K_t} \right) \right]^2 \geq G''' \left(\frac{I_t}{K_t} \right) G' \left(\frac{I_t}{K_t} \right)$. ■

Proof of Proposition 6 First, when $\Pi_t - \Phi_t \leq 0$ or $C_t = 0$, the two derivatives in the proposition are exactly zero. Now consider the case when $C_t > 0$; so I can ignore the indicator function. Equation (25) implies that

$$\frac{C_t}{K_t} = \frac{\Pi_t}{K_t} - G\left(\frac{I_t}{K_t}\right) K_t^{\alpha-1} \quad \text{or} \quad \frac{I_t}{K_t} = G^{-1}\left[\left(\frac{\Pi_t}{K_t} - \frac{C_t}{K_t} \right) K_t^{1-\alpha} \right] \tag{B23}$$

where $G^{-1}(\cdot)$ is the inverse function of G, and is also an increasing function because G is around the neighborhood of optimal investment rate.

Now by the chain rule,

$$\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (C_t/K_t)} = \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} \frac{\partial (I_t/K_t)}{\partial (C_t/K_t)} = -\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} \frac{K^{1-\alpha}}{G'(I_t/K_t)} > 0$$

where the inequality follows because Proposition 3 says that $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial (I_t/K_t)} < 0$.

Next, again by the chain rule,

$$\begin{aligned}
\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(C_t/K_t)\partial Q_t} &= -\partial\left(\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial Q_t}{\partial(C_t/K_t)}\right)/\partial Q_t \\
&= -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2}\frac{\partial Q_t}{\partial(C_t/K_t)} + \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial^2 Q_t}{\partial(C_t/K_t)\partial Q_t} \\
&= \frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t^2}\frac{\partial Q_t}{\partial(I_t/K_t)}\frac{\partial(I_t/K_t)}{\partial(C_t/K_t)} < 0
\end{aligned}$$

where the inequality follows from the inequality (B18) and equation (B20). ■

Proof of Proposition 7 When $O_t = 0$, the two derivatives in the proposition are exactly zero. Consider the case when $O_t > 0$. Now equation (26) implies that

$$\frac{O_t}{K_t} = G\left(\frac{I_t}{K_t}\right)K_t^{\alpha-1} - \frac{\Pi_t}{K_t} \quad \text{or} \quad \frac{I_t}{K_t} = G^{-1}\left[\left(\frac{O_t}{K_t} + \frac{\Pi_t}{K_t}\right)K_t^{1-\alpha}\right] \tag{B24}$$

Now by the chain rule and Proposition 3,

$$\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(O_t/K_t)} = \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\frac{\partial(I_t/K_t)}{\partial(O_t/K_t)} = \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(I_t/K_t)}\frac{K_t^{1-\alpha}}{G'(I_t/K_t)} < 0$$

And again by the chain rule,

$$\begin{aligned}
\partial\left|\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial(O_t/K_t)}\right|/\partial P_t &= -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial(O_t/K_t)\partial P_t} = -\partial\left(\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial Q_t}{\partial(O_t/K_t)}\right)/\partial P_t \\
&= -\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t \partial P_t}\frac{\partial Q_t}{\partial(O_t/K_t)} - \frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t}\frac{\partial^2 Q_t}{\partial(O_t/K_t)\partial P_t}
\end{aligned}$$

To show the left-hand-side is negative, it suffices to show that $\frac{\partial^2 Q_t}{\partial(O_t/K_t)\partial P_t} < 0$. The reason is that $\frac{\partial^2 \mathrm{E}_t[r_{t+1}]}{\partial Q_t \partial P_t} > 0$ from equation (B22), $\frac{\partial Q_t}{\partial(O_t/K_t)} = \frac{\partial Q_t}{\partial(I_t/K_t)}\frac{\partial(I_t/K_t)}{\partial(O_t/K_t)} > 0$, and $\frac{\partial \mathrm{E}_t[r_{t+1}]}{\partial Q_t} < 0$. But,

$$\frac{\partial^2 Q_t}{\partial(O_t/K_t)\partial P_t} = \frac{\partial(1/K_{t+1})}{\partial(I_t/K_t)}\frac{\partial(I_t/K_t)}{\partial(O_t/K_t)} < 0$$

where the inequality follows from equation (B21) and $\frac{\partial(I_t/K_t)}{\partial(O_t/K_t)} > 0$. ■