The Analytics of Investment, $q$, and Cash Flow*

Andrew B. Abel

Wharton School of the University of Pennsylvania
National Bureau of Economic Research

First draft, September 2012
Current draft, July 2014

Abstract

This paper analyzes the relationships among investment, $q$, and cash flow in a tractable stochastic model in which marginal $q$ and average $q$ are identically equal. In the special, but widely used, case of quadratic adjustment costs, it derives an expression for $q$ that is in closed form, up to an additive constant. After analyzing the impact of changes in the distribution of the marginal operating profit of capital, the paper extends the model to include measurement error and then analyzes the cash flow coefficient in regressions of investment on $q$ and cash flow. The coefficient on cash flow is typically estimated to be positive and to be larger for firms, such as rapidly growing firms, that are likely to face financial frictions. These findings are typically interpreted as evidence of financial frictions facing the firm. This paper derives closed-form expressions for the cash flow coefficient in the model presented here and shows that it is positive and is larger for more rapidly growing firms, even though there are no financial frictions in the model.

*I thank Joao Gomes, Richard Kihlstrom, and Stavros Panageas for helpful discussion, Colin Ward for excellent research assistance, and participants in seminars at University of California at Santa Barbara, University of Southern California, Columbia University, and the Penn Macro Lunch Group for helpful comments.
Empirical investment equations typically find that Tobin’s $q$ has a positive effect on capital investment by firms, and that even after taking account of the effect of Tobin’s $q$ on investment, cash flow has a positive effect on investment. Of course, the interpretations of these results rely on some theoretical model of investment. Typically, the theoretical model that underlies the relationship between Tobin’s $q$ and investment is based on convex capital adjustment costs. In this framework, marginal $q$ is a sufficient statistic for investment. No other variables, in particular, cash flow, should have any explanatory power for investment, once account is taken of marginal $q$. The fact that cash flow has a positive impact on investment, even after taking account of $q$, is interpreted by many researchers as evidence of financing constraints facing firms. That interpretation is bolstered by the finding that the cash flow coefficient is larger for firms that are likely to be financially constrained, such as rapidly growing firms.

In this paper, I develop and analyze a tractable stochastic model of investment, $q$, and cash flow and use it to interpret the empirical results described above. In modeling adjustment costs, the first choice is whether to specify these costs as a function of investment only (usually either gross investment or net investment) or to specify these costs as a function of the capital stock as well as of investment. The former specification is more tractable and easier to analyze in some ways, especially in the context of perfect competition and constant returns to scale in production. In that context, the marginal contribution of capital to operating profits is a function only of exogenous factors such as the price of output, the wage rate, and the level of productivity. Marginal $q$ equals the expected present value of the stream of marginal contributions to operating profit accruing to the undepreciated portion of a unit of capital installed today. When this stream of marginal operating profits depends only on exogenous factors, the value of marginal $q$ is exogenous to the firm, and in particular, does not depend on current or future investment decisions of the firm. In Abel (1983), I exploit this exogeneity of the stream of marginal operating profits to derive closed-form expressions for marginal $q$ and for the value of the firm.

One unfortunate implication of specifying adjustment costs to depend only on investment, and not on the capital stock also, is that the optimal level of investment is independent of the size of the firm. Two firms with the same value of marginal $q$ would undertake the same level of investment even if one firm’s capital stock is a thousand times the size of the other firm’s capital stock. Alternatively, as shown by Lucas (1967), if the net profit of the firm, after deducting all costs associated with investment, is linearly homogeneous in capital, labor, and investment, the growth rate of the firm is independent of its size. Later, Hayashi (1982) showed that this linear homogeneity implies that Tobin’s $q$, often called average $q$, is identically equal to marginal $q$. This equality of marginal $q$ and average $q$ is particularly powerful, because average $q$, which is in principle observable, can be used to measure marginal $q$, which is the appropriate shadow value of capital that determines the optimal

---

1 Lucas and Prescott (1971) and Mussa (1977) first demonstrated the link between securities prices, which are related to Tobin’s $q$, and investment in an adjustment cost framework.
rate of investment. In addition, this linearly homogeneous framework relates the investment-capital ratio to \( q \) and most empirical analyses, in fact, use the investment-capital ratio as the dependent variable in regressions.

Although the linearly homogeneous framework has some convenient properties, it can be less tractable because the adjustment cost function depends on the firm’s capital stock as well as on its rate of investment. In general, an additional unit of capital stock reduces the adjustment cost, and this marginal benefit of capital must be added to the marginal operating profit of capital in computing marginal \( q \). Even though the marginal operating profit of capital is exogenous, the marginal reduction in the adjustment cost depends on the firm’s choice of investment. This dependence on the firm’s decisions complicates the calculation of marginal \( q \), and has been a barrier to deriving closed-form expressions for \( q \), even under quadratic adjustment costs. In this paper, I derive a closed-form solution for \( q \) under quadratic adjustment costs in the special case in which the marginal operating capital is constant.

Ideally, to analyze the response of investment to \( q \) requires a framework with variation in the firm’s marginal operating profit, which induces variation in \( q \), and in optimal investment. In this paper, I develop a model in which stochastic variation in the marginal operating profit of capital is generated by a Markov regime-switching process. With this stochastic specification, the model turns out to be very tractable. I present closed-form solutions for optimal investment and \( q \), up to a scalar constant, in the case of quadratic adjustment costs. More importantly, however, the framework is tractable enough to permit straightforward analysis of the effects on \( q \) and investment of changes in the marginal operating profit for a particular firm, even if the adjustment cost function is not quadratic. The model can also be used to compare \( q \) and investment across firms that face different interest rates, different depreciation rates, and different stochastic processes for the exogenous marginal operating profit of capital. I apply this tractable framework to analyze the impact on marginal \( q \) and investment of a mean-preserving spread in the unconditional distribution of marginal operating profit of capital, as well as the impact of a change in the persistence of the Markov regime-switching process generating these marginal operating profits.

As mentioned earlier, a common feature of adjustment cost models of investment is that marginal \( q \) is a sufficient statistic for investment. In particular, cash flow should not add any explanatory power for investment after taking account of marginal \( q \). This feature holds in the model I present here and might appear to be an obstacle to accounting for the empirical cash flow effect on investment described above. To overcome that obstacle, I introduce classical measurement error in Section 5. It has been argued in the literature that if \( q \) is measured with error, then since the true value of \( q \) is an increasing function of cash flow, cash flow will have some additional explanatory power for investment, and the coefficient on cash flow will be positive in a regression of investment on \( q \) and cash flow. I derive a simple expression to illustrate the impact of measurement error on
estimated coefficients on \( q \) and cash flow. The model I present here allows the analysis to go beyond the existing argument for a positive cash flow coefficient by showing that the size of the cash flow coefficient can be larger for firms that grow more rapidly. Since rapidly growing firms are more likely to be classified as facing binding financial constraints, the model’s implication that rapidly growing firms can have larger cash flow coefficients is consistent with empirical studies that find larger cash flow coefficients for firms classified as financially constrained. However, because the model has perfect capital markets, without any financial frictions, the results described here imply that the finding of positive cash flow coefficients that are larger for faster-growing firms cannot be taken as evidence of financing constraints.

Because the analysis of the model relies on the equality of marginal \( q \) and average \( q \), I begin, in Section 1, by re-stating, and extending to a stochastic framework, the Hayashi condition under which average \( q \) and marginal \( q \) are equal. Section 2 introduces the model of the firm and analyzes the valuation of a unit of capital and the optimal investment decision in the case in which the marginal operating profit of capital is known to be constant forever. More than simply serving as a warm up to the stochastic model, Section 2 introduces a function that facilitates the analysis of the stochastic model that follows in later sections. I introduce a Markov regime-switching process for the marginal operating profit of capital in Section 3 to generate stochastic variation in \( q \) and optimal investment. In Section 4, I analyze the impact of changes in the stochastic properties of the marginal operating profit of capital, specifically, changes to the unconditional distribution and changes to the persistence of this exogenous random variable. In order to account for the positive impact of cash flow on investment, even after taking account of \( q \), I introduce classical measurement error in Section 5. In addition, I show that this cash flow coefficient tends to be larger for firms that are growing more rapidly. Concluding remarks are in Section 6. The proofs of lemmas, propositions, and corollaries are in the Appendix.

1 The Hayashi Condition

Before describing the specific framework that I analyze in this paper, it is useful to begin with a simple, yet more general, description of the conditions under which average \( q \) and marginal \( q \) are equal. Consider a competitive firm with capital stock \( K_t \) at time \( t \), where time is continuous. The firm accumulates capital by undertaking gross investment \( I_t \) at time \( t \), and capital depreciates at rate \( \delta_t \), so the capital stock evolves according to

\[
\frac{dK_t}{dt} = I_t - \delta_t K_t. \tag{1}
\]

The firm uses capital, \( K_t \), and labor, \( L_t \), to produce and sell output at time \( t \). I assume that the price of capital goods is constant and normalize it to be one. Define \( \pi_t(K_t, I_t) = \)
Proposition 1 (extension of Hayashi) If \( \pi_s(K_s, I_s) \) is linearly homogeneous in \( K_s \) and \( I_s \), then for any \( \omega \geq 0 \), \( V_t(\omega K_t) = \omega V_t(K_t) \), i.e., the value function is linearly homogeneous in \( K_t \), so that average \( q \), \( V_t(K_t) \), and marginal \( q \), \( V_t'(K_t) \), are identically equal.

For the remainder of this paper, I will assume that \( \pi_s(K_s, I_s) \) is linearly homogeneous in \( K_s \) and \( I_s \) so that average \( q \) and marginal \( q \) are equal.

2 Model of the Firm

Consider a competitive firm that faces convex costs of adjustment that are separable from the production function. The firm uses capital, \( K_t \), and labor, \( L_t \), to produce non-storable output, \( Y_t \), at time \( t \) according to the production function \( Y_t = A_t f(K_t, L_t) \), where \( f(K_t, L_t) \) is linearly homogeneous in \( K_t \) and \( L_t \), and \( A_t \) is the exogenous level of total factor productivity. If the amount of labor is costlessly and instantaneously adjustable, the firm chooses \( L_t \) at time \( t \) to maximize instantaneous revenue less wages \( p_t A_t f(K_t, L_t) - w_t L_t \), where \( p_t \) is the price of the firm’s output at time \( t \) and \( w_t \) is the wage rate per unit of labor at time \( t \). The linear homogeneity of \( f(K_t, L_t) \) and the assumption that the firm is a price-taker in the markets for its output and its labor together imply that the maximized value of revenue less wages is \( \Phi_t K_t \), where \( \Phi_t \equiv \max_l [p_t A_t f(1, l) - w_t l] \).

The marginal (and average) operating profit of capital, \( \Phi_t \), is a deterministic function of \( A_t \), \( p_t \), and \( w_t \), all of which are exogenous to the firm and possibly stochastic. Therefore, \( \Phi_t \) is exogenous to the firm and, henceforth, I will treat \( \Phi_t \) as the fundamental exogenous variable facing the firm, comprising the effects of productivity, output price and the wage rate.

I assume that the depreciation rate of capital is constant, so that net investment, \( \frac{dK_t}{dt} \), is given by equation (1) with \( \delta \) equal to the constant \( \delta \). Define \( \gamma_t \equiv \frac{K_t}{dt} \) to be the investment-capital ratio at time \( t \). Therefore, the growth rate of the capital stock, \( g_t \), is

\[
g_t \equiv \frac{1}{K_t} \frac{dK_t}{dt} = \gamma_t - \delta, \tag{3}
\]
so that for $s \geq t$

$$K_s = K_t \exp \left( \int_t^s g_u du \right).$$  \hfill (4)

Finally, I will specify the stochastic discount factor, $M(t, s)$, to be simply $\exp(-r(s-t))$, so that net cash flows are discounted at the constant rate $r$.

At time $t$, the firm chooses gross investment, $I_t$. The cost of this investment has two components. The first component is the cost of purchasing capital at a price per unit that I assume to be constant over time and normalize to be one. Thus, this component of the cost of gross investment at rate $I_t$ is simply $I_t = \gamma_t K_t$, which, of course, would be negative if the firm sells capital so that $I_t < 0$. The second component is the cost of adjustment, $C(I_t, K_t)$, which I assume to be linearly homogeneous in $I_t$ and $K_t$. It will be convenient to use the definition of the investment-capital ratio, $\gamma_t \equiv \frac{I_t}{K_t}$, to write the adjustment cost function as $c(\gamma_t) K_t$, where $c(\gamma) \geq 0$ is strictly convex and at least twice differentiable, $c'(r+\delta) > 0$, and for some $\gamma^m < r+\delta$, $c'(\gamma^m) = -1$. Therefore, $\gamma_t + c(\gamma_t)$ is strictly convex and attains its minimum value at $\gamma^m$. Finally, after choosing the optimal usage of labor, the amount of revenue less wages and less the cost of investment is

$$\pi_t(K_t, I_t) \equiv \left[ \Phi_t - \gamma_t - c(\gamma_t) \right] K_t.$$  \hfill (5)

### 2.1 Constant $\Phi_t$

Consider the case in which the marginal operating profit of capital, $\Phi_t$, is constant forever. This case is simple enough that closed-form solutions for the value of the firm and optimal investment are readily available when the adjustment cost function $c(\gamma_t)$ is quadratic. More importantly, however, the analytic apparatus developed in the case of certainty will prove to be useful in later sections when $\Phi_t$ evolves according to a Markov regime-switching process.

I begin by defining $G$, which is an admissible set of values for $\Phi$, the constant marginal operating profit of capital,\footnote{Since $c'(\gamma^m) = -1$ and $c'(r+\delta) > 0$, the strict convexity of $c(\gamma)$ implies that $\gamma^m < r+\delta$ and that $c(\gamma) + \gamma_t$ is strictly increasing in $\gamma_t$ for all $\gamma_t > \gamma^m$. Therefore, $c(r+\delta) + r+\delta > c(\gamma^m) + \gamma^m$ so that $G$ is non-vacuous.}

$$G \equiv \{ \Phi : c(\gamma^m) + \gamma^m < \Phi < c(r+\delta) + r+\delta \}.$$  \hfill (6)

The lower bound on $G$ ensures that there is a value of $\gamma_t$ such that $\pi_t(K_t, I_t) > 0$ when $K_t > 0$, so that the value of the firm is positive.\footnote{If $\Phi - c(\gamma) - \gamma > 0$ and (2) $\gamma - \delta > r$, the value of the firm, with $K_t > 0$, would be positive and infinite because the growth rate of $[\Phi - c(\gamma) - \gamma] K_t$, which is $\gamma - \delta$, would exceed $r$. The upper bound implies that if $\gamma \geq r+\delta$, then $c(\gamma) + \gamma \geq c(r+\delta) + r+\delta$, so $\Phi - c(\gamma) - \gamma < 0$.} The upper bound on $G$ keeps the value of the firm finite when it is positive.\footnote{If $1) \Phi - c(\gamma) - \gamma > 0$ and (2) $\gamma - \delta > r$, the value of the firm, with $K_t > 0$, would be positive and infinite because the growth rate of $[\Phi - c(\gamma) - \gamma] K_t$, which is $\gamma - \delta$, would exceed $r$. The upper bound implies that if $\gamma \geq r+\delta$, then $c(\gamma) + \gamma \geq c(r+\delta) + r+\delta$, so $\Phi - c(\gamma) - \gamma < 0$.}
2.1.1 The Value of a Unit of Capital

With a constant marginal operating profit of capital, \( \Phi \), constant depreciation rate, \( \delta \), and constant discount rate, \( r \), the optimal investment-capital ratio, \( \gamma_t \), is constant also. In this case, the value of the firm in equation (2) can be written as

\[
V_t (K_t) = \max_\gamma \int_0^\infty [\Phi - \gamma - c(\gamma)] K_t e^{-r(s-t)} ds. \tag{7}
\]

Dividing both sides of equation (7) by \( K_t \) and using equation (4) with \( g_n = \gamma - \delta \) yields an expression for the average value of a unit of capital, \( v \equiv \frac{V_t(K_t)}{K_t} \), which is

\[
v = \max_\gamma \frac{\Phi - \gamma - c(\gamma)}{r + \delta - \gamma}. \tag{8}
\]

The value of a unit of capital shown in equation (8) equals \( \pi_t (K_t, I_t) / K_t \) divided by the excess of the interest rate, \( r \), over the growth rate, \( \gamma - \delta \). Differentiating the maximand on the right-hand side of equation (8) with respect to \( \gamma \) and setting the derivative equal to zero yields\(^5\)

\[
1 + c' (\gamma) = \frac{\Phi - \gamma - c(\gamma)}{r + \delta - \gamma}. \tag{9}
\]

Rewriting equation (9) yields

\[
\Phi - (r + \delta) - c(\gamma) - (r + \delta - \gamma) c' (\gamma) = 0. \tag{10}
\]

It will be useful to define a function \( H(\gamma, \Phi, \rho) \) to characterize the optimal investment-capital ratio. Specifically,

\[
H(\gamma, \Phi, \rho) \equiv \Phi - \rho - c(\gamma) - (\rho - \gamma) c' (\gamma), \tag{11}
\]

and the optimal value of \( \gamma \), characterized by equation (10), satisfies

\[
H(\gamma, \Phi, r + \delta) = 0. \tag{12}
\]

**Lemma 1** Define \( H(\gamma, \Phi, \rho) \equiv \Phi - \rho - c(\gamma) - (\rho - \gamma) c' (\gamma) \) and assume that \( \Phi \in G \) and \( \rho \geq r + \delta \). Then

1. \( H(\gamma, \Phi, \rho) \) is an increasing, linear function of \( \Phi \).
2. \( H(\gamma, \Phi, \rho) \) is a decreasing, linear function of \( \rho \) for \( \gamma > \gamma^m \), where \( c' (\gamma^m) = -1 \).
3. \( H(\gamma, \Phi, \rho) \) is strictly quasi-convex in \( \gamma \).
4. \( H(\gamma^m, \Phi, \rho) > 0 \), where \( c' (\gamma^m) = -1 \).

---

\(^5\)Define \( h(\gamma) \equiv \frac{\Phi - \gamma - c(\gamma)}{r + \delta - \gamma} \) and observe that \( h'(\gamma) = \frac{1}{r + \delta - \gamma} [-1 - c'(\gamma) + h(\gamma)] \) and that \( h''(\gamma) = \frac{1}{(r + \delta - \gamma)^2} [-c''(\gamma) + h'(\gamma)] \). Therefore, if \( h'(\gamma) = 0 \), then \( h''(\gamma) = \frac{1}{(r + \delta - \gamma)^2} c''(\gamma) \). If \( h'(\gamma_1) = 0 \) and \( \gamma_1 < r + \delta \), then \( h''(\gamma) < 0 \) so \( h(\gamma_1) \) is a local maximum. However, if \( h'(\gamma_2) = 0 \) and \( \gamma_2 > r + \delta \), then \( h''(\gamma) > 0 \) so \( h(\gamma_2) \) is a local minimum.
5. \( \min \gamma \) is known with certainty to be constant forever, then the optimal investment-capital ratio, \( \gamma^c(\Phi, r+\delta) \), is the unique value of \( \gamma \in (\gamma^m, r+\delta) \) that satisfies \( H(\gamma, \Phi, r+\delta) = 0 \).

**Corollary 1** If \( r \geq r+\delta \), then \( \frac{\partial^2 \gamma(\Phi, r)}{\partial \Phi^2} \gamma^c(\Phi, r) > 0 \) and \( \frac{\partial^2 \gamma(\Phi, r)}{\partial \rho^2} \gamma^c(\Phi, r) = -\frac{1+r'(\gamma^c)}{\rho-\gamma^c} \gamma^c(\Phi, r) < 0 \).

Corollary 1 states that a firm with a higher deterministic value of marginal operating profit of capital, \( \Phi \), will have a higher optimal value of the investment-capital ratio. It also states that a firm
with a higher user cost of capital, \( r + \delta \), will have a lower optimal value of the investment-capital ratio.

**Corollary 2** If \( \rho \geq r + \delta \), then for given \( \rho \), \( \frac{\partial v}{\partial \rho} = \frac{1}{\rho - \gamma^2} > 0 \), \( \frac{\partial^2 v}{(\partial \rho)^2} = \frac{1}{(\rho - \gamma^2)^2} > 0 \), and \( \frac{\partial v}{\partial \rho} = -\frac{1 + \gamma^2(1) (r^2)}{\rho - \gamma^2} < 0 \).

Corollary 2 states that \( v \), which is the common value of marginal \( q \) and average \( q \), is an increasing convex function of the marginal operating profit of capital, \( \Phi \), and a decreasing function of the user cost of capital.

### 2.2 Example: Quadratic Adjustment Cost

Now consider an example with a quadratic adjustment cost function. Assume that \( c(\gamma) = \theta \gamma^2 \), with \( \theta > 0 \), which implies that \( G \), the set of admissible marginal operating profits of capital, \( \Phi \), is \( \{ \Phi : -\frac{1}{\theta} < \Phi < \theta (r + \delta)^2 + r + \delta \} \). With this quadratic adjustment cost function, and setting \( \rho = r + \delta \), the function \( H(\gamma, \Phi, \rho) \) in equation (11) is \( H(\gamma, \Phi, r + \delta) = \Phi - (r + \delta) - 2\theta \gamma (r + \delta) + \theta \gamma^2 \).

The root of \( H(\gamma, \Phi, r + \delta) = 0 \) that is smaller than \( r + \delta \) is

\[
\gamma^c = \left[ 1 - \sqrt{\frac{1 - \Phi - (r + \delta)}{\theta (r + \delta)^2}} \right] (r + \delta) < r + \delta, \tag{13}
\]

where, as defined earlier, \( \gamma^c \) is the optimal value of \( \gamma \) under certainty.

Since \( v = 1 + c'(\gamma) = 1 + 2\theta \gamma \) at the optimal value of \( \gamma \), equation (13) implies\(^6\)

\[
v = 1 + 2\theta \left[ 1 - \sqrt{\frac{1 - \Phi - (r + \delta)}{\theta (r + \delta)^2}} \right] (r + \delta). \tag{14}
\]

If the marginal operating profit of capital, \( \Phi \), exceeds the user cost of capital, \( r + \delta \), then \( v \), the common value of average \( q \) and marginal \( q \), is greater than one and the optimal investment-capital ratio, \( \gamma^c \), is positive. However, if the marginal operating profit of capital, \( \Phi \), is smaller than the user cost of capital, \( r + \delta \), then \( v \), the common value of average \( q \) and marginal \( q \), is smaller than one and the optimal investment-capital ratio, \( \gamma^c \), is negative. Finally, if the marginal operating profit of capital, \( \Phi \), equals the user cost of capital, \( r + \delta \), then the common value of average \( q \) and marginal \( q \) equals one and the optimal investment-capital ratio is zero.

### 3 Markov Regime-Switching Process for \( \Phi_t \)

In this section I develop and analyze a model of a firm facing stochastic variation in the marginal operating profit of capital, \( \Phi_t \), governed by a Markov regime-switching process. Specifically, a regime is defined by a constant value of \( \Phi_t \). If the marginal operating profit of capital at time \( t \),

---

\(^6\)The values of \( \gamma^c \) and \( v \) in equations (13) and (14) are real because \( \Phi \in G \) implies that \( \Phi < \theta (r + \delta)^2 + r + \delta \), which implies that \( 1 - \frac{\Phi - (r + \delta)}{\theta (r + \delta)^2} > 0 \).
is proportional to the capital stock. Therefore, the average value of the capital stock, \( \Phi_t \), equals \( \phi \) until a new regime arrives. The arrival of a new regime is a Poisson process with probability \( \lambda dt \) of a new arrival during a time interval of length \( dt \). When a new regime arrives, a new value of the marginal operating profit of capital, \( \Phi \), is drawn from a distribution with c.d.f \( F(\Phi) \), where the support of \( F(\Phi) \) is in \( G \), defined in equation (6). \( F(\Phi) \) can be continuous or not continuous, so the random variable \( \Phi \) can be continuous or discrete. The values of \( \Phi \) are drawn independently across regimes.

The Markovian nature of \( \Phi \) implies that the value of the firm at time \( t \) depends only on the capital stock at time \( t \), \( K_t \), and the value of the marginal operating profit at time \( t \), \( \phi \). The value of the firm \( V(K_t, \phi) \) is

\[
V(K_t, \phi) = \max_{\gamma} \int_t^{t+dt} [\phi - \gamma - c(\gamma)] K_t e^{-r(s-t)} ds \\
+ e^{-\lambda dt} e^{-r dt} V(K_{t+dt}, \phi) \\
+ (1 - e^{-\lambda dt}) e^{-r dt} \int_G V(K_{t+dt}, \Phi) dF(\Phi),
\]

which is the maximized sum of three terms. The first term is the present value of \( \pi(K_s, I_s) = [\phi - \gamma_s - c(\gamma_s)] K_s \) over the infinitesimal interval of time from \( t \) to \( t + dt \). The second term is the present value of the firm at time \( t + dt \), conditional on \( \Phi \) remaining equal to \( \phi \) at time \( t + dt \), weighted by the probability, \( e^{-\lambda dt} \), that \( \Phi_{t+dt} = \phi \). The third term is the present value of the expected value of the firm at time \( t + dt \) conditional on a new regime for \( \Phi \) at time \( t + dt \), weighted by the probability that a new regime will arrive by time \( t + dt \).

The Hayashi conditions in Proposition 1 hold in this framework so that the value of the firm is proportional to the capital stock. Therefore, the average value of the capital stock, \( \frac{V(K_t, \phi)}{K_t} \), is independent of the capital stock and depends only on \( \phi \). I will define \( v(\phi) \equiv \frac{V(K_t, \phi)}{K_t} \) to be Tobin’s \( q \), or equivalently, the average value of the capital stock. Since average \( q \) and marginal \( q \) are identically equal in this framework, \( v(\phi) \) is also marginal \( q \).

Use the definition \( v(\phi) \equiv \frac{V(K_t, \phi)}{K_t} \) and the fact that \( \frac{K_{t+dt}}{K_t} = e^{\gamma dt} = e^{(\gamma - \delta) dt} \) and perform the first integration on the right-hand side of equation (15) to obtain

\[
v(\phi) = \max_{\gamma} [\phi - \gamma - c(\gamma)] \frac{1 - e^{-r dt}}{r} \\
+ e^{-\lambda dt} e^{-r dt} v(\phi) \\
+ (1 - e^{-\lambda dt}) e^{-r dt} \int_G v(\Phi) dF(\Phi).
\]

Take the limit of equation (16) as \( dt \) goes to zero to obtain

\[
0 = \max_{\gamma} \phi - \gamma - c(\gamma) - (r + \delta + \lambda - \gamma) v(\phi) + \lambda \Phi,
\]

(17)
where
\[ \pi \equiv \int_G v(\Phi) dF(\Phi) \quad (18) \]
is the unconditional expected value of a unit of capital, which is also the unconditional expected value of both average \( q \) and marginal \( q \).

The maximization in equation (17) has the first-order condition
\[ 1 + c'(\gamma) = v(\phi). \quad (19) \]
Thus, the optimal value of \( \gamma \) equates the marginal cost of investment, including the purchase price of capital and the marginal adjustment cost, with marginal \( q \) and average \( q \).

### 3.1 Marginal \( q \) and Average \( q \)

In this section I present alternative expressions for marginal \( q \) and average \( q \). Because the model presented here is a special case of Proposition 1, marginal \( q \) and average \( q \) are identically equal. Nevertheless, it is helpful to examine different expressions for marginal \( q \) and average \( q \) and to understand why these expressions, which at first glance look different, are equivalent.

Marginal \( q \) at time \( t \) is commonly expressed as the expected present value of the stream of contributions to revenue, less wages and investment costs, of the remaining undepreciated portion of a unit of capital installed at time \( t \), which is
\[ q(\phi) = E \left\{ \int_t^{\infty} \frac{\partial \pi_s(K_s, I_s)}{\partial K_s} e^{-(r+\delta)(s-t)} ds | \Phi_t = \phi \right\}. \quad (20) \]

Average \( q \) at time \( t \) is the value of the firm at time \( t \) divided by \( K_t \). Dividing both sides of equation (2) by \( K_t \), using the linear homogeneity of \( \pi_s(K_s, I_s) \), and using equation (4) and \( M(t, s) = \exp(-r(s-t)) \) yields
\[ v(\phi) = E \left\{ \int_t^{\infty} \pi_s(1, \gamma_s) \exp \left( -\int_t^s (r - g_u) du \right) ds | \Phi_t = \phi \right\}. \quad (21) \]

**Proposition 3** The value of marginal \( q \) is
\[ q(\phi) = \frac{\phi - c(\gamma) + \gamma c'(\gamma)}{r + \delta + \lambda} + \frac{\lambda}{r + \delta + \lambda} \pi, \]
where \( \gamma \) is the optimal value of \( \gamma \) when \( \Phi = \phi \) and \( \pi \equiv \int_G q(\Phi) dF(\Phi) \) is the unconditional expected value of marginal \( q \). The value of average \( q \) is
\[ v(\phi) = \frac{\phi - \gamma - c(\gamma)}{r + \delta + \lambda - \gamma} + \frac{\lambda}{r + \delta + \lambda - \gamma} \pi, \]
where \( \pi \equiv \int_G v(\Phi) dF(\Phi) \) is the unconditional expected value of average \( q \).

Proposition 1 implies that \( q(\phi) \equiv v(\phi) \). However, at first glance, the expressions for \( q(\phi) \) and
Proposition 4 implies that the quadratic equation in \( \gamma \) has two real roots.\(^7\) The optimal value of \( \gamma \) is the smaller root, which is

\[
\gamma = \frac{1}{2} - \frac{\phi - \frac{1}{2} \cdot (r + \delta + \lambda)}{\theta (r + \delta + \lambda)} (r + \delta + \lambda). \tag{27}
\]

Therefore, equations (25) and (27) imply that

\[
q(\phi) = 1 + 2\theta \left[ 1 - \sqrt{1 - \frac{\phi - \frac{1}{2} \cdot (r + \delta + \lambda)}{\theta (r + \delta + \lambda)}^2} \right] (r + \delta + \lambda). \tag{28}
\]

\(^7\)To prove that equation (26) has two real roots, it suffices to show that \( \phi - \frac{1}{2} \cdot (r + \delta + \lambda) < \phi - \frac{1}{2} \cdot (r + \delta + \lambda) \), or equivalently, \( \lambda \gamma < \Omega \equiv (r + \delta + \lambda) + \theta (r + \delta + \lambda)^2 - \phi \). Since \( \phi < \Omega \), \( \phi < \Omega \), and hence \( \Omega > \lambda + \theta (r + \delta + \lambda)^2 - (r + \delta)^2 = \lambda (1 + \theta (r + \delta)) > \lambda (1 + 2\theta (r + \delta)) = \lambda [1 + c'(r + \delta)]. \)

Proposition 4 implies that \( \gamma < 1 + c'(r + \delta) \) so \( \lambda \gamma < \lambda [1 + c'(r + \delta)] < \Omega \).
The expressions for optimal $\gamma$ and $q$ in equations (27) and (28) are closed-form functions of parameters plus one other variable, $\nu$, which is constant over time for any given firm. I discuss $\nu$ further in subsection 3.4.

3.3 Optimal Investment

In this section I exploit the first-order condition for optimal investment in equation (19) to analyze several properties of optimal $\gamma$. The optimal value of $\gamma$ depends on $\nu$. For now, I will treat $\nu$ as a parameter and defer further analysis of $\nu$ to subsection 3.4.

To analyze optimal $\gamma$, substitute the first-order condition for optimal $\gamma$ from equation (19) into equation (17) to obtain

$$0 = \phi - c(\gamma) - (r + \delta + \lambda) - (r + \delta + \lambda - \gamma)c'(\gamma) + \lambda \nu.$$  \hspace{2cm} (29)

Using the definition of $H(\gamma, \Phi, \rho)$ in equation (11), rewrite equation (29) as

$$H(\gamma, \Phi, r + \delta + \lambda) = -\lambda \nu.$$  \hspace{2cm} (30)

Equation (30) characterizes the optimal value of $\gamma$ when there is a constant instantaneous probability, $\lambda$, of a regime switch. Of course, when $\lambda = 0$, this equation is equivalent to equation (12),
which characterizes the optimal value of $\gamma$ under certainty. The optimal value of $\gamma$ when $\lambda = 0$ is shown in Figure 2 as point A where $H(\gamma, \phi, r + \delta) = 0$. The introduction of a positive value of $\lambda$, which introduces stochastic variation in the future values of $\pi_s(1, \gamma_s)$ and $\frac{\partial \pi_s(1, \gamma_s)}{\partial \pi_s}$, has two opposing effects on optimal $\gamma$ in equation (30). First, the introduction of a positive value of $\lambda$ increases the effective user cost of capital, $\rho$, from $r + \delta$ to $r + \delta + \lambda$. This increase in the user cost, $\rho$, reduces the value of $H(\gamma, \phi, \rho)$ by $\lambda (1 + c'(\gamma))$ at each value of $\gamma$, which induces the downward shift of the curve shown in Figure 2. This downward shift of the curve reduces the value of $\gamma$ for which $H(\gamma, \phi, \rho) = 0$, as illustrated by the movement from point A to point B. The value of $\gamma$ for which $H(\gamma, \phi, r + \delta + \lambda) = 0$ is the optimal value of $\gamma$ that would arise if the firm were to disappear, with zero salvage value, when the regime switches. Thus, not surprisingly, the introduction of the possibility of a stochastic death of the firm reduces the value of a unit of capital and reduces the optimal investment-capital ratio. However, if the new regime does not eliminate the firm, there is a second impact on optimal $\gamma$ of the introduction of a positive value of $\lambda$. Specifically, if the firm receives a new draw of $\Phi$ from the unconditional distribution $F(\Phi)$ when the regime changes, then $\pi$ is the expected value of a unit of capital in the new regime. With $\pi > 0$, the term $-\lambda \pi$ on the right-hand side of equation (30) is negative, so that $H(\gamma, \phi, r + \delta + \lambda) < 0$ at the optimal value of $\gamma$. Reducing the value of $H(\gamma, \phi, r + \delta + \lambda)$ from zero to a negative value requires an increase in $\gamma$, as shown in Figure 2 by the movement from point B to point C. To summarize, the introduction of stochastic variation in future $\Phi$ has two opposing effects on the optimal value of $\gamma$. For some values of $\phi$ the introduction of uncertainty will increase the optimal value of $\gamma$, and for other values of $\phi$ it will decrease the optimal value of $\gamma$.

Define $\gamma(\phi, \kappa, r + \delta, \lambda)$ to be the optimal value of $\gamma$ for given values of $r + \delta$ and $\lambda$ if $\Phi = \phi$ and $\pi = \kappa$. Formally, $\gamma(\phi, \kappa, r + \delta, \lambda)$ is defined by

$$H(\gamma(\phi, \kappa, r + \delta, \lambda), \phi, r + \delta + \lambda) = -\lambda \kappa.$$  \hspace{1cm} (31)

Of course, this definition is meaningful only if $\min_{\gamma} H(\gamma, \phi, r + \delta + \lambda) \leq -\lambda \kappa$. The following Lemma identifies an interval of non-negative values of $\kappa$ for which this definition is meaningful.

**Lemma 2** If $0 \leq \kappa < \frac{1}{\lambda}[(r + \delta + \lambda) + c(r + \delta + \lambda) - \phi]$ and $\phi \in G$, then there exists a unique $\gamma(\phi, \kappa, r + \delta, \lambda) \in (\gamma^m, r + \delta + \lambda)$ for which $H(\gamma(\phi, \kappa, r + \delta, \lambda), \phi, r + \delta + \lambda) = -\lambda \kappa$.

Note that $\gamma(\phi, 0, r + \delta, 0) = \gamma^c(\phi, r + \delta)$, which is the optimal value of the investment-capital ratio, $\gamma$, in the case in which $\Phi_t = \phi$ with certainty forever.

The following lemma and its corollary list several properties of the optimal investment-capital ratio $\gamma(\phi, \kappa, r + \delta, \lambda)$ and $c'(\gamma(\phi, \kappa, r + \delta, \lambda))$.

**Lemma 3** Define $\rho \equiv r + \delta + \lambda$. If $\phi \in G$ and if $0 \leq \kappa < \frac{1}{\lambda} (\rho + c(\rho) - \phi)$, then

1. $\frac{\partial \gamma(\phi, \kappa, r + \delta + \lambda)}{\partial \phi} = \frac{1}{(\rho - \gamma)c'(\gamma)} > 0,$
2. \[ \frac{\partial^2 \gamma(\phi, r + \delta, \lambda)}{\partial n} = \frac{\lambda}{(\rho - \gamma)c''(\gamma)} > 0, \]

3. \[ \frac{\partial^2 \gamma(\phi, r + \delta, \lambda)}{\partial t} = - \frac{1 + c'(\gamma)}{(\rho - \gamma)c'(\gamma)} < 0, \]

4. \[ \frac{\partial^2 \gamma(\phi, r + \delta, \lambda)}{\partial \lambda} = - \frac{1 + c'(\gamma) - \kappa}{(\rho - \gamma)c'(\gamma)}. \]

**Corollary 3** Define \( \rho \equiv r + \delta + \lambda \). If \( \phi \in G \) and if \( 0 \leq \kappa < \frac{1}{2} (\rho + c(\rho) - \phi) \), then

1. \[ \frac{\partial c'(\gamma(\phi, r + \delta, \lambda))}{\partial \phi} = \frac{1}{\rho - \gamma} > 0, \]

2. \[ \frac{\partial c'(\gamma(\phi, r + \delta, \lambda))}{\partial n} = \frac{\lambda}{\rho - \gamma} > 0, \]

3. \[ \frac{\partial c'(\gamma(\phi, r + \delta, \lambda))}{\partial t} = - \frac{1 + c'(\gamma)}{\rho - \gamma} < 0, \]

4. \[ \frac{\partial c'(\gamma(\phi, r + \delta, \lambda))}{\partial \lambda} = - \frac{1 + c'(\gamma) - \kappa}{\rho - \gamma} \]

5. \[ \frac{\partial^2 c'(\gamma(\phi, r + \delta, \lambda))}{\partial (\phi)^2} = \frac{1}{(\rho - \gamma)c''(\gamma)} > 0. \]

Lemma 3 and its corollary show for any \( \kappa \in \left[0, \frac{1}{2} (\rho + c(\rho) - \phi)\right] \) and \( \phi \in G \), both \( \gamma(\phi, \kappa, r + \delta, \lambda) \) and \( c'(\gamma(\phi, \kappa, r + \delta, \lambda)) \) are increasing functions of \( \phi \) and \( \kappa \), and decreasing functions of \( r + \delta \). The impact of a higher value of \( \lambda \) depends on the size of \( \phi \). This result is easiest to articulate for the case in which \( \kappa = \pi \), so that \( \kappa \) is the expected value of a unit of capital when \( \phi \) is drawn from the unconditional distribution. In this case, an increase in \( \lambda \) hastens the arrival of a new regime in which the expected value of a unit of capital is \( \kappa \). For values of \( \phi \) that are small enough that \( 1 + c'(\gamma(\phi, \pi, r + \delta, \lambda)) < \kappa = \pi \), hastening the arrival of a new regime increases the value of a unit of capital, thereby increasing optimal \( \gamma \) and the optimal value of \( c'(\gamma) \). Alternatively, for values of \( \phi \) that are large enough that \( 1 + c'(\gamma(\phi, \pi, r + \delta, \lambda)) > \kappa = \pi \), hastening the arrival of a new regime means an earlier end to the current regime with a high \( \phi \). As a result, capital is less valuable and the optimal values of \( \gamma \) and \( c'(\gamma) \) decline. Finally, the corollary shows that \( c'(\gamma(\phi, \kappa, r + \delta, \lambda)) \) is strictly convex in \( \phi \). This convexity will be helpful in subsection 4.2 when I analyze the impact on the value of a unit of capital of a mean-preserving spread in the unconditional distribution \( F(\Phi) \).

### 3.4 The Unconditional Expectation of a Unit of Capital

Equation (30) is a simple expression that characterizes the optimal value of \( \gamma \). However, this expression depends on \( \pi \equiv \int_G v(\Phi) dF(\Phi) \), which is the unconditional expectation of the optimal value of a unit of installed capital. In this subsection, I prove that \( \pi \) is the unique fixed point of particular function and show that this property helps analyze the impact on optimal investment of changes in the distribution \( F(\Phi) \) and changes in \( \lambda \).

Define

\[ \alpha(\kappa) \equiv 1 + \int_G c'(\gamma(\Phi, \kappa, r + \delta, \lambda)) dF(\Phi) \]

as the unconditional expectation of the marginal cost of investment, including the purchase cost of capital and the marginal adjustment cost, where \( \gamma(\phi, \kappa, r + \delta, \lambda) \) is defined in equation (31) as the
optimal value of the investment-capital ratio if $\Phi = \phi$ and $\varpi = \kappa$. Since the value of a unit of capital when $\Phi = \phi$ is $v(\phi) = 1 + c'(\gamma(\phi, \varpi, r + \delta, \lambda))$, optimal behavior by the firm implies that $\varpi$ satisfies $\alpha(\varpi) = \varpi$.

**Lemma 4** Suppose that the support of the distribution $F(\Phi)$ is contained in $G$. The function $\alpha(\kappa) \equiv 1 + \int_G c'(\gamma(\Phi, \kappa, r + \delta, \lambda)) \, dF(\Phi)$ has the following three properties: (1) $\alpha(0) > 0$; (2) $\alpha(1 + c'(r + \delta)) < 1 + c'(r + \delta)$; and (3) $0 < \alpha'(\kappa) < 1$ for $\kappa \in [0, 1 + c'(r + \delta)]$.

Lemma 4 together with the continuity of $\alpha(\kappa)$ leads to the following proposition.

**Proposition 4** Suppose that the support of the distribution $F(\Phi)$ is contained in $G$. Then $\varpi$ is the unique positive value of $\kappa \in (0, 1 + c'(r + \delta)]$ that satisfies $\alpha(\kappa) = \kappa$.

Lemma 4 also leads to the following corollary, which will prove useful in analyzing the effects of changes in the distribution $F(\Phi)$ and changes in $\lambda$.

**Corollary 4** For any $\kappa^* \in [0, 1 + c'(r + \delta)]$, $\text{sign}[\alpha(\kappa^*) - \kappa^*] = \text{sign}[\varpi - \kappa^*]$.

Corollary 4 helps determine the impact on $\varpi$ of changes in the distribution $F(\phi)$ or in $\lambda$. Let $\varpi_0$ be the initial value of $\varpi$ before the change in $F(\phi)$ or in $\lambda$. Then any change that increases $\alpha(\varpi_0)$ will increase $\varpi$, and any change that decreases $\alpha(\varpi_0)$ will decrease $\varpi$.

### 4 Effect of Changing the Stochastic Properties of $\Phi$

In this section I consider the impact of changing the stochastic properties of the marginal operating profit of capital, $\Phi$. Specifically, I consider three changes: (1) replacing the original distribution $F(\phi)$ by a distribution that first-order stochastically dominates the original distribution; (2) introducing a mean-preserving spread on $F(\phi)$; and (3) increasing $\lambda$, the arrival rate of a new value of $\Phi$, which reduces the persistence of $\Phi$.

#### 4.1 $F_2(\Phi)$ First-Order Stochastically Dominates $F_1(\Phi)$

In this subsection, I analyze a change in the distribution $F(\Phi)$ from $F_1(\Phi)$ to $F_2(\Phi)$, where $F_2(\Phi)$ first-order stochastically dominates $F_1(\Phi)$. Let $\varpi_i$ be the unconditional expected value of a unit of capital when the distribution of $\Phi$ is $F_i(\Phi)$, $i = 1, 2$. Also, let $\gamma(\Phi, \varpi_i, r + \delta, \lambda)$ be the optimal value of $\gamma$ for given $\Phi$ when the distribution of $\Phi$ is $F_i(\Phi)$, and let $\Gamma_i(\gamma)$ be the induced distribution of the optimal value of $\gamma$ when the distribution of $\Phi$ is $F_i(\Phi)$, $i = 1, 2$.

**Proposition 5** If $F_2(\Phi)$ strictly first-order stochastically dominates $F_1(\Phi)$, then $\varpi_2 > \varpi_1$ and $\Gamma_2(\gamma)$ strictly first-order stochastically dominates $\Gamma_1(\gamma)$.

Proposition 5 states that moving to a more favorable distribution of $\Phi$ that first-order stochastically dominates the original distribution will increase $\varpi$, the average value of a unit of capital. The
increase in \( \tau \) will increase the optimal value of \( \gamma \) at each value of \( \Phi \), and because the distribution of \( \Phi \) becomes more favorable and optimal \( \gamma \) is increasing in \( \Phi \), the distribution of optimal \( \gamma \) also moves toward larger values in the sense of first-order stochastic dominance.

### 4.2 A Mean-Preserving Spread on \( F(\Phi) \)

Now consider the effect on optimal investment of a mean-preserving spread on the distribution \( F(\Phi) \). This question was first addressed in a model with convex costs of adjustment by Hartman (1972) and then by Abel (1983). In both papers, the production function is linearly homogeneous in capital and labor and the firm is perfectly competitive, so that, as in this paper, the marginal operating profit of capital, \( \Phi \), is independent of the capital stock. Hartman and Abel both found that an increase in the variance of the price of output leads to an increase in the optimal rate of investment.\(^8\)

The channel through which this effect operates is the convexity of \( \Phi_t \equiv \max_t [p_t A_t f(1, l) - w_t l] \) in \( p_t A_t \) and \( w_t \). This convexity implies that a mean-preserving spread on \( p_t A_t \) or \( w_t \) at some future time \( t \) increases the expected value of future \( \Phi_t \) and thus increases the expected present value of the stream of future \( \Phi_t \), which increases (marginal) \( q \) and hence increases investment. In the current paper, I analyze a different channel for increased uncertainty to affect investment. To focus on that channel, I analyze mean-preserving spreads in the distribution of \( \Phi_t \) directly. Since the expected value of \( \Phi_t \) remains unchanged by construction, any effects on the optimal value of \( \gamma \) will operate through a different channel than in Hartman (1972) and Abel (1983).

**Proposition 6** A mean-preserving spread of \( F(\Phi) \) that maintains the support within \( G \) increases \( \tau \).

The proof of Proposition 6 is in the Appendix, but it is helpful to examine a key step to get a sense for what is driving the result. As shown in the Appendix, this result relies on the fact that \( c'(\gamma(\Phi, \kappa, r + \delta, \lambda)) \) is convex in \( \Phi \), even though \( c'(\gamma) \) may not be convex in \( \gamma \) and \( \gamma(\Phi, \kappa, r + \delta, \lambda) \) may not be convex in \( \Phi \). Notice that \( c'(\gamma(\Phi, \kappa, r + \delta, \lambda)) \) will be convex in \( \Phi \) if \( \frac{\partial c'(\gamma(\Phi, \kappa, r + \delta, \lambda))}{\partial \Phi} \) is increasing in \( \Phi \). However, neither \( c''(\gamma(\Phi, \kappa, r + \delta, \lambda)) \) nor \( \frac{\partial c(\Phi, \kappa, r + \delta, \lambda)}{\partial \Phi} \) is necessarily increasing in \( \Phi \). But their product, \( \frac{1}{\rho - \gamma(\Phi, \kappa, r + \delta, \lambda)} \), is increasing in \( \Phi \), so \( c'(\gamma(\Phi, \kappa, r + \delta, \lambda)) \) is convex in \( \Phi \). Therefore, a mean-preserving spread on \( \Phi \) increases the unconditional expected value of \( c'(\gamma(\Phi, \kappa, r + \delta, \lambda)) \) and hence increases \( \tau \).

The following corollary uses the fact that with a quadratic adjustment cost function \( c(\gamma) \), the marginal adjustment cost function is linear in \( \gamma \), so that optimal \( \gamma \) is a linear function of \( q \), or equivalently, \( v \).

---

\(^8\) Caballero (1991) showed that positive impact of uncertainty on optimal investment can be reversed by relaxing the assumption of perfect competition or by relaxing the linear homogeneity of the production function in capital and labor.
Corollary 5  If the adjustment cost function \( c(\gamma) \) is quadratic, then a mean-preserving spread of \( F(\Phi) \) that maintains the support within \( G \) increases \( \mathcal{T} \equiv \int_G \gamma(\Phi, \tau, r + \delta, \lambda) dF(\Phi) \), the unconditional expected value of \( \gamma \).

4.3  A Change in Persistence of Regimes

Now consider a change in the persistence of regimes governing \( \Phi \). With a constant probability \( \lambda \) of a switch in the regime, the expected life of a regime is \( \frac{1}{\lambda} \), so an increase in \( \lambda \) reduces the persistence of the regime.

Proposition 7  If \( F(\Phi) \) is non-degenerate, then \( \frac{\partial \mathcal{R}}{\partial \lambda} < 0 \), so that an increase in the persistence of regimes (which is a reduction in \( \lambda \)) increases \( \mathcal{T} \).

An increase in \( \lambda \) hastens the arrival of a new regime and therefore diminishes the contribution of the current regime to the expected present value of the future cash flows to the firm. For low values of current \( \phi \), diminishing the contribution of the current regime increases the value of the firm, and for high values of current \( \phi \), diminishing the contribution of the current regime reduces the value of the firm. Since the optimal value of the investment-capital ratio \( \gamma \) is higher when \( \phi \) is higher, the growth rate of the capital stock in the current regime is higher when \( \phi \) is high than when \( \phi \) is low. Therefore, for a given capital stock at the beginning of the current regime, the future capital stock will be higher throughout the current regime when the current value of \( \phi \) is high than when the current value of \( \phi \) is low. Therefore, as shown in the proof of Proposition 7 in the Appendix, the reduction in the value of the firm resulting from an increase in \( \lambda \) when \( \phi \) is high outweighs the increase in the value of the firm resulting from an increase in \( \lambda \) when \( \phi \) is low. Hence, an increase in \( \lambda \) reduces \( \mathcal{T} \).

For each regime, the optimal value of the investment-capital ratio, \( \gamma \), moves in the same direction as the value of a unit of capital moves when \( \lambda \) increases. The following corollary exploits the fact that in the case of quadratic adjustment costs, the optimal value of \( \gamma \) is a linear function of \( v(\phi) \).

Corollary 6  If \( F(\Phi) \) is non-degenerate and if \( c(\gamma) \) is quadratic, then \( \frac{\partial \mathcal{T}}{\partial \lambda} < 0 \).

5  Measurement Error and the Cash Flow Effect on Investment

The model developed in this paper focuses on three variables that are often used in empirical studies of investment, specifically, the investment-capital ratio, \( \gamma \), the value of a unit of capital, \( v \), which is Tobin’s \( q \), and cash flow per unit of capital, \( \Phi \). This model, like most existing models, uses the first-order condition for optimal investment, \( 1 + c'(\gamma) = v(\phi) \) (equation 19), to draw a tight link between \( \gamma \) and \( v \). This link is often described by saying that \( v \) is a sufficient statistic for \( \gamma \), meaning that if an observer knows the adjustment cost function and the value of \( v \), then the value
of $\gamma$ can be computed in a straightforward manner without any additional information or knowledge of the values of any other variables. Indeed, if the adjustment cost function, $c(\gamma)$, is quadratic, the marginal adjustment cost function is linear, and optimal $\gamma$ is a linear function of $v$.

The empirical literature has a long history of finding that $v$ is not a sufficient statistic for $\gamma$. In particular, at least since the work of Fazzari, Hubbard, and Petersen (1988), researchers have found that in a regression of $\gamma$ on $v$ and $\Phi$, estimated coefficients on both $v$ and $\Phi$ tend to be positive and statistically significant. The finding of a positive significant coefficient on cash flow, $\Phi$, is often interpreted as evidence that firms face financing constraints or some other imperfection in financial markets. This interpretation of financial frictions, as they are sometimes known, is bolstered by the finding that in firms that one might suspect to be more likely to face these frictions, the cash flow effects tend to be more substantial. For instance, as the argument goes, firms that are growing rapidly may encounter more substantial financial frictions, and it turns out that cash flow coefficients are often larger for such firms.\(^9\)

In this section, I will offer a different interpretation of the cash flow coefficients. I will demonstrate that if $v$ is observed with classical measurement error, then the coefficient of $v$ is biased toward zero and, more importantly, the coefficient on cash flow, $\Phi$, will be positive, even though in the absence of measurement error in $v$, the coefficient on $\Phi$ would be zero. The fact that measurement error in $v$ can affect the coefficient estimates in this way been pointed out by Erickson and Whited (2000) and Gilchrist and Himmelberg (1995) and others, though the particular simple expressions I present in this paper appear to be new. More novel, however, is the analytical demonstration that cash flow coefficients will be larger in firms that grow more rapidly. The finding that measurement error in $v$ can lead to a positive cash flow coefficient does not use the particular model in this paper, other than the result that $v$ and $\Phi$ are positively correlated with each other. However, the model in this paper is used in the more novel demonstration that cash flow coefficients are larger for firms that have higher growth rates. Although the literature interprets the empirical finding of larger cash flow coefficients for more rapidly growing firms as further evidence of financial frictions, the model here has no financial frictions whatsoever, and yet leads to the same finding. Therefore, the finding of positive cash flow coefficients, including larger coefficients for firms that are growing more rapidly, does not necessarily show that financial frictions are important or operative.

To isolate measurement error from specification error, I assume that the adjustment cost function is quadratic so that optimal $\gamma$ is a linear function of $v$. As before, the quadratic adjustment cost function is $c(\gamma) = \theta \gamma^2$ so the first-order condition for optimal $\gamma$ in equation (19) implies that

$$\gamma = \frac{v - 1}{2\theta}.$$  \hspace{1cm} (33)

\(^9\)For instance, Deveraux and Schiantarelli (1990) state "The perhaps surprising result from table 11.7 is that the coefficient on cash flow is greater for firms operating in growing sectors." (p. 298).
Assume that the manager of the firm can observe \( v, \Phi, \) and \( \gamma \) without error, but people outside the firm, including the econometrician, observe these variables with classical measurement error. Specifically, the econometrician observes the value of a unit of capital as \( \bar{q} = v + \varepsilon_q \), the investment-capital ratio as \( \bar{\gamma} = \gamma + \varepsilon_\gamma \), and cash flow as \( \bar{c} = \Phi + \varepsilon_c \), where the observation errors \( \varepsilon_q, \varepsilon_\gamma, \) and \( \varepsilon_c \) are mean zero, mutually independent, and independent of \( v, \Phi, \) and \( \gamma \). Erickson and Whited (2000) offer a useful taxonomy of reasons for measurement error in \( q \), and except for differences between marginal \( q \) and average \( q \) (which are non-existent in the model presented here), those reasons could apply here.

Consider a linear regression of \( \bar{\gamma} \) on \( \bar{q} \) and \( \bar{c} \), after all variables have been de-meaned. Let \( b_q \) and \( b_c \) be the plims of the estimated coefficients on \( \bar{q} \) and \( \bar{c} \), respectively, so

\[
\begin{bmatrix}
    b_q \\
    b_c 
\end{bmatrix} = \begin{bmatrix}
    \text{Var}(\bar{q}) & \text{Cov}(\bar{q}, \bar{c}) \\
    \text{Cov}(\bar{q}, \bar{c}) & \text{Var}(\bar{c})
\end{bmatrix}^{-1} \begin{bmatrix}
    \text{Cov}(\bar{q}, \bar{\gamma}) \\
    \text{Cov}(\bar{c}, \bar{\gamma})
\end{bmatrix}.
\] (34)

The variance-covariance matrix, \( A \), of \( (\bar{q}, \bar{c}, \bar{\gamma}) \) conveniently displays the variances and covariances in equation (34), where

\[
A = \begin{bmatrix}
    \text{Var}(v) + \text{Var}(\varepsilon_q) & \text{Cov}(v, \Phi) & \frac{1}{2\pi} \text{Var}(v) \\
    \text{Cov}(v, \Phi) & \text{Var}(\Phi) + \text{Var}(\varepsilon_c) & \frac{1}{2\pi} \text{Cov}(v, \Phi) \\
    \frac{1}{2\pi} \text{Var}(v) & \frac{1}{2\pi} \text{Cov}(v, \Phi) & \frac{1}{2\pi} \text{Var}(v) + \text{Var}(\varepsilon_\gamma)
\end{bmatrix}.
\] (35)

Substituting the relevant second moments from equation (35) into equation (34), and performing the indicated matrix inversion and matrix multiplication yields

\[
\begin{bmatrix}
    b_q \\
    b_c 
\end{bmatrix} = \frac{1}{2\pi} \begin{bmatrix}
    \text{Var}(v) + \text{Var}(\varepsilon_q) & \text{Cov}(v, \Phi) & \frac{1}{2\pi} \text{Var}(v) \\
    \text{Cov}(v, \Phi) & \text{Var}(\Phi) + \text{Var}(\varepsilon_c) & \frac{1}{2\pi} \text{Cov}(v, \Phi) \\
    \frac{1}{2\pi} \text{Var}(v) & \frac{1}{2\pi} \text{Cov}(v, \Phi) & \frac{1}{2\pi} \text{Var}(v) + \text{Var}(\varepsilon_\gamma)
\end{bmatrix}^{-1} \begin{bmatrix}
    \text{Cov}(\bar{q}, \bar{\gamma}) \\
    \text{Cov}(\bar{c}, \bar{\gamma})
\end{bmatrix}.
\] (36)

Define \( s_q^2 \equiv \frac{\text{Var}(\varepsilon_q)}{\text{Var}(v)} \) as the variance of the measurement error in \( \bar{q} \), normalized by \( \text{Var}(v) \), which is the variance of the true value of \( q \), \( s_c^2 \equiv \frac{\text{Var}(\varepsilon_c)}{\text{Var}(\Phi)} \) as the variance of the measurement error in cash flow normalized by the variance of the true value of cash flow, and \( R^2 = \frac{[\text{Cov}(v, \Phi) \text{Var}(v)]^2}{\text{Var}(\Phi) \text{Var}(v)} \) as the squared correlation between the true values of \( q \) and cash flow. Dividing both the numerators and denominators of \( b_q \) and \( b_c \) in equation (36) by \( \text{Var}(\Phi) \text{Var}(v) \) yields

\[
\begin{bmatrix}
    b_q \\
    b_c 
\end{bmatrix} = \frac{1}{2\pi} \begin{bmatrix}
    \text{Var}(v) + \text{Var}(\varepsilon_q) & \text{Cov}(v, \Phi) & \frac{1}{2\pi} \text{Var}(v) \\
    \text{Cov}(v, \Phi) & \text{Var}(\Phi) + \text{Var}(\varepsilon_c) & \frac{1}{2\pi} \text{Cov}(v, \Phi) \\
    \frac{1}{2\pi} \text{Var}(v) & \frac{1}{2\pi} \text{Cov}(v, \Phi) & \frac{1}{2\pi} \text{Var}(v) + \text{Var}(\varepsilon_\gamma)
\end{bmatrix}^{-1} \begin{bmatrix}
    \text{Cov}(\bar{q}, \bar{\gamma}) \\
    \text{Cov}(\bar{c}, \bar{\gamma})
\end{bmatrix}.
\] (37)

Equation (37) shows the impact of measurement error in \( q \). If \( q \) is perfectly measured, then
$s^2_q = 0$ and, regardless of whether cash flow is measured with error, equation (37) immediately yields $b_q = \frac{1}{\sigma^2}$ and $b_c = 0$. Thus, if $q$ is perfectly measured, $b_q$ equals the derivative of the optimal value of $\gamma$ with respect to $v$ in the first-order condition in equation (33). In addition, the estimated effect of cash flow on investment, $b_c$, is zero. However, if $q$ is measured with error, so that $s^2_q > 0$, then, $b_q$, the estimated coefficient on $q$ is smaller than $\frac{1}{\sigma^2}$, the true derivative of $\gamma$ with respect to $v$. Moreover, if $s^2_q > 0$, then $b_c$, the estimated coefficient on cash flow can be nonzero; in fact, if $q$ and cash flow are positively correlated, the estimated cash flow coefficient, $b_c$, is positive. Much of the investment literature interprets a significantly positive coefficient on cash flow as evidence of financing constraints. Yet equation (37) demonstrates that measurement error in $q$ will lead to a positive coefficient on cash flow, provided that $q$ and cash flow are positively correlated, even if there are no financial frictions. This argument is not restricted to the particular specification of the firm in this model, and has been made less formally by, for example, Gilchrist and Himmelberg (1995).11 The model in this paper allows the analysis to go one step further and to account for differences in the estimated cash flow coefficients for firms with different growth rates, as I discuss next.

Proponents of the view that positive cash flow coefficients are evidence of financing constraints bolster their view by showing that firms that are likely to face binding financing constraints are likely to exhibit larger, more significant positive cash flow coefficients. For instance, they argue that firms that are growing more quickly are more likely to face binding financing constraints. Empirical evidence that rapidly growing firms have larger, significant positive cash flow coefficients is then presented as evidence of financing constraints. However, the model in this paper offers an alternative interpretation. Equation (37) shows that the cash flow coefficient is proportional to $\frac{\text{Cov}(v, \Phi)}{\text{Var}(\Phi)}$, which is the population regression coefficient of $v$ on $\Phi$. The analog of this coefficient in the model is $\frac{\partial c'(\gamma)}{\partial \Phi}$, which equals $\frac{\partial c'(\gamma(\Phi, \tau, \lambda))}{\partial \Phi}$ because $v(\Phi) = 1 + c'(\gamma(\Phi, \tau, \lambda))$. Since $\frac{\partial c'(\gamma(\Phi, \tau, \lambda))}{\partial \Phi} = \frac{1}{\rho - \gamma}$ (Corollary 3), $\frac{\partial c'(\gamma)}{\partial \Phi} = \frac{1}{\rho - \gamma}$, which is increasing in the growth rate of capital, $\gamma - \delta$, for a given depreciation rate $\delta$. Therefore, the cash flow coefficient is increasing in the growth rate of the firm.

To use the model to compare the investment behavior of a slowly growing firm and a rapidly growing firm, I will consider firms that face different unconditional distributions, $F(\Phi)$, of $\Phi$, that endogenously lead to different growth rates. The following Proposition states that the more rapidly growing firm will have a higher cash flow coefficient, which is proportional to $\frac{\partial c'(\gamma)}{\partial \Phi}$, than the more slowly growing firm, even though there are no financial frictions in the model.

---


11 Gilchrist and Himmelberg, p. 544, state "More generally, anything that systematically reduces the signal-to-noise ratio of Tobin’s Q (for example, measurement error or ‘excess volatility’ of stock prices) will shift explanatory power away from Tobin’s Q toward cash flow, thus making such firms appear to be financially constrained when in fact they are not."
Proposition 8 Consider two firms with identical quadratic adjustment cost functions but with different unconditional distributions of $\Phi$, $F_1(\Phi)$ and $F_2(\Phi)$, which imply different unconditional values of capital, $\bar{\nu}_1$ and $\bar{\nu}_2$. If $F_2(\Phi)$ strictly first-order stochastically dominates $F_1(\Phi)$, then

1. $\gamma(\Phi, \bar{\nu}_2, r + \delta, \lambda) > \gamma(\Phi, \bar{\nu}_1, r + \delta, \lambda)$
2. $\int_G \gamma(\Phi, \bar{\nu}_2, r + \delta, \lambda) dF_2(\Phi) > \int_G \gamma(\Phi, \bar{\nu}_1, r + \delta, \lambda) dF_1(\Phi)$
3. $\frac{d\nu_2(\Phi)}{d\Phi} > \frac{d\nu_1(\Phi)}{d\Phi}$ and
4. $\int_G \frac{d\nu_2(\Phi)}{d\Phi} dF_2(\Phi) > \int_G \frac{d\nu_1(\Phi)}{d\Phi} dF_1(\Phi)$.

Proposition 8 states that the firm with distribution $F_2(\Phi)$ is the faster-growing firm, whether the speed of growth is measured by the investment-capital ratio at any given value of $\Phi$ (statement 1) or by the unconditional expectation of the investment-capital ratio (statement 2). This proposition also states that the firm with the distribution $F_2(\Phi)$ has the higher value of $\frac{\partial \nu_2(\Phi)}{\partial \Phi}$ for a given value of $\Phi$ (statement 3) and the higher unconditional expected value of $\frac{\partial \nu(\Phi)}{\partial \Phi}$ (statement 4). Therefore, the firm with the distribution $F_2(\Phi)$ has the higher value of $\frac{\partial \nu(\Phi)}{\partial \Phi}$ and hence the higher cash flow coefficient. To summarize, the firm that is growing more rapidly has the larger coefficient on cash flow, even though the are no financial frictions in this model.

6 Concluding Remarks

This paper develops a model of a competitive firm with constant returns to scale to provide a tractable and useful stochastic framework to analyze the behavior and interrelationships of optimal investment, $q$, and cash flow that are widely studied in the empirical literature. As first shown by Hayashi (1982), average $q$ and marginal $q$ are identically equal in this framework. Within the class of models for which average $q$ and marginal $q$ are equal, the model presented here places only one additional restriction on technology, namely that adjustment costs, which are a function of investment and the capital stock, are additively separable from the production function for output, which is a function of capital and labor. For convenience, the model specifies a constant discount rate and a constant depreciation rate of capital. Finally, the analysis of the stochastic model is greatly facilitated by the simple Markov regime-switching specification for the marginal operating profit capital.

The model developed here is tractable enough to analyze various aspects of optimal investment behavior in a framework that is consistent with empirical analyses that use average $q$ to measure marginal $q$ and that specify the investment-capital ratio as a function of $q$. A closed-form solution for optimal investment and $q$ is derived only for the case in which the marginal operating profit of capital is known to be constant and the cost of adjustment function is quadratic. When the marginal operating profit of capital follows a Markov regime-switching process, I present analytic expressions
for the optimal investment-capital ratio and the value of a unit of capital that are closed-form up to a single undetermined scalar constant.

After demonstrating various properties of optimal investment and \( q \), I use the model to analyze the effects of three changes in the stochastic environment facing the firm. First, a favorable shift in the unconditional distribution of the marginal operating profit of capital, in the sense of first-order stochastic dominance, increases the expected value of a unit of capital, and shifts the distribution of the optimal investment-capital ratio in a first-order stochastically dominating way. Second, a mean-preserving spread in the unconditional distribution of the marginal operating profit of capital increases the average value of a unit of capital, as in the existing literature, though the channel of the effect is different than in previous studies. Third, an increase in the persistence of regimes increases the average value of a unit of capital.

To address the common empirical finding of a positive coefficient on cash flow in a regression of the investment-capital ratio on \( q \) and cash flow, I introduce classical measurement error. Consistent with existing arguments, I show that measurement error in \( q \) can lead to a positive coefficient on cash flow. However, I use the model to go a step further and demonstrate that the model can account for the finding of larger cash flow coefficients for firms that grow more rapidly. Proponents of the importance of financing constraints point to the positive coefficient on cash flow as evidence of the importance of these constraints. Moreover, they argue that larger cash flow coefficients for firms likely to be constrained, such as rapidly growing firms, support the interpretation that positive cash flow coefficients indicate the importance of financing constraints. However, the model presented here has no financing constraints at all, yet in the presence of classical measurement error, it predicts coefficients on cash flow that are both positive and are larger for firms that grow more rapidly.
References


Appendix: Proofs of Lemmas, Propositions and Corollaries

**Proof of Proposition 1**: Let \( \{K^A_s, I^A_s\}_{s=t}^{s=\infty} \) satisfy the capital accumulation equation in (1) and attain the maximum on the right-hand side of (2). Let \( \{K^B_s, I^B_s\}_{s=t}^{s=\infty} = \{\omega K^A_s, \omega I^A_s\}_{s=t}^{s=\infty} \) for an arbitrary \( \omega > 0 \) and note that \( \{K^B_s, I^B_s\}_{s=t}^{s=\infty} \) satisfies the capital accumulation equation in (1). Then \( V_t(\omega K^A_t) = V_t(K^B_t) \geq E_t \{ \int_t^\infty \pi_s (K^B_t)^M (t, s) \, ds \} = E_t \{ \int_t^\infty \pi_s (\omega K^A_s, \omega I^A_s) \, ds \} = \omega E_t \{ \int_t^\infty \pi_s (K^A_s, I^A_s) \, ds \} = \omega V_t(K^A_t) \). Since \( V_t(\omega K_i) \geq \omega V_t(K_i) \) for any \( \omega > 0 \) and any \( K_i > 0 \), we have \( V_t(\frac{\omega}{\omega} K^B_t) \geq V_t(K^B_t) \), which implies \( V_t(K^A_t) = V_t(K^B_t) \geq V_t(K^A_t) \). Therefore, \( V_t(\omega K^A_t) \geq \omega V_t(K^A_t) \geq V_t(\omega K^A_t) \), which implies \( V_t(\omega K^A_t) = V_t(K^A_t) \).

**Proof of Lemma 1**: (1) Inspection of the definition of \( H(\gamma, \Phi, \rho) \) immediately reveals that \( H(\gamma, \Phi, \rho) \) is an increasing linear function of \( \Phi \). (2) Differentiating \( H(\gamma, \Phi, \rho) \) with respect to \( \rho \) yields \( H_\rho = -(1 + c'(\gamma)) \), which is negative for \( \gamma > \gamma_m \). (3) Differentiating \( H(\gamma, \Phi, \rho) \) with respect to \( \gamma \) yields \( H_\gamma(\gamma, \Phi, \rho) = -(\rho - \gamma) c''(\gamma) \). Since \( c'(\gamma) \) is strictly convex, \( c''(\gamma) > 0 \). Therefore, \( H(\gamma, \Phi, \rho) \) is strictly decreasing in \( \gamma \) for \( \gamma < \rho \), strictly increasing in \( \gamma \) for \( \gamma > \rho \), and minimized with respect to \( \gamma \) at \( \gamma = \rho \). That is, \( H(\gamma, \Phi, \rho) \) is strictly quasi-convex in \( \gamma \). (4) Since \( c'(\gamma_m) = -1 \), \( H(\gamma_m, \Phi, \rho) = \Phi - (\rho - \gamma_m) c'(\gamma_m) \). (5) Since \( H(\gamma, \Phi, \rho) = -(\rho - \gamma) c''(\gamma) \) and \( H(\gamma, \Phi, \rho) \) is strictly quasi-convex in \( \gamma \), \( H(\gamma, \Phi, \rho) \) is minimized with respect to \( \gamma \) at \( \gamma = \rho \). Therefore, \( \min H(\gamma, \Phi, \rho) = H(\rho, \Phi, \rho) = \Phi - c(\gamma_m) \leq \Phi - (\rho + \delta) - c(\rho + \delta) < 0 \), where the first inequality follows from \( c'(\rho + \delta) > 0 \), the second inequality follows from the assumption that \( \Phi \in G \). (6) \( H(\gamma_m, \Phi, \rho) > 0 > H(\rho, \Phi, \rho) \) and the strict quasi-convexity of \( H(\gamma, \Phi, \rho) \) in \( \gamma \) imply that \( H(\gamma, \Phi, \rho) = 0 \) for exactly two distinct values of \( \gamma \), denoted \( \gamma_1 \) and \( \gamma_2 \), with \( \gamma_m < \gamma_1 < \rho \) and \( \gamma_2 > \rho \). \( H(\gamma_i, \Phi, \rho) = 0 \), for \( i = 1, 2 \), implies \( \Phi - c(\gamma_1 i) = (\rho - \gamma_i) [1 + c'(\gamma_i)], \ i = 1, 2 \). Since \( 1 + c'(\gamma_m) > 0 \), \( c''(\gamma_m) = 0 \), \( \Phi - \gamma_1 - c(\gamma_1) > 0 \) for \( \gamma_1 \in (\gamma_m, \rho) \) and \( \Phi - \gamma_2 - c(\gamma_2) < 0 \) for \( \gamma_2 \in (\rho, \infty) \). ■

**Proof of Proposition 2**: The optimal value of \( \gamma \) is a root of \( H(\gamma, \Phi, \rho) = 0 \). Lemma 1 states that this equation has two roots: \( \gamma_1 \in (\gamma_m, \rho) \) and \( \gamma_2 \in (\rho, \infty) \) and that \( \Phi - \gamma_1 - c(\gamma_1) > 0 > \Phi - \gamma_2 - c(\gamma_2) \). Therefore, \( \gamma_1 \) is the optimal value of the investment-capital ratio. ■

**Proof of Corollary 1**: Apply the implicit function theorem to \( H(\gamma^c, \Phi, \rho) = 0 \) and use the facts that \( c'(\gamma) < r + \delta \), \( 1 + c'(\gamma^c) > 0 \), and \( c''(\gamma^c) > 0 \) to obtain \( \frac{\partial H(\gamma, \Phi, \rho)}{\partial \rho} = -H_\rho = -\frac{1}{(\rho - \gamma)^2 c''(\gamma^c)} > 0 \) and \( \frac{\partial H(\gamma, \Phi, \rho)}{\partial \gamma} = -\frac{1}{(\rho - \gamma)^2 c''(\gamma^c)} < 0 \). ■

**Proof of Corollary 2**: Differentiate the first-order condition for optimal investment, which holds at all points of time, \( v = 1 + c'(\gamma^c) \), with respect to \( \Phi \) to obtain \( \frac{\partial v}{\partial \Phi} = c''(\gamma^c) \frac{\partial^2 H(\gamma, \Phi, \rho)}{\partial \rho^2} = c''(\gamma^c) \frac{1}{(\rho - \gamma)^2 c''(\gamma^c)} > 0 \). Differentiate the expression for \( \frac{\partial v}{\partial \rho} \) with respect to \( \Phi \) to obtain \( \frac{\partial^2 v}{\partial \Phi^2} = \frac{1}{(\rho - \gamma)^3 c''(\gamma^c)} > 0 \). Differentiate \( v = 1 + c'(\gamma^c) \), with respect to \( \rho \) to obtain \( \frac{\partial v}{\partial \rho} = c''(\gamma^c) \frac{1}{(\rho - \gamma)^2 c''(\gamma^c)} > 0 \). ■
Proof of Proposition 3: Calculation of $q(\phi)$: Suppose that $\Phi_s = \phi$ for all $t \leq s < x$ and the regime switches at time $x$, with a new drawing of $\Phi$, say $\Phi_x$, from the unconditional distribution $F(\Phi)$. The expression for $\pi_s(K_s, I_s)$ in equation (5) can be written as $\pi_s(K_s, I_s) = \left[ \phi - c \left( \frac{I_s}{K_s} \right) \right] K_s - I_s$. Therefore, $\frac{\partial \pi_s(K_s, I_s)}{\partial K_s} = \phi - c(\gamma_\delta) + \gamma_\delta c'(\gamma_\delta)$, which equals $\phi - c(\gamma_\delta) + \gamma_\delta c'(\gamma_\delta)$ for all $t \leq s < x$. As for the stream of marginal contributions of capital accruing from time $x$ onward, their expected present value as of time $x$ is $e^{-\delta(x-t)}q(\Phi_x)$; the expected present value of $e^{-\delta(x-t)}q(\Phi_x)$ as of time $t$ is $e^{-(r+\delta)(x-t)}q$, where $q = \int_G q(\Phi) dF(\Phi)$ is the unconditional expected value of $q(\Phi_x)$. Therefore, $q_{t|x}$, the value of $q_t$, conditional on the next regime switch occurring at time $x > t$ is

$$q_{t|x} = 1 - e^{-(r+\delta)(x-t)} \left( \frac{\phi - c(\gamma_\delta) + \gamma_\delta c'(\gamma_\delta)}{r + \delta} \right) + e^{-(r+\delta)(x-t)}q.$$

The first term on the right-hand side of equation (38) is the present value of $\frac{\partial \pi_s(K_s, I_s)}{\partial K_s} e^{-\delta(s-t)}$ from time $t$ to time $x$. The second term is the expected present value, discounted to time $t$, of $\frac{\partial \pi_s(K_s, I_s)}{\partial K_s} e^{-\delta(s-t)}$ from time $x$ onward.

The probability that the first switch in the regime after time $t$ occurs at time $x > t$ is $\lambda e^{-\lambda(x-t)}$, so that

$$q(\phi) = \int_t^\infty \lambda e^{-\lambda(x-t)} q_{t|x} dx.$$  

Substituting equation (38) into equation (39) and performing the integration yields

$$q(\phi) = \frac{\phi - c(\gamma_\delta) + \gamma_\delta c'(\gamma_\delta)}{r + \delta + \lambda} + \frac{\lambda}{r + \delta + \lambda} q,$$

where $\gamma$ in this equation is the optimal value of $\gamma$ when $\Phi = \phi$.

Calculation of $v(\phi)$: Continue to suppose that $\Phi_s = \phi$ for all $t \leq s < x$ and the regime switches at time $x$, with a new drawing of $\Phi$ from the unconditional distribution $F(\Phi)$. With $\pi_s(K_s, I_s)$ specified as in equation (5), $\pi_s(1, \gamma_\delta) = \phi - \gamma_\delta - c(\gamma_\delta)$ for all $t \leq s < x$ and $g_u = g_t = \gamma_\delta - \delta$ for all $t \leq u < x$. Therefore, $v_{t|x}$, the value of $v_t$, conditional on the next regime switch occurring at time $x > t$, is

$$v_{t|x} = 1 - e^{-(r+\delta-\gamma)\gamma_\delta}(x-t)} \frac{\phi - c(\gamma_\delta) - c(\gamma_\delta) - c(\gamma_\delta)}{r + \delta + \gamma} + e^{-(r+\delta-\gamma)\gamma_\delta}(x-t)}q.$$

where $q = \int_G q(\Phi) dF(\Phi)$ is defined in equation (18). The first term on the right-hand side of equation (41) is the present value of $\pi_s(1, \gamma_\delta) \exp \left( - \int_t^x g_u du \right) = \pi_t(1, \gamma_\delta) e^{(\gamma_\delta - \delta)(s-t)}$ from time $t$ to time $x$. The second term on the right-hand side of equation (41) is the expected present value, discounted to time $t$, of $\pi_s(1, \gamma_\delta) \exp \left( - \int_t^x g_u du \right)$ from time $x$ onward.

Since the probability that the first switch in the regime after time $t$ occurs at time $x > t$ is
\footnote{The equality of $\frac{\partial \pi_s(K_s, I_s)}{\partial K_s} e^{-\delta(s-t)}$ and $\frac{\partial \pi_s(1, \gamma_\delta)}{\partial K_s} e^{-\delta(s-t)}$ is an implication of the linear homogeneity of $\pi_s(K_s, I_s)$.}
\(\lambda e^{-\lambda(x-t)}\), the average value of capital is

\[
v(\phi) = \int_t^\infty \lambda e^{-\lambda(x-t)} v_t(x) dx.
\]  

(42)

Next substitute equation (41) into equation (42) and perform the integration to obtain

\[
v(\phi) = \frac{\phi - \gamma - c(\gamma)}{r + \delta + \lambda - \gamma} + \frac{\lambda}{r + \delta + \lambda - \gamma} \gamma,
\]  

(43)

where \(\gamma\) in this equation is the optimal value of \(\gamma\) when \(\Phi = \phi\). ■

**Proof of Lemma 2:** From Lemma 1, \(H(\gamma^m, \phi, \rho) > 0\) and \(H(\gamma, \phi, \rho) = \phi - \rho - c(\gamma) < 0\) for any \(\Phi \in G\) and \(\rho \geq r + \delta\). Since \(H(\gamma, \phi, \rho)\) is strictly quasi-convex in \(\gamma\), there is a unique \(\gamma \in (\gamma^m, \rho)\) such that \(H(\gamma, \phi, \rho) = -\lambda k\) for any \(-\lambda k \in (H(\rho, \phi, \rho), H(\gamma^m, \phi, \rho))\) or, equivalently, any \(\kappa \in \left(\frac{1}{H(\gamma^m, \phi, \rho), -\frac{1}{\lambda k}(H(\rho, \phi, \rho))\right)\), which contains the non-degenerate interval \([0, \frac{1}{\lambda k}((r + \delta + \lambda) + c(r + \delta + \lambda) - \phi))\).

**Proof of Lemma 3:** Differentiate \(H(\gamma(\phi, k, r + \delta, \lambda), \phi, r + \delta + \lambda) = -\lambda k\) with respect to \(\Phi\) to obtain \(H_\gamma(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho) = \frac{\partial H_\gamma(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho)}{\partial \phi} = 0\). Therefore,

\[
\frac{\partial \gamma(\phi, k, r + \delta, \lambda)}{\partial \kappa} = -\frac{H_\kappa(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho)}{H_\gamma(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho)}.
\]

Use the facts that \(H_\rho(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho) = -1 + c'(\gamma)\) and \(H_\gamma(\gamma, \phi, \rho) = -(\rho - \gamma)c''(\gamma)\) to obtain \(\frac{\partial \gamma(\phi, k, r + \delta, \lambda)}{\partial (r + \delta)} = \frac{1 + c'(\gamma)}{(\rho - \gamma)c''(\gamma)} = 0\) since optimal \(\gamma < \rho\).

Differentiate \(H(\gamma(\phi, k, r + \delta, \lambda), \phi, r + \delta + \lambda) = -\lambda k\) with respect to \(\rho\) to obtain

\[
\frac{\partial \gamma(\phi, k, r + \delta, \lambda)}{\partial \kappa} = -\frac{H_\kappa(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho)}{H_\gamma(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho)} = -\frac{1 + c'(\gamma)}{(\rho - \gamma)c''(\gamma)} < 0\) since optimal \(\gamma < \rho\).

Differentiate \(H(\gamma(\phi, k, r + \delta, \lambda), \phi, r + \delta + \lambda) = -\lambda k\) with respect to \(\lambda\) to obtain

\[
\frac{\partial \gamma(\phi, k, r + \delta, \lambda)}{\partial \kappa} = -\frac{H_\kappa(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho)}{H_\gamma(\gamma(\phi, k, r + \delta, \lambda), \phi, \rho)} = -\frac{1 + c'(\gamma)}{(\rho - \gamma)c''(\gamma)} < 0\) since optimal \(\gamma < \rho\).

**Proof of Corollary 3:** Use the chain rule to obtain \(\frac{\partial c'(\gamma(\phi, k, r + \delta, \lambda))}{\partial x} = \frac{\partial c'(\gamma(\phi, k, r + \delta, \lambda))}{\partial \phi} \frac{\partial \phi}{\partial x} = c''(\gamma(\phi, k, r + \delta, \lambda)) \frac{\partial x}{\partial \phi} = c''(\gamma(\phi, k, r + \delta, \lambda)) \frac{1}{(\rho - \gamma)c''(\gamma)} > 0\) since optimal \(\gamma < \rho\). ■

**Proof of Lemma 4:** To prove property (1), let \(k = 0\) and use the definitions of \(\alpha(\kappa)\) and \(\gamma(\phi, k, r + \delta, \lambda)\) to obtain \(\alpha(0) = 1 + \int_G c'(\gamma(\Phi, 0, r + \delta, \lambda)) dF(\Phi)\). Lemma 2 implies that \(\gamma(\Phi, 0, r + \delta, \lambda) > \gamma^m\) and the convexity of \(c(\gamma)\) implies \(c'(\gamma)\) is strictly increasing so \(\alpha(0) > 1 + \int_G c'(\gamma^m) dF(\Phi) = 1 + c'(\gamma^m) = 0\).
To prove property (2), let $\kappa = 1 + c'(r + \delta)$ and use the definitions of $\alpha(\kappa)$ and $\gamma(\phi, \kappa, r + \delta, \lambda)$ to obtain $\alpha (1 + c'(r + \delta)) = 1 + \int_G c'(\gamma(\Phi, 1 + c'(r + \delta), r + \delta, \lambda))dF(\Phi)$. The definition of $H(\gamma, \phi, \rho)$ implies that $H(\gamma, \phi, r + \delta + \lambda) = H(\gamma, \phi, r + \delta) - \lambda(1 + c'(\gamma))$. In particular, this equation holds for $\gamma = r + \delta$, so that $H(r + \delta, \phi, r + \delta + \lambda) = H(r + \delta, \phi, r + \delta) - \lambda(1 + c'(r + \delta))$. Since $\Phi \in G \equiv \{\Phi : c(\gamma^m) + \gamma^m < \Phi < c(r + \delta) + r + \delta\}$, $H(r + \delta, \phi, r + \delta) = \phi - (r + \delta) - c(r + \delta) < 0$. Therefore, $H(r + \delta, \phi, r + \delta + \lambda) = \phi - (r + \delta) - c(r + \delta) < 0$ for $\gamma < \rho$, implies $\gamma(\phi, 1 + c'(r + \delta), r + \delta + \lambda) < r + \delta$. Therefore, the convexity of $c(\gamma)$ implies $\alpha (1 + c'(r + \delta)) = 1 + \int_G c'(\gamma(\Phi, 1 + c'(r + \delta), r + \delta, \lambda))dF(\Phi) < 1 + \int_G c'(r + \delta)dF(\Phi) = 1 + c'(r + \delta).

To prove property (3), it is helpful to first prove that $\gamma(\phi, \kappa, r + \delta, \lambda) < r + \delta$ for all $\kappa$ in $[0, 1 + c'(r + \delta)]$. The proof of property (2) above includes a proof that $\gamma(\phi, 1 + c'(r + \delta), r + \delta, \lambda) < r + \delta$. Therefore, Statement 2 of Lemma 3, i.e., $\frac{\partial \gamma(\phi, \kappa, r + \delta, \lambda)}{\partial \kappa} = \frac{\lambda}{(r - \gamma(\phi, \kappa, r + \delta, \lambda))} > 0$, implies that $\gamma(\phi, \kappa, r + \delta, \lambda) < r + \delta$ for all $\kappa$ in $[0, 1 + c'(r + \delta)]$. Use the definition of $\alpha(\kappa)$ to obtain $\alpha'(\kappa) = \int_G \frac{\partial \gamma(\phi, \kappa, r + \delta, \lambda)}{\partial \kappa}dF(\Phi)$. Use Statement 2 in Corollary 3, i.e., $\frac{\partial \gamma(\phi, \kappa, r + \delta, \lambda)}{\partial \kappa} = \frac{\lambda}{(r + \delta - \gamma(\phi, \kappa, r + \delta, \lambda))} > 0$, to obtain $\alpha'(\kappa) < \int_G dF(\Phi) = 1$, which completes the proof that $0 < \alpha'(\kappa) < 1$ for $\kappa \in [0, 1 + c'(r + \delta)]$.

Proof of Proposition 4: The function $\alpha(\kappa)$ is continuous over the domain $[0, 1 + c'(r + \delta)]$ and has the three properties listed in Lemma 4. Therefore, there exists a unique positive value of $\kappa < 1 + c'(r + \delta)$ that satisfies $\alpha(\kappa) = \kappa$. For that value of $\kappa$, $\int_G [1 + c'(\gamma(\Phi, \kappa, \rho))]dF(\Phi) = \kappa = \bar{\tau}$. ■

Proof of Corollary 4: Suppose that $\alpha(\kappa^*) < \kappa^*$. Property 3 of Lemma 4 implies that $\alpha(\kappa) < \kappa$ for all $\kappa \in [\kappa^*, 1 + c'(r + \delta)]$. Therefore, the unique value of $\bar{\tau}$ for which $\alpha(\bar{\tau}) = \bar{\tau}$ is less than $\kappa^*$, so $\bar{\tau} - \kappa^* < 0$ when $\alpha(\kappa^*) - \kappa^* < 0$. A similar argument proves that $\bar{\tau} - \kappa^* > 0$ when $\alpha(\kappa^*) - \kappa^* > 0$. ■

Proof of Proposition 5: Let $\kappa^* = \bar{\tau}_1$, which implies that $\bar{\tau}_1 = \int_G [1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))]dF_1(\Phi)$. Since $c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))$ is strictly increasing in $\Phi$ (Statement 1 of Corollary 3), the assumption that $F_2(\Phi)$ strictly first-order stochastically dominates $F_1(\Phi)$ implies that $\kappa^* = \bar{\tau}_1 = \int_G [1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))]dF_1(\Phi) < \int_G [1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))]dF_2(\Phi) = \alpha_2(\kappa^*)$, using the definition in (32). Therefore, Corollary 4 implies that $\bar{\tau}_2 > \kappa^* = \bar{\tau}_1$.

Define $\omega(\gamma^*, \kappa, r + \delta, \lambda)$ to be the value of $\Phi$ for which $\gamma(\Phi, \kappa, r + \delta, \lambda) = \gamma^*$. Since $\gamma(\Phi, \kappa, r + \delta, \lambda)$ is strictly increasing in $\Phi$ and strictly increasing in $\kappa$, it follows that $\omega(\gamma^*, \kappa, r + \delta, \lambda)$ is strictly increasing in $\gamma^*$ and is strictly decreasing in $\kappa$. Note that $\Gamma_1(\gamma^*) = F_1(\omega(\gamma^*, \bar{\tau}_1, r + \delta, \lambda)) > F_2(\omega(\gamma^*, \bar{\tau}_1, r + \delta, \lambda)) \geq F_2(\omega(\gamma^*, \bar{\tau}_2, r + \delta, \lambda)) = \Gamma_2(\gamma^*)$, where the first inequality follows from the assumption that $F_2(\Phi)$ first-order stochastically dominates $F_2(\Phi)$, and the second inequality follows from the facts that $\bar{\tau}_2 > \bar{\tau}_1$, $\omega(\gamma^*, \kappa, r + \delta, \lambda)$ is strictly decreasing in $\kappa$, and $F_2(\Phi)$ is increasing.
Since $F_2(\Phi)$ strictly first-order stochastically dominates $F_1(\Phi)$, the inequality $\Gamma_1(\gamma^*) \geq \Gamma_2(\gamma^*)$, which holds for all $\gamma^*$, holds strictly for some $\gamma^*$. Therefore, $\Gamma_2(\gamma)$ strictly first-order stochastically dominates $\Gamma_1(\gamma)$. ■

**Proof of Proposition 6:** Suppose that initially $\pi = \kappa^*$. Since statement 5 of Corollary 3 states that
\[
\frac{\partial^2 c'((\Phi, \pi, r + \delta, \lambda))}{(\partial \pi)^2} = \frac{1}{(p-\gamma)(\gamma)} > 0, 
\]
cr(\Phi, \kappa^*, r + \delta, \lambda)) is a convex function of $\phi$. Therefore, a mean-preserving spread of $F(\Phi)$ increases the value of $\int_G c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda)) \ dF(\Phi)$, which increases the value of $\alpha(\kappa^*) \equiv 1 + \int_G c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda)) \ dF(\Phi)$, so that $\alpha(\kappa^*) > \kappa^*$. Corollary 4 implies that $\pi > \kappa^*$. ■

**Proof of Corollary 5:** Recall that optimal investment implies that $\pi = \int_G (1 + c'(\gamma(\Phi, \pi, r + \delta, \lambda))) \ dF(\Phi)$. If $c(\gamma)$ is quadratic and convex, it can be written as $c(\gamma) = \frac{1}{2} a\gamma^2 + b\gamma + d$ where $a > 0$. Therefore, $c'(\gamma) = a\gamma + b$ so that $\int_G (1 + c'(\gamma(\Phi, \pi, r + \delta, \lambda))) \ dF(\Phi) = \int_G (1 + a\gamma(\Phi, \pi, r + \delta, \lambda) + b) \ dF(\Phi) = 1 + a\pi + b$, where $\pi = \int_G \gamma \ dF(\Phi)$ is the expected value of $\gamma$. Therefore, $\pi = 1 + a\pi + b$, and since Proposition 6 implies that a mean-preserving spread of $F(\Phi)$ increases $\pi$, it also increases $\pi$. ■

**Proof of Proposition 7:** Since $\alpha(\kappa) \equiv 1 + \int_G c'(\gamma(\Phi, \kappa, r + \delta, \lambda)) \ dF(\Phi)$, we have
\[
\frac{\partial^2 \alpha(\kappa)}{\partial \lambda^2} = \int_G \frac{\partial^2 c'(\gamma(\Phi, \kappa, r + \delta, \lambda))}{\partial \lambda^2} \ dF(\Phi). 
\]
Use statement 4 of Corollary 3 to obtain
\[
\frac{\partial^2 \alpha(\kappa)}{\partial \lambda^2} = \int_G \frac{\kappa - 1 + c'(\gamma(\Phi, \kappa, r + \delta, \lambda))}{1 - \gamma(\Phi, \kappa, r + \delta, \lambda)} \ dF(\Phi). 
\]
Let $\kappa^* = \pi$ and define $\phi^*$ as the unique value of $\Phi$ for which $1 + c'(\gamma(\phi^*, \kappa^*, r + \delta, \lambda)) = \kappa^*$, so $\kappa^* - (1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))) > 0$ if $\Phi < \phi^*$ and $\kappa^* - (1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))) < 0$ if $\Phi > \phi^*$. Since $\gamma(\Phi, \kappa^*, r + \delta, \lambda)$ is increasing in $\Phi$, $\frac{\kappa - 1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))}{1 - \gamma(\Phi, \kappa^*, r + \delta, \lambda)} < \frac{\kappa - 1 + c'(\gamma(\Phi, \phi^*, r + \delta, \lambda))}{1 - \gamma(\Phi, \phi^*, r + \delta, \lambda)}$ if $\Phi < \phi^*$ and $\frac{\kappa - 1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))}{1 - \gamma(\Phi, \kappa^*, r + \delta, \lambda)} > \frac{\kappa - 1 + c'(\gamma(\Phi, \phi^*, r + \delta, \lambda))}{1 - \gamma(\Phi, \phi^*, r + \delta, \lambda)}$ if $\Phi > \phi^*$. Therefore, $\frac{\kappa - 1 + c'(\gamma(\Phi, \kappa^*, r + \delta, \lambda))}{1 - \gamma(\Phi, \kappa^*, r + \delta, \lambda)} \ dF(\Phi) = \frac{\kappa - 1 + c'(\gamma(\Phi, \phi^*, r + \delta, \lambda))}{1 - \gamma(\Phi, \phi^*, r + \delta, \lambda)} \ dF(\Phi)$. Since an increase in $\lambda$ reduces $\alpha(\kappa)$ for all $\kappa$, Corollary 4 implies that an increase in $\lambda$ also reduces $\pi$. ■

**Proof of Corollary 6:** If $c(\gamma)$ is quadratic and convex, it can be written as $c(\gamma) = \frac{1}{2} a\gamma^2 + b\gamma + d$ where $a > 0$. Therefore, $c'(\gamma) = a\gamma + b$. The first-order condition in equation (19) can be written as $1 + a\gamma + b = \pi$, which implies that $1 + a\pi + b = \pi$. Therefore, $\frac{\partial^2 \alpha(\kappa)}{\partial \lambda^2} = \frac{1}{a} \frac{\partial \alpha(\kappa)}{\partial \lambda} < 0$, where the inequality follows from $a > 0$ and Proposition 7, which states that $\frac{\partial \alpha(\kappa)}{\partial \lambda} < 0$. ■

**Proof of Proposition 8:** Proposition 5 states that $\pi_2 > \pi_1$ and statement 2 of Lemma 3 states that $\gamma(\phi, \kappa, r + \delta, \lambda)$ is increasing in $\kappa$. Therefore, $\gamma(\Phi, \pi_2, r + \delta, \lambda) > \gamma(\Phi, \pi_1, r + \delta, \lambda)$, which proves statement 1. Statement 1 implies $\int_G \gamma(\Phi, \pi_2, r + \delta, \lambda) \ dF_2(\Phi) > \int_G \gamma(\Phi, \pi_1, r + \delta, \lambda) \ dF_2(\Phi)$ and statement 1 of Lemma 3 that $\gamma(\phi, \kappa, r + \delta, \lambda)$ is increasing in $\phi$ implies that $\int_G \gamma(\Phi, \pi_1, r + \delta, \lambda) \ dF_2(\Phi) > \int_G \gamma(\Phi, \pi_2, r + \delta, \lambda) \ dF_2(\Phi)$. Therefore, $\int_G \gamma(\Phi, \pi_2, r + \delta, \lambda) \ dF_2(\Phi) > \int_G \gamma(\Phi, \pi_1, r + \delta, \lambda) \ dF_1(\Phi)$, which proves statement 2.

Since $v(\Phi) = 1 + c'(\gamma(\Phi, \pi, r + \delta, \lambda))$, $\frac{dv(\Phi)}{d\Phi} = \frac{d\gamma}{d\Phi}$, $i = 1, 2$. Statement 1 of Corollary 3 is $\frac{\partial \gamma(\phi, \kappa, r + \delta, \lambda)}{\partial \phi} = \frac{1}{p-\gamma}$, so $\frac{dv(\Phi)}{d\Phi} = \frac{1}{p-\gamma}$. Therefore, statement 1 of Proposition 8 implies that $\frac{dv_2(\Phi)}{d\Phi} > \frac{dv_1(\Phi)}{d\Phi}$, which proves statement 3.

Statement 1 of Lemma 3 is that $\gamma(\phi, \kappa, r + \delta, \lambda)$ is increasing in $\phi$, so that $\frac{1}{p-\gamma}$ is in-
creasing in $\Phi$. Therefore, 

$$R^2(\Phi) - \rho(\Phi, \mu_2, r + \delta, \lambda)^2 > R^2(\Phi) - \rho(\Phi, \mu_1, r + \delta, \lambda)^2 = R^2(\Phi) - \rho(\Phi, \mu_2, r + \delta, \lambda)^2.$$ 

Since $\gamma(\phi, \kappa, r + \delta, \lambda)$ is increasing in $\kappa$, 

$$\int_G \frac{dv_2(\Phi)}{d\Phi} dF_2(\Phi) > \int_G \frac{dv_1(\Phi)}{d\Phi} dF_1(\Phi)$$ 

Putting together the inequalities in the two preceding sentences implies 

$$\int_G \frac{dv_2(\Phi)}{d\Phi} dF_2(\Phi) > \int_G \frac{dv_1(\Phi)}{d\Phi} dF_1(\Phi),$$ 

which proves statement 4. ■