Research in Economics (1998) **52**, 3–37

Irreversible investment†

GIUSEPPE BERTOLA‡

Università di Torino and Department of Economics, Istituto Universitario Europeo,
I-50016 San Domenico di Fiesole (FI), Italy

Summary

Investment is often irreversible: once installed, capital has little or no value unless used in production. This paper proposes, solves and characterizes a model of sequential irreversible investment by a firm facing uncertainty in technology, demand and price of capital. The solution can be found in closed form if simple (but not totally unrealistic) functional forms are assumed, and can be given an optimal stopping interpretation. The marginal revenue product of capital that induces additional investment is higher, under irreversibility, than the conventionally measured user cost of capital. In ergodic steady state, however, the former quantity is on average lower than the latter.

J.E.L. Classification: C61, D92.
Keywords: Option values, uncertainty.

1. Introduction and literature review

An essential part of the economic theory of production is the study of the investment process, or of the relationship between the stocks of capital goods at different points in time on the one hand, and the evolution of the firm's business conditions on the other. If it were possible to rent capital services on a smoothly functioning spot market, the user cost of capital as defined by Jorgenson (1963)§ could be

† June, 1988. This is an extensively revised version of a previously circulated paper, with the same title. The material also appears in a doctoral dissertation at the Massachusetts Institute of Technology.

‡ Università di Torino. The participants in numerous seminars offered very helpful comments on this paper; discussions with Olivier Blanchard, Ricardo Caballero and Robert Pindyck were particularly insightful. All errors are the author's. Partial support from the Sloan Foundation is gratefully acknowledged.

§ The neo-classical cost of capital is the opportunity cost of holding one unit of capital for a period. If δ is the economic depreciation rate of capital, r is the discount rate the firm applies to cash flows, P is the price of new and used capital, and $E\{dP\}/P$ is the expected rate of increase of P, the cost of capital equals $(r+\delta)P-E\{dP\}$.

1090–9443/98/010003 + 35 $25.00/0/re970153

used in modelling demand for capital, like the wage rate is used in modelling demand for labour. In reality, however, firms own rather than rent their capital stock, and an explicitly dynamic analysis of the investment process is necessary.

Most economic theories of investment are based on the assumption that variations in capital input are subject to convex adjustment costs—either internal to the firm, due to increasing costs of installing more capital in shorter intervals of time, or external to it and due to decreasing returns in production of capital goods. The adjustment-costs literature often assumes that firms are perfectly competitive and operate under constant returns to scale; the marginal contribution of capital to operating cash flow is then independent of the installed capital stock. Certainty about the dynamic path followed by exogenous processes is often assumed (e.g. in Hayashi, 1982), or costs of adjustment are given a quadratic form (e.g. in Sargent, 1979), yielding certainty equivalence. Abel (1983, 1985) solves for the optimal investment policy under uncertainty, choosing a formal specification close to the one of this paper, and Lucas and Prescott (1971) prove existence of competitive equilibria under uncertainty, giving some characterization results.

Models of investment based on convex adjustment costs have not been very successful empirically (Abel & Blanchard, 1986; Hall, 1987). In fact, the realism of smooth adjustment costs as the source of investment dynamics is doubtful. From a microeconomic point of view, the cost of additions to the capital stock is likely to be linear in investment, even if it does not display increasing returns to scale; disinvestment, on the other hand, is costly, if at all possible: many productive facilities are firm-specific, reconversion to alternative uses of industrial real estate is difficult and markets for most used machinery are thin and discount it heavily. From a macroeconomic point of view, industrial plants are next to worthless unless used in production, as their direct consumption value is clearly very low: Sargent (1979, 1980) studies irreversible investment in the framework of stochastic growth theory, providing existence results as well as a characterization of the accumulation process similar to the one found below.

Two separate strands of literature have focused on the realism and importance of investment irreversibility. Arrow (1968), Nickell (1974) and others have studied irreversible investment decisions in continuous-time dynamic optimization models, assuming that firms hold certain expectations about the continuous cyclical path of exogenous variables.† It can be shown that in this framework

† The same certainty equivalence assumption underlies the (mostly empirical) literature on "putty-clay" models of investment, which assume not only that machine tools have no value unless used in production, but that the labour requirement of existing machines is fixed (see, for example, Ando *et al.*, 1974).

the marginal revenue product of capital equals the cost of capital whenever gross investment is positive. But investment is not necessarily always positive if it is irreversible: it generally ceases before a cyclical peak is reached, and starts again after the cyclical trough. Irreversibility then drives a wedge (negative during booms, and positive during troughs) between the cost of capital and its marginal contribution to profits.

Irreversible investment *under uncertainty* has been studied by financial economists (see McDonald and Siegel (1986) and their references, as well as Ingersoll and Ross (1987) for the case of interest rate uncertainty). Option pricing techniques provide elegant solutions in the case of a single irreversible investment project with uncertain pay-offs: such a project will be adopted only when the expected discounted pay-off from investment *exceeds* the cost by an amount that depends on the level of uncertainty, and can be impressively large for plausible parameter values. Even risk-neutral firms are, in a sense, reluctant to invest when projects are irreversible and the future is uncertain. When the irreversible project is adopted, the option to wait for some of the uncertainty to be resolved is forsaken: options are valuable even to risk-neutral agents, the more so the more uncertain is the future. These results are clearly relevant to the study of the investment process. Bernanke (1983) notes that the level of uncertainty perceived by firms is likely to vary cyclically, and emphasizes that irreversibility effects are important for the understanding of the cyclical behaviour of aggregate investment.

Most of the option valuation models so far available consider only the optimal timing for the adoption of an individual project with given characteristics. Neglecting the availability of many investment projects of different sizes and with different characteristics at different points in time, these models do not provide a proper dynamic investment function. Adoption of an investment project today changes the menu of projects available tomorrow: this has to be explicitly taken into account to clarify the relationship of the option pricing models to the more usual dynamic models based on convex adjustment costs under certainty. Pindyck (1988) applies option pricing techniques to the *marginal* investment decision: in his model, the firm sequentially decides the optimal amount of capacity to be installed, knowing that future demand (and production) are uncertain and follow a geometric Brownian motion stochastic process with known parameters. This paper proposes a model of *irreversible* putty–putty investment under uncertainty: capital has no value unless used in production, but can *ex post* be optimally combined with other factors as more is learned about the path of exogenous variables. Section 2.1 lays out a general model of a firm with these characteristics. If functional forms traditionally used in dynamic investment models are combined with the formalization of continuous-time uncertainty typically used in finance theory, the irreversible investment problem can be solved in closed form. Section 2.2 obtains the solution, and Section 2.3 offers an interpretation based on option pricing techniques. The results bridge the gap between the option-based financial literature on the one hand, and Arrow's and Nickell's results under certainty on the other.

Section 3.1 discusses the characteristics of the investment process implied by the decision rule. Section 3.2 studies the relationship between the neo-classical cost of capital and its marginal contribution to profits, and Sections 3.3 and 3.4 characterize the dynamic behaviour and ergodic distribution of the value of the firm, and of Tobin's average and marginal Q.

2. Optimal sequential investment under uncertainty

2.1. A MODEL OF THE FIRM: PRODUCTION, SALES AND CAPITAL

In a partial-equilibrium model of production, a firm is defined by the production and demand functions and by the form of the stochastic processes it takes as given. Consider a firm endowed with a Cobb–Douglas production function†

$$Q_t=(K_t^{\alpha}(A_tL_t)^{1-\alpha})^{\varphi} \quad 0<\alpha\leqslant 1,\ \varphi>0 \tag{2.1}$$

where Q_t denotes production and sales at time t (inventories are assumed away for simplicity), φ indexes the return to scale in production (constant returns to scale are given by $\varphi=1$), K_t is the capital stock and L_t is input of labour, a perfectly flexible factor that can be rented at the instantaneous price w_t; A_t is an index of technological progress.

The firm is faced by a constant elasticity demand function

$$B_t=D_tQ_t^{\mu-1} \quad 0<\mu\varphi<1 \tag{2.2}$$

† The Cobb–Douglas production function is the workhorse of investment theory. Recent applications related to this paper, in that the uncertainty facing the firm is modelled in continuous time, include: Dietrich and Heckerman (1980), who solve for the one-time choice of capital stock by a competitive firm producing with decreasing returns to scale; Abel (1983, 1985), who studies the investment problem under constant returns, perfect competition and constant-elasticity costs of capital stock adjustment; McDonald and Siegel (1985), who extend Dietrich and Heckerman's model by imposing a fixed cost of production, so that the factory may be shut down; and McDonald and Siegel (1986), who derive the value of an investment opportunity in a Cobb–Douglas plant of fixed size, and the optimal timing of its adoption.

where B_t is the product price at time t, and μ indexes the firm's monopoly power: μ equals the inverse of the markup factor and the firm's monopoly power increases as μ approaches zero, while $\mu=1$ for a competitive firm. The parameter D_t indexes the position of the demand curve (it may be a function of the consumers' income, or of a price index for substitutes, that the firm takes as given in its optimization). If $\mu=1$, $D_t=B_t$ is the market price taken as given by a competitive firm.

Given the capital stock, the operating cash flow function is defined as

$$\Pi(K_t, w_t, D_t, A_t)=\operatorname*{Max}_{L_t} B_tQ_t-w_tL_t$$

subject to equations (2.1), (2.2).

It can be shown by simple algebra that

$$\Pi(K_t, Z_t)=\frac{1}{1+\beta}K_t^{1+\beta}Z_t \tag{2.3}$$

where

$$\beta\equiv\frac{\varphi\mu-1}{1-(1-\alpha)\varphi\mu}, \quad -1<\beta<0 \tag{2.4}$$

$$Z_t=f(w_t, D_t, A_t)\equiv\frac{\alpha\varphi\mu}{1-(1-\alpha)\varphi\mu}$$

$$\left[(\varphi\mu(1-\alpha))^{\frac{(1-\alpha)\varphi\mu}{1-(1-\alpha)\varphi\mu}}-(\varphi\mu(1-\alpha))^{\frac{1}{1-(1-\alpha)\varphi\mu}}\right](1+\beta)$$

$$D_t^{\frac{1}{1-(1-\alpha)\varphi\mu}}\left(\frac{w_t}{A_t}\right)^{\frac{-(1-\alpha)\varphi\mu}{1-(1-\alpha)\varphi\mu}}. \tag{2.5}$$

$\Pi(K_t, Z_t)$ is strictly concave in K as long as $\varphi\mu<1$. The new variable Z_t summarizes at every instant the business conditions for the firm: it is higher the higher is the demand indicator D_t, and the lower is the ratio of the flexible-factor rental cost w_t to its productivity A_t. Cobb–Douglas production and constant-elasticity demand functions yield a log-linear functional form for $\Pi(K_t, Z_t)$.

Uncertainty is introduced in the model by the assumption that $\{w_t\}$, $\{A_t\}$ and $\{D_t\}$ follow geometric Brownian motion stochastic processes:

$$\mathrm{d}w_t=\theta_1w_t\mathrm{d}t+\sigma_1w_t\mathrm{d}W_{1t}$$

$$\mathrm{d}A_t = \theta_2 A_t \mathrm{d}t + \sigma_2 A_t \mathrm{d}W_{2t}$$
$$\mathrm{d}D_t = \theta_3 D_t \mathrm{d}t + \sigma_3 D_t \mathrm{d}W_{3t}$$

where $[\theta_i]$ and $[\sigma_i]$ are vectors of real constants, and $[\mathbf{W}_i]$ is a vector of standard (possibly correlated) Wiener processes. In words, it is assumed that wages, productivity and demand are always expected to grow at some constant mean rate—but the realized growth rates are random, normally distributed and independent over time. The exogenous variables are non-stationary and have no unconditional distribution. Conditional on current values, future values of the exogenous variables are jointly log-normally distributed, with variance proportional to the length of the forecast interval.

The summary indicator of business conditions $\{Z_t\}$, being a constant-elasticity (log–linear) function of geometric Brownian motion processes, follows itself a geometric Brownian motion process,

$$\mathrm{d}Z_t = Z_t \theta_z \mathrm{d}t + Z_t \sigma_z \mathrm{d}W_{zt} \tag{2.6}$$

as is easily shown by an application of Ito's lemma, or by the fact that log-linear functions of log-normal random variables are log-normally distributed. The drift and standard deviation parameters of the $\{Z_t\}$ process are linear combinations of the corresponding parameters of the primitive processes $\{w_t\}$, $\{A_t\}$, $\{D_t\}$, with weights given by the exponents in equation (2.5) and by the correlation between the driving Wiener processes. In empirical work it would be necessary to estimate the importance of each source of uncertainty for the firm; in the remainder of this paper $\{Z_t\}$, the shifter of the reduced-form profit function, is taken as the primitive exogenous variable in the firm's problem.

2.2. OPTIMAL IRREVERSIBLE INVESTMENT DECISIONS

Let capital accumulation be irreversible: at any time t capital can be purchased at unit price P_t, but installed capital has no resale value. Reduction of the capital stock only occurs via depreciation, as the installed capital stock depreciates at the constant exponential rate δ

Let $\{P_t\}$ follow a geometric Brownian motion process

$$\mathrm{d}P_t = P_t \theta_p \mathrm{d}t + P_t \sigma_p \mathrm{d}W_{pt} \tag{2.7}$$

and let $\rho = (\mathrm{d}W_{pt}\mathrm{d}W_{zt})/\mathrm{d}t$ be the correlation between the increments of the price of capital and its marginal revenue product for a given capital stock.

The firm's managers are risk neutral and attempt to maximize the expected value of cash flows, discounted at a given, constant required rate of return r: let the resulting value function be

$$V^*(K_t, Z_t, P_t) = \underset{\{X_\tau\}}{\mathrm{Max}}\, E_t\left\{\int_t^\infty e^{-r(\tau-t)}[\Pi(K_\tau, Z_\tau)\mathrm{d}\tau - P_\tau \mathrm{d}X_\tau]\right\} \quad (2.8)$$

subject to

$$\mathrm{d}K_\tau = -\delta K_\tau \mathrm{d}\tau + \mathrm{d}X_\tau, \quad \text{all } \tau \quad (2.8a)$$

$$\mathrm{d}X_\tau \geqslant 0, \quad \text{all } \tau \quad (2.8b)$$

where $\{X_t\}$ is the gross investment process. In the absence of convex installation costs, the rate of growth of capital is unbounded and the second integral in equation (2.8) is defined in the Stieltjes sense, with $\{X_t\}$ the integrating function.

It is necessary to find the contingent investment rule $\{X_t\}$, or—which is the same—the stochastic process $\{K_t\}$, that achieves the maximum in equation (2.8). The expectation E_t in equation (2.8) is taken over the joint distribution of the $\{K_t\}$, $\{P_t\}$ and $\{Z_t\}$ processes, conditional on the information available at time t, taking into account that investment decisions will be taken optimally (subject to the irreversibility constraint) in the future. At any time t, the irreversibility constraint imposes that $\mathrm{d}X_t \geqslant 0$, or, which is the same, $K_{t+} \geqslant K_{t-}$. The following Proposition provides necessary conditions of optimality:

PROPOSITION 1: *if the firm follows the optimal irreversible investment policy, the following must almost surely hold at all* t:

$$E_t\left\{\int_t^\infty e^{-(r+\delta)(\tau-t)} \frac{\partial \Pi(K_\tau, Z_\tau)}{\partial K_\tau} \mathrm{d}\tau\right\} = P_t \quad \text{if } \mathrm{d}X_t > 0 \quad (2.9)$$

$$E_t\left\{\int_t^\infty e^{-(r+\delta)(\tau-t)} \frac{\partial \Pi(K_\tau, Z_\tau)}{\partial K_\tau} \mathrm{d}\tau\right\} \leqslant P_t \quad \text{if } \mathrm{d}X_t = 0. \quad (2.10)$$

PROOF: let $\{K_\tau;\ 0 \leqslant \tau \leqslant \infty\}$ be the capital stock stochastic process corresponding to the optimal feedback rule. Denote Ω the sample space from which $\{Z_t\}$ and $\{P_t\}$ paths are drawn, $p(\cdot)$ the probability measure induced on Ω by the Wiener processes $\{W_{zt}\}$ and $\{W_{pt}\}$, and $\{\mathrm{F}_t\}$ the family of σ-fields on Ω produced by observation of $\{Z_\tau;\ \tau \leqslant t)$ and $\{P_\tau;\ \tau \leqslant t\}$.

Let A be a subset of Ω, with $p(A) > 0$, such that if $\omega \in A$ neither equation (2.9) nor equation (2.10) are satisfied at $t = T < \infty$. Perturb the original investment policy $\{dX_t\}$ by an amount Δ ($\Delta > -dX_T$, $\Delta \neq 0$) at time T if A occurs, without otherwise modifying the feedback rule: $dX_T = dX_T + \Delta$ if A occurs, and $d\tilde{x}_\tau = dX_t$ almost surely for $\tau \neq T$. $A \in F_T$ (all expressions in equations (2.9) and (2.10) are observable at $t = T$), so the modified policy is still adapted to F_t. Denote $\{\tilde{K}_\tau;\ 0 \leqslant \tau \leqslant \infty\}$ the capital stock paths for the perturbed policy; integrating equation (2.8a) we obtain $\tilde{K}_\tau = K_\tau$ almost surely for $0 \leqslant \tau \leqslant T$, and for $\tau \geqslant T$ $\tilde{K}_\tau = K_\tau + \Delta e^{-\delta(\tau - t)}$ if A occurs, $\tilde{K}_\tau = K_\tau$ otherwise.

Proceed to evaluate at $t = T$ the difference between the value function corresponding to the perturbed policy and the value function corresponding to the unperturbed policy:

$$
\begin{aligned}
&\tilde{V}(K_t, Z_t, P_t) - V^*(K_t, Z_t, P_t) \\
&= E_t\left\{\int_t^\infty e^{-r(\tau - t)}\,(\Pi(\tilde{K}_\tau, Z_\tau) - \Pi(K_\tau, Z_\tau))\,d\tau\right\} - P_t\Delta \\
&= E_t\left\{\int_t^\infty e^{-r(\tau - t)}\,(\Pi(K_\tau + \Delta e^{-\delta(\tau - t)}, Z_\tau) - \Pi(K_\tau, Z_\tau))\,d\tau\right\} - P_t\Delta \\
&> \left[E_t\left\{\int_t^\infty e^{-r(\tau - t)}\,\frac{\partial \Pi(K_\tau, Z_\tau)}{\partial K_\tau}\,e^{-\delta(\tau - t)}\,d\tau\right\} - P_t\right]\Delta
\end{aligned}
\tag{2.11}
$$

where the inequality follows, for $\Delta \neq 0$, from strict concavity of $\Pi(K, Z)$ in its first argument. For any failure of equations (2.9) and (2.10), it is possible to choose a Δ to obtain a positive right-hand side for equation (2.11), contradicting the assumed optimality of $\{K_t\}$: if

$$
E_t\left\{\int_t^\infty e^{-(r+\delta)(\tau - t)}\,\frac{\partial \Pi(K_\tau, Z_\tau)}{\partial K_\tau}\,d\tau\right\} > P_t,
$$

choose $\Delta > 0$; if

$$
E_t\left\{\int_t^\infty e^{-(r+\delta)(\tau - t)}\,\frac{\partial \Pi(K_\tau, Z_\tau)}{\partial K_\tau}\,d\tau\right\} < P_t
$$

but $dX_t > 0$, choose $-dX_t < \Delta < 0$.

As long as $p(A) > 0$, one can show using iterated expectations that a Δ-modification of the candidate policy strictly increases the value of the firm not only at the time T when equations (2.9) and (2.10) fail if A occurs, but at all $t < T$ as well.

Any failure of equations (2.9) and (2.10) that occurs with positive probability implies that the firm is not following the optimal feedback rule in its control policy, in that a feasible feedback policy would yield a strictly larger value function.

QED

Conditions (2.9) and (2.10) have simple interpretations as Kuhn–Tucker complementary slackness conditions for the constrained value maximization problem. The left-hand side of equations (2.9) and (2.10) is the shadow price of capital, or the derivative of V^* with respect to K_{t+}. By the envelope theorem, when taking the derivative all controls but dX_t are taken as given (in probability distribution): the currently marginal unit of installed capital is viewed, allowing for depreciation, as the marginal unit throughout the infinite future. The firm knows that equations (2.9) and (2.10) will be satisfied at all future times, and this defines the (yet to be found) probability distribution of future capital stocks, which is used in taking the expectation in equations (2.9) and (2.10).

The functional forms assumed above make it possible to solve the irreversible investment problem. The following propositions provide sufficient conditions for existence of a solution to equation (2.8), and a proof of uniqueness of the optimal investment rule.

PROPOSITION 2: *provided that*

$$r > \left(\theta_p\left(1+\frac{1}{\beta}\right)-\theta_z\frac{1}{\beta}+\frac{\sigma^2}{2}\frac{1+\beta}{\beta}\frac{1}{\beta}\right)$$

where $\sigma^2 \equiv \sigma_z^2+\sigma_p^2-2\rho\sigma_z\sigma_p$ *is the variance of the rate of increase of the process* $\{Z_t/P_t\}$, *the irreversible investment problem has a solution—i.e. the value function* V^* *defined in equation (2.8) is bounded.*

PROOF: if the firm could at any time buy or sell capital at the same price P_t, the risk-neutral manager's problem would be the same as if there were a rental market for capital, with instantaneous rental rate $(r-\theta_p+\delta)P_t$; then the capital stock would always satisfy the first-order condition $K_\tau^\beta Z_\tau=(r-\theta_p+\delta)P_\tau$, implying that

$$K_\tau=\left[(r-\theta_p+\delta)\frac{P_\tau}{Z}\right]^{1/\beta}$$

and that

$$\Pi(K_\tau, Z_\tau) = \frac{1}{1+\beta}\left[(r-\theta_p+\delta)\frac{P_\tau}{Z}\right]^{(1+\beta)/\beta} Z_\tau.$$

Consider the present discounted value of operating profits under such conditions of *reversible* investment. By Fubini's theorem, the order of the expectation and integration operators can be reversed as long as both operations are well defined, to obtain

$$E_t\left\{\int_t^\infty e^{-r(\tau-t)}\,\Pi(K_\tau, Z_\tau)\mathrm{d}\tau\right\}$$

$$= \int_t^\infty e^{-r(\tau-t)}\frac{1}{1+\beta}(r-\theta_p+\delta)^{(1+\beta)/\beta} E_t\left\{\left[\frac{P_\tau}{Z_\tau}\right]^{(1+\beta)/\beta} Z_\tau\right\}\mathrm{d}\tau. \qquad (2.12)$$

Conditional on the information available at $t<\tau$, $[P_\tau/Z_\tau]^{(1+\beta)/\beta} Z_\tau$ is a constant-elasticity combination of log-normal random variables, and is itself log-normally distributed. By the properties of the log-normal distribution, or by a simple application of Ito's lemma, one finds

$$E_t\left\{\left[\frac{P_\tau}{Z_\tau}\right]^{(1+\beta)/\beta} Z_\tau\right\}$$

$$= \left[\frac{P_t}{Z_t}\right]^{(1+\beta)/\beta} Z_t \mathrm{e}^{\left(\theta_p\left(1+\frac{1}{\beta}\right)-\theta_z\frac{1}{\beta}+\frac{\sigma^2}{2}\frac{1+\beta}{\beta}\frac{1}{\beta}\right)(\tau-t)}$$

and it is easy to see that the integral in equation (2.12) converges if and only if

$$r > \left(\theta_p\left(1+\frac{1}{\beta}\right)-\theta_z\frac{1}{\beta}+\frac{\sigma^2}{2}\frac{1+\beta}{\beta}\frac{1}{\beta}\right). \qquad (2.13)$$

Imposing the irreversibility constraint and subtracting investment expenditures can only decrease the value of the firm, which therefore is bounded.†

QED

† Equation (2.13) turns out to be *necessary*, as well as sufficient, for existence of the value function. See the expression for the value of the firm below.

PROPOSITION 3: *if an investment rule can be found to satisfy equations (2.9) and (2.10) at all times, that is the unique solution of the optimization problem (equation (2.8)).*

PROOF:† by standard Kuhn–Tucker theory, uniqueness of the optimal investment policy follows from concavity of the maximand, V^*, at times when investment is positive. Concavity of $V^*(K_t, Z_t, P_t)$ follows from concavity in K_t of the instantaneous operating profit function $\Pi(K_t, Z_t)$. Since K_{t+} increases linearly with $\mathrm{d}X_t$, the instantaneous cash flow function $\Pi(K_t, Z_t)-P_t\mathrm{d}X_t$ is concave in $\mathrm{d}X_t$. The value function $V^*(K_t, Z_t, P_t)$ is the integral of instantaneous cash flows sample paths, over states of nature and time, with positive measure (the joint probability measure of density of Z_τ, P_τ, K_τ and the discount factor). To prove concavity of $V^*(K, Z, P)$ in its first argument it is sufficient to show that K_t and $-\mathrm{d}X_\tau$ are non-decreasing in K_t, $t<\tau$.‡ But investment irreversibility implies that more investment at time t, all else being equal, will never result in a lower K_τ or in a higher $\mathrm{d}X_\tau$ ($\tau>t$), and concavity of $V^*(K_t, Z_t)$ is therefore proved.

Having established existence and uniqueness of the solution, the next Proposition finds the investment rule in closed form.

QED

PROPOSITION 4: *under the assumptions given above, the optimal investment policy is the one that obtains*

$$\frac{\partial\Pi(K_t, Z_t)}{\partial K_t} \leqslant P_t c \text{ always} \tag{2.14a}$$

$$\frac{\partial\Pi(K_t, Z_t)}{\partial K_t} = P_t c \text{ if } \mathrm{d}X_t>0 \tag{2.14b}$$

where the constant c *is defined as*

$$c \equiv \frac{A}{A-1}(r+\delta+\delta\beta-\theta_z)$$

$$A \equiv \frac{-\left(\theta_z-\delta\beta-\theta_p-\dfrac{\sigma^2}{2}\right)+\sqrt{\left(\theta_z-\delta\beta-\theta_p-\dfrac{\sigma^2}{2}\right)^2+2(r+\delta-\theta_p)\sigma^2}}{\sigma^2}$$

$$\sigma^2 \equiv \sigma_z^2+\sigma_p^2-2\rho\sigma_z\sigma_p.$$

† Ricardo Caballero suggested this line of proof.

‡ A non-decreasing function of a concave function is weakly concave; the *current* cash flow is strictly concave; and sums (or integrals) of concave functions are strictly concave if one of the elements is strictly concave.

PROOF: Appendix 1 derives the shadow value of capital under the assumption that equations (2.14a) and (2.14b) describe the investment policy:

$$E_t\left\{\int_t^\infty \mathrm{e}^{-(r+\delta)(\tau-t)}K_\tau^\beta Z_\tau \mathrm{d}\tau\right\}$$

$$=\frac{K^\beta Z_t-\dfrac{1}{A}\left(\dfrac{K_t^\beta Z_t}{cP_t}\right)^{A-1}K_t^\beta Z_t}{r+\delta+\delta\beta-\theta_z}. \tag{2.15}$$

Considering equations (2.14a) and (2.14b), it can be immediately verified that this expression satisfies equations (2.9) and (2.10).

QED

Note, for reference below, that $-\beta A$ is the positive solution to the quadratic equation

$$\frac{\sigma^2}{2\beta^2}X^2-\left(\theta_z\frac{1}{\beta}+\delta-\theta_p\frac{1}{\beta}-\frac{\sigma^2}{2}\frac{1}{\beta}\right)X-(r+\delta-\theta_p)=0$$

and that, as long as the condition in equation (2.13) holds, the positive root of this equation is larger than unity: it follows that $A>-1/\beta$.

As shown in Appendix 1, the optimal irreversible investment policy controls the stochastic process followed by the marginal revenue product of capital, not allowing it to become larger than a constant proportion c of the purchase price of capital. Section 3 discusses in detail the stochastic characteristics of the investment process; the remainder of this section proposes an economic interpretation of the investment rule, and an alternative approach based on optimal stopping arguments.

Figure A plots c as a function of σ for several values of β; values for the other parameters are given in the figure. The higher the variability of its environment, the more reluctant the firm is to invest; this is summarized by σ (note that σ is a function of variances and covariances of the processes for demand, wage, productivity and capital price, and is monotonically increasing in the variability of each of these primitive processes).

Such reluctance to invest is not surprising: higher variability worsens the "worst case" scenario, in which the firm regrets the irreversible investment decision. Higher variance does not symmetrically improve the "best case" scenario. Whenever demand increases, or the wage decreases, or the price of capital falls,

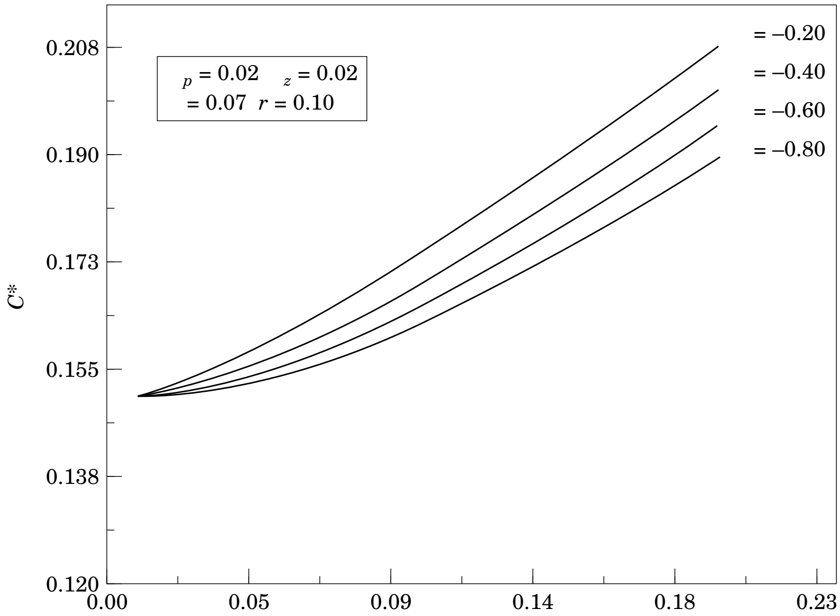


FIGURE A. *C**.

the firm can easily increase the capital stock. The irreversibility constraint only binds in the case of adverse realizations of uncertainty, and from the point of view of the investing firm effectively truncates the (log-normal) probability distribution of future states of nature. Bernanke (1983) refers to this insight as the "' bad news principle of irreversible investment' . . . of all possible future outcomes, only the unfavourable ones have a bearing on the current propensity to undertake a given project . . .".

When $\sigma \to 0$ it appears from the figure that $c \to (r - \theta_p + \delta)$, the Jorgenson (1963) rental cost of capital. This can be verified† to be true as long as

$$\theta_z - \delta\beta - \theta_p > 0 \tag{2.16}$$

implying that, under certainty, the irreversibility constraint is *never* binding (the firm's desired dotation of capital steadily increases).

If equation (2.16) does not hold, in the absence of uncertainty the firm never wants to increase the capital stock—except when it is set up and a capital stock is chosen taking the irreversibility

† If $\sigma^2 = 0$ then A is the solution to $(\theta_z - \beta\delta - \theta_p)A - (r + \delta - \theta_p) = 0$. Insertion of $A = (r + \delta - \theta_p)/(\theta_z - \beta\delta - \theta_p)$ in equation (2.14) yields the result.

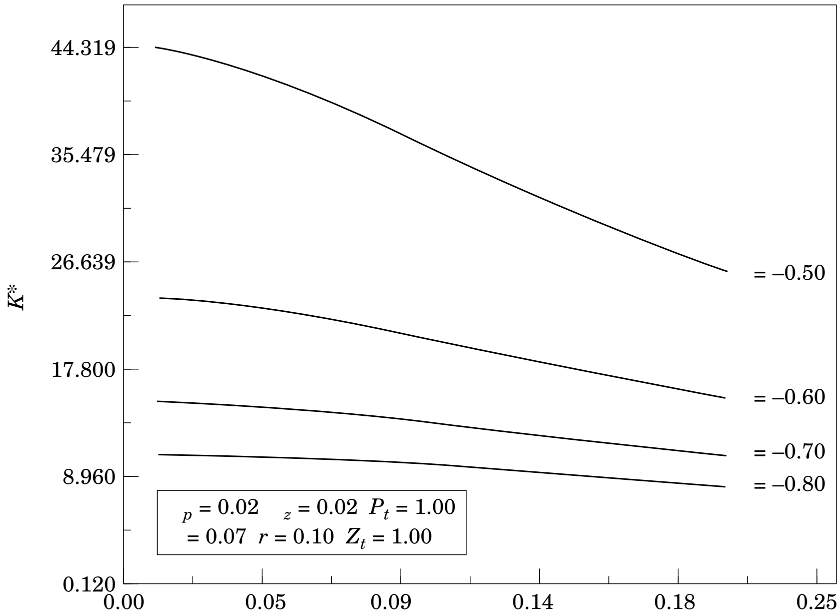


FIGURE B. K^*.

constraint into account: it is possible to verify that equations (2.14a) and (2.14b) yield the appropriate limit for this case as well.

The value of the marginal productivity of capital that triggers investment is higher, for a given σ, if β is lower in absolute value, that is, if the instantaneous operating profit function is less concave in K: for a given $\delta>0$, a large $|\beta|$ implies that the marginal profitability of capital increases more rapidly when gross investment is zero and the installed capital stock depreciates. It is therefore less likely that the firm will be stuck with an excessive stock of capital *ex post*.

The marginal condition in equation (2.14b) can be inverted to find an expression for the firm's desired capital stock as a function of the current value of Z and P:

$$K^*(z_t, P_t) = \left(c\frac{P_t}{Z_t}\right)^{1/\beta} = \left(\frac{A}{A-1}(r+\delta+\delta\beta-\theta_z)\frac{P_t}{Z_t}\right)^{1/\beta}.$$

Since $\beta<0$, the desired capital stock is higher when Z is higher (the lower w_t, the stronger the demand), and P_t is lower. Figure B plots the desired capital stock for *given* Z and P, as a function of σ, for several values of β: higher uncertainty implies a lower desired capital stock, because the firm knows that the "worst case" is very likely to occur and hedges against possible decreases in Z or P. Desired capital decreases in $|\beta|$ for given P, Z and σ: a firm with high $|\beta|$ has more monopoly power and/or more strongly decreasing returns to scale, hence supplies less and uses less capital to maximize its profits.

The optimal irreversible investment rule is then a simple one under the assumptions made above: if the currently installed capital stock K_t is smaller than $K^*(Z_t, P_t)$, the firm immediately invests so that $K_t = K^*$; if K is larger than K^*, the firm does not invest, and lets the capital stock be reduced by depreciation. Of course, since downward fluctuations of Z are possible, the firm will sometimes regret the investment decision; similarly, the firm will regret having invested when P decreases fast: investments made in the past could have been profitably delayed.

2.3. OPTIMAL STOPPING APPROACH

The dynamic program defined in equation (2.8) can be decomposed in a sequence of optimal stopping problems: rather than deciding *how much* to invest at any given time, the firm can decide *when* infinitesimal particles of capital should be installed. This approach to the problem has been used by Pindyck (1988).

At any time, installation of an additional infinitesimal unit of capital will produce a stream of marginal pay-offs, regardless of future investment decisions. Given business conditions Z_t, the Kth unit of capital contributes $\partial\Pi(K, Z_t)/\partial K = K^\beta Z_t$ to the firm's operating cash flow. Addition of one marginal unit at time t increases the capital stock at $\tau > t$ by $e^{-\delta(\tau-t)}$ units, and (taking depreciation of the current capital stock into account) the expected marginal discounted cash flow produced by installation of the Kth unit at time t equals

$$
\begin{aligned}
v(K, Z_t) &= E_t\left\{\int_t^\infty e^{-r(\tau-t)} e^{-\delta(\tau-t)} (K e^{-\delta(\tau-t)})^\beta Z_\tau \, d\tau\right\} \\
&= \int_t^\infty e^{-(r+\delta)(\tau-t)} (K e^{-\delta(\tau-t)})^\beta Z_t e^{\theta_z(\tau-t)} \, d\tau \\
&= \frac{K^\beta Z_t}{r+\delta-(\theta_z-\delta\beta)}. \qquad (2.17)
\end{aligned}
$$

The decision to install a single, marginal unit of capital defines an optimal stopping problem of the form considered by McDonald and Siegel (1986). At every time t, the firm holds the right to pay P_t, acquire the currently marginal unit of K and receive the stream of marginal operating cash flow whose discounted expectation is given in equation (2.17)—*regardless* of future investment decisions. If the right to purchase is not exercised for a unit of capital, a "dividend" (equal to the marginal revenue product of capital) is given up: uninstalled units do not produce. On the other hand, if the right to invest is exercised the firm gives up the option to wait and to learn about the evolution of $\{Z_t\}$ and $\{P_t\}$: often installation will turn out *ex post* to have been a bad idea. This defines an optimal risk-taking problem. In deciding when to install additional units of capital, the firm has to trade-off the forgone dividends against the risk of acting too soon.

Appendix 2 uses standard optimal stopping arguments to prove the following:

PROPOSITION 5: *the value of the opportunity to install the currently marginal* K*th unit of capital is*

$$F(K, Z_t, P_t) = \frac{1}{A-1}\left(\frac{K^{\beta} Z_t}{cP_t}\right)^{A-1} \frac{K^{\beta} Z_t}{c}, \quad K \geqslant K_t \tag{2.18}$$

and the right to invest is optimally exercised when

$$K_t^{\beta} Z_t = P_t c \tag{2.19}$$

where c is the same constant defined in Proposition 4; equation (2.19) implies that

$$v(K_t, Z_t) = P_t + F(K_t, Z_t, P_t) \tag{2.20}$$

at times of positive investment.

PROOF: see Appendix 2.

According to equation (2.20), installation of a marginal unit of capital is optimal when $v(K_t, Z_t)$ (the expected discounted value of marginal profits from *that* unit) is so large as to cover not only the out-of-pocket cost of installation, P_t, but also the opportunity cost of immediate installation, or "value of waiting", $F(K_t, Z_t, P_t)$.

It has been noted by financial economists (McDonald & Siegel, 1986, Ingersoll & Ross, 1987) that standard present-value decision rules fail to take the option value of waiting into account, and therefore induce premature exercise of investment opportunities, neglecting the fact that every investment project competes for funds with itself delayed. In the fully dynamic model of this paper, this insight is found to be equivalent with the more traditional approach based on the shadow value of capital, computed (in equations (2.9), (2.10) and (2.15) above) as the expected discounted contribution to cash flow of the marginal unit of capital in the future—rather than of the unit that is currently marginal, but will be sub-marginal as soon as more investment will be undertaken.

3. Characterization of irreversible capital accumulation

3.1. THE INVESTMENT PROCESS

The investment rule in equations (2.14a) and (2.14b) maintains the ratio of the marginal revenue product of capital ($K_t^{\beta} Z_t$) to the price of capital (P_t) in the $(0, c)$ region. For notational convenience, let $\eta_t \equiv K_t^{\beta} Z_t$ denote the marginal revenue product of capital process.

Given the assumptions of Section 2.1, in the absence of investment η_t / P_t follows a geometric Brownian motion process—with drift determined by the expected rate of growth of the business conditions $\{Z_t\}$, by the expected rate of increase of P_t and by the rate of capital depreciation δ (weighed by the concavity of the marginal revenue product function, β). Dynamic optimization models in whose solution an Ito process is constrained never to leave a region have been extensively studied in the Operations Research literature (see Benes *et al.*, 1980; Harrison, 1985; Karatzas & Shreve, 1988 and references therein).

The sample paths of $\{\eta_t / P_t\}$ have *infinite variation*, and, though they are continuous almost everywhere, they are differentiable almost nowhere: Brownian motion sample paths move both up *and* down in any interval of time, no matter how small. In the model of the previous sections, the firm has to be very fast in exercising control to maintain $\{\eta_t / P_t\}$ below the control barrier c. The control process X_t increases, or gross investment is positive, *only* when $\eta_t / P_t = c$. Because of infinite variation, this only happens at distinct moments in time: control is never exercised throughout the length of any non-empty interval $[\tau_1, \tau_2]$. Investment occurs in spurts, whenever the price of capital and business conditions are sufficiently favourable, in an extreme continuous time version of the investment accelerator.

The stochastic process followed by the control process $\{X_t\}$ is *singular*: though continuous (it is a continuous transformation of Brownian motion), it only increases on a time set that has total measure zero, being a collection of distinct points. At its points of increase or decrease, X_t moves infinitely fast, though it never jumps: the rate of control is infinite, making the classic Hamiltonian analysis inapplicable. The engineering literature solves abstract cost-minimization problems by Hamilton, Jacobi, Bellman methods, using a limit argument to allow the rate of control to become infinite (see, for example, Chow *et al.*, 1985); the equivalence between a singular control problem and a sequence of optimal stopping problems has been noted in this literature (Karatzas & Shreve, 1984, 1985).

3.2. IRREVERSIBLE CAPITAL ACCUMULATION IN ERGODIC STEADY STATE

Exogenous processes $\{Z_t\}$ and $\{P_t\}$ are assumed above to be non-stationary, and do not possess a steady-state distribution. The installed capital stock is, however, cointegrated (in logarithms) with the processes taken as given by the firm, in the sense that there are functions of the exogenous processes and of the installed capital stock which do possess a steady-state distribution.

The ratio of the marginal profitability of currently installed capital to the current purchase price of capital plays a fundamental role in the investment rule. It is convenient to define

$$\xi_t \equiv K_t^\beta Z_t / P_t = \eta_t / P_t. \tag{3.1}$$

Ito's lemma can be used to derive the stochastic differential of ξ_t when the firm is not investing:

$$\begin{aligned} \mathrm{d}\xi_t &= \mathrm{d}[K_t^\beta Z_t P_t^{-1}] \\ &= \beta K_t^{\beta-1} Z_t P_t^{-1}(-\delta K_t \mathrm{d}t) + K_t^\beta P_t^{-1}(\mathrm{d}Z_t) - K_t^\beta Z_t P_t^{-2} \\ &\quad \times (\mathrm{d}P_t) + 2\frac{1}{2} K_t^\beta Z_t P_t^{-3}(\mathrm{d}P_t)^2 - K_t^\beta P_t^{-2}(\mathrm{d}Z_t \mathrm{d}P_t) \\ &= \xi_t(-\delta\beta + \theta_z - \theta_p + \sigma_p^2 - \rho\sigma_z\sigma_p)\mathrm{d}t + \xi_t(\sigma_z \mathrm{d}W_{zt} - \sigma_p \mathrm{d}W_{pt}). \end{aligned}$$

The drift and standard deviation parameters of $\{\xi_t\}$ are

$$m \equiv -\delta\beta + \theta_z - \theta_p + \sigma_p^2 - \rho\sigma_z\sigma_p), \quad \sigma \equiv \sqrt{\sigma_z^2 + \sigma_p^2 - 2\rho\sigma_z\sigma_p}.$$

The investment policy (equations (2.14a) and (2.14b)) imposes on $\{\xi_t\}$ an upper control barrier at c. $\{\xi_t\}$ then follows a regulated geometric Brownian motion with a reflecting barrier at c. Then, $-\ln(\xi_t) + \ln(c)$ is a *linear* Brownian motion process with a control barrier at zero and (by Ito's lemma) drift $-(m - \sigma^2/2)$ and standard deviation σ. The ergodic ditribution for such a process is known to exist as long as the drift is negative, i.e. (from the definitions above for m and σ) as long as

$$-\delta\beta + \theta_z - (\sigma_z^2/2) > \theta_p - (\sigma_p^2/2). \tag{3.2}$$

This requires both that $\{\xi_t\}$ have a tendency to drift (upwards) towards the investment point, *and* that there not be too much noise in the model: σ_z^2 and σ_p^2 should be small compared to the drift parameters δ, θ_z and θ_p. If equation (3.2) is not satisfied, the density of $\{-\ln(\xi_t)+\ln(c)\}$ goes to zero everywhere in $[0, \infty)$, implying that the density of ξ degenerates to a spike arbitrarily close to zero: ξ converges to zero in probability for all initial conditions. In the certainty case, if $\sigma_z^2=\sigma_p^2=\sigma^2=0$ and equation (3.2) is not satisfied, the firm would never undertake a dynamic investment strategy, but would limit itself to a once-and-for-all acquisition of capital when it is set up (compare equation (3.2) with equation (2.16)). The ratio of capital's marginal profitability to its purchase price would then certainly converge to zero as $t\to\infty$. In the presence of uncertainty, the firm would invest not only at the beginning of time but also at other points in time, when business conditions and/or the price of capital are favourable enough to obtain $\xi_t=c$ even though equation (3.2) fails to hold. In the limit the probability of observing $\xi>0$ goes to zero all the same, because good business conditions and/or low price of capital are very unlikely if equation (3.2) is not true.

If equation (3.2) is satisfied, the ergodic distribution of the $\{\xi_t\}$ process is well defined and is exponential (see Cox & Miller, 1965, page 225):

$$\text{Prob}(-\ln(\xi)+\ln(c)\leqslant x)=1-1(x\geqslant 0)\mathrm{e}^{\left(\frac{2(m-s^2/2)}{s^2}\right)x}$$

where $1(\cdot)$ is the indicator function.

It is then a simple matter to invert the monotonic function $f(\xi)=-\ln(\xi_t)+\ln(c)$ and find the ergodic cumulative distribution function of ξ:

$$\text{Prob}(\xi\leqslant x)=\left(\frac{x}{c}\right)^{2(m/\sigma^2)-1}1(0\leqslant x\leqslant c). \tag{3.3}$$

The steady-state density of ξ is plotted in Figure C for different levels of uncertainty. If there were no uncertainty the density would degenerate to a spike located at $(r-\theta_p+\delta)$, the neo-classical rental cost of capital. As uncertainty becomes more important, c increases (see Figure A), the upper limit of the distribution shifts to the right and more probability density is located at low values of ξ.

Simple integration shows that, in the ergodic steady state, the mean of ξ is

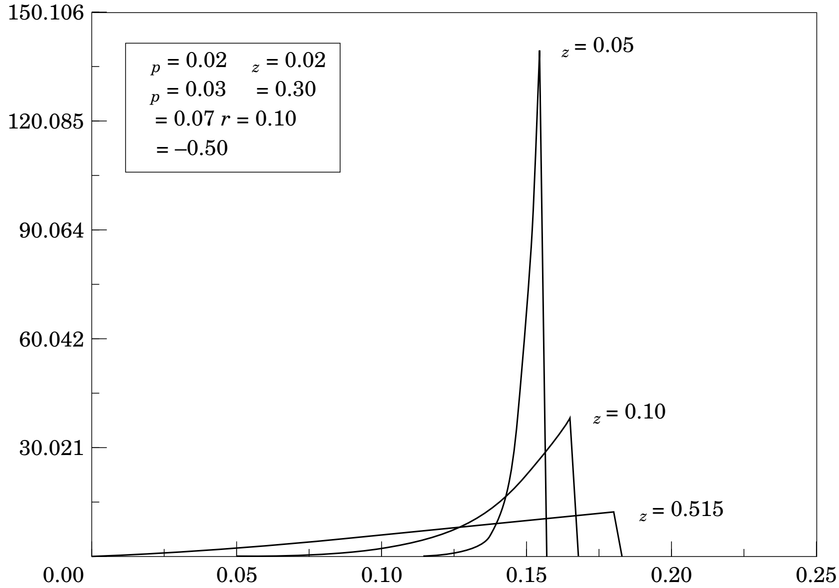


FIGURE C. Ergodic density of ξ.

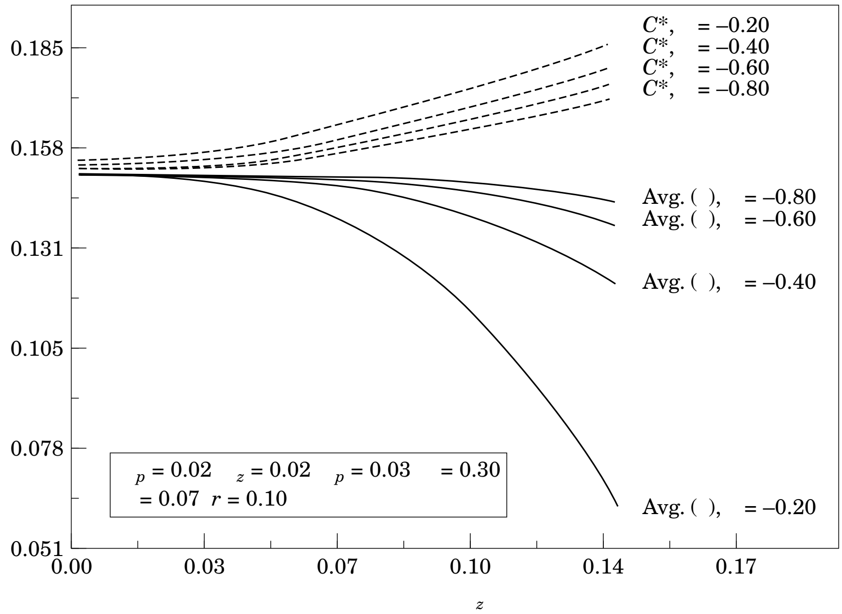
$$\bar{\xi} = \left(\frac{m - \sigma^2/2}{m}\right) c = \frac{-\delta\beta + \theta_z - (\sigma_z^2/2) - \theta_p + (\sigma_p^2/2)}{-\delta\beta + \theta_z - \theta_p + \sigma_p^2 - \rho\sigma_z\sigma_p} c. \tag{3.4}$$

Of course, $0 < \bar{\xi} < c$ (as long as equation (3.2) is satisfied and the expectation is well defined).

If equation (3.2) is satisfied, this expression can be shown to be equal to $r + \delta - \theta_p$ when there is no uncertainty ($\sigma_z^2 = \sigma_p^2 = 0$), and to be strictly *less* than that when there is uncertainty and the irreversibility constraint is sometimes binding.

Figures D(i) and D(ii) plot the ergodic mean of ξ as a function of σ_z and σ_p, for several values of the concavity parameter β. It is apparent from the figures that when the parameters are such as to make the firm more reluctant to invest *ex ante*, they are also such that *ex post* the marginal revenue product of capital is *lower* compared to the reversible investment case.

Uncertainty, while making the firm more reluctant to undertake irreversible investment *ex ante*, also makes adverse realizations of business conditions or decreases in the price of capital so likely that, *ex post*, irreversible capital accumulation results on average in *higher* capital intensity of production. Hall (1987) tests empirically the equality of marginal revenue product and rental cost of capital in the long run, under the assumption of constant returns



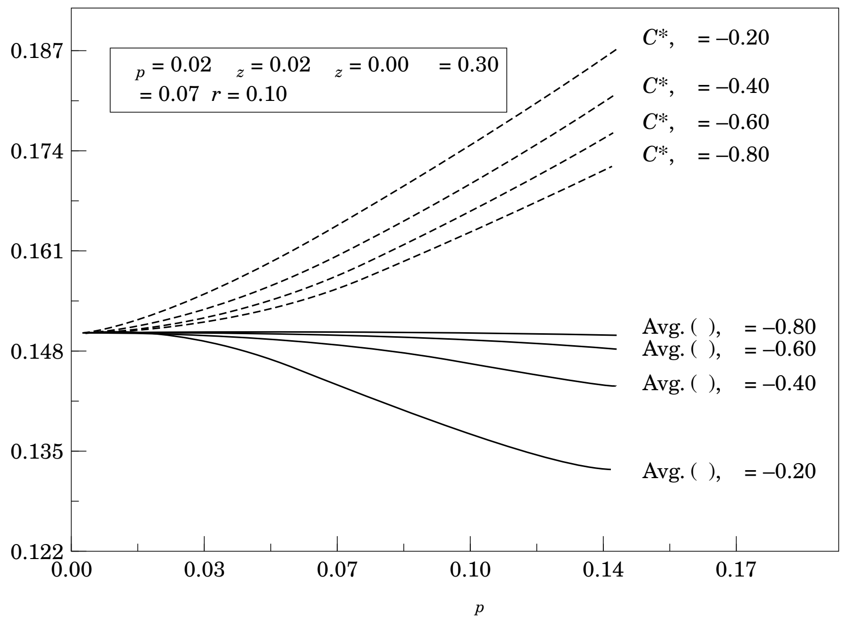


FIGURE D(i). *Z*—uncertainty. FIGURE D(ii). *P*—uncertainty.

to scale. He finds that the former is significantly lower than the latter, and interprets this finding as evidence of purposeful over-investment, possibly as an entry deterrent on the part of incumbent firms, or as evidence of increasing returns. While these explanations may be true in reality, the optimality condition tested by Hall does not hold if capital accumulation is irreversible at the firm level: investment irreversibility implies that in the long run capital's marginal profitability *should* be lower than the conventionally measured user cost of capital, even under constant returns to scale ($\theta = 1$), and this may well explain Hall's empirical findings.

3.3. THE VALUE OF THE FIRM

The value of the firm can be found by integrating the value of all infinitesimal units of capital, both those installed and those yet to be installed, given in equations (2.17) and (2.18). If the firm were never to install any more capital, each of the currently installed units would still (while progressively depreciating) produce a cash flow with present expected value as given in equation (2.17). But the firm does hold the option to install more capital: an option is always valuable, since it provides its owner with the right, but not the obligation, to acquire an asset. In the framework considered here, the option's value is given by equation (2.19) for every infinitesimal unit of capital.

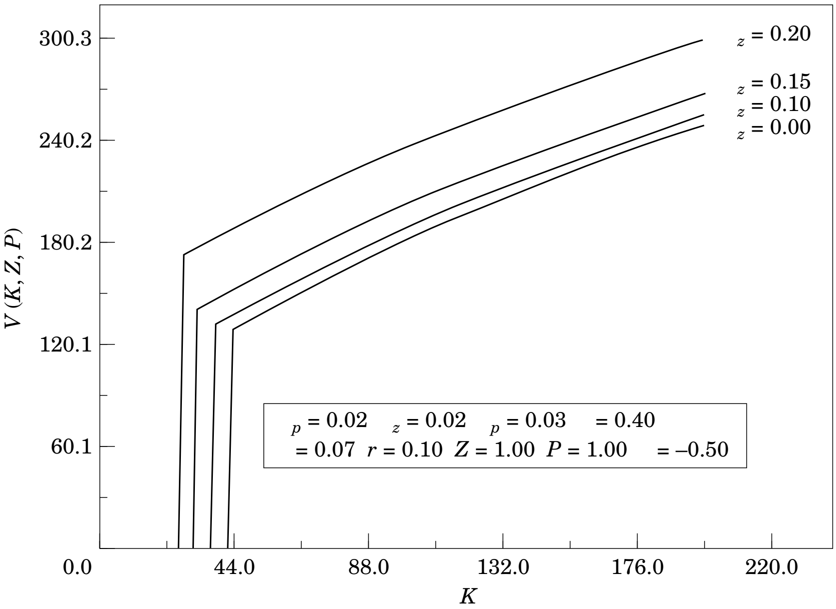
Simple integration obtains the following expression for the firm's value:

$$\begin{aligned} V^*(K_t, Z_t, P_t) &= \int_0^{K_t} v(\chi, Z_t)\mathrm{d}\chi + \int_0^{\infty} F(\chi, Z_t, P_t)\mathrm{d}\chi \qquad (3.5) \\ &= \int_0^{K_t} \frac{\chi^\beta Z_t}{r+\delta-(\theta_z-\delta\beta)}\mathrm{d}\chi + \int_{K_t}^{\infty} \frac{1}{A-1}\left(\frac{\chi^\beta Z_t}{c^* P_t}\right)^{A-1} \frac{\chi^\beta Z_t}{c^*}\mathrm{d}\chi \\ &= \frac{1}{1+\beta}\frac{K_t^{1+\beta} Z_t}{r+\delta-(\theta_z-\delta\beta)} + \int_0^{\infty} \frac{1}{A-1}\frac{1}{-\beta A-1}\left(\frac{K_t^\beta Z_t}{c^* P_t}\right)^A P_t K_t. \end{aligned}$$

Note that $-\beta A>1$ is necessary and sufficient for convergence of the second integral above, and that, as noted below Proposition 4, this is guaranteed to be true as long as the condition in equation (2.13) above is satisfied. If $\beta A \leqslant 1$, the value function fails to exist because the value of the options to install capital in the future does not converge.

Also note that $K_t<K^*$ is never observed. If that were the case the capital stock would instantaneously be increased to K^*; in the $K<K^*$ region, the value function can be defined as

$$V^*(K^*(Z_t, P_t), Z_t, P_t) - P_t(K^*(Z_t, P_t) - K).$$



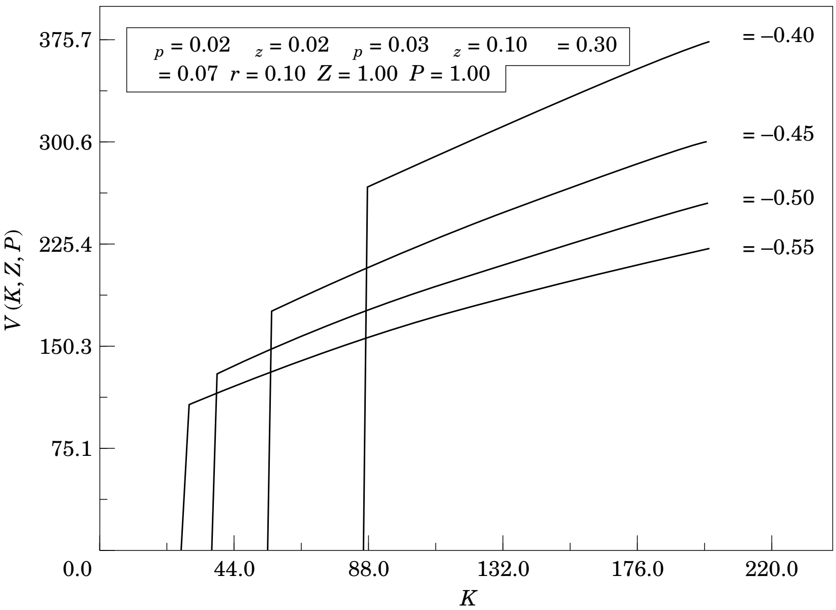


FIGURE E(i). Value of firm. FIGURE E(ii). Value of firm.

If for any reason the firm finds that $K<K^*$,(K^*-K) units of capital are immediately purchased and $V^*(K^*, Z, P)$ is obtained. V^* is continuously differentiable, is concave in the relevant region and its slope equals P_t at or below $K^*(P_t, Z_t)$.

Figure E(i) plots the value of the firm against K for several values of σ_z (the values for the other parameters are given in the figure); and Figure E(ii) performs the same experiment for several values of β.

For given K, Z and P, the firm is more valuable the more volatile the business conditions process. As Pindyck (1988) notes, when demand is very volatile the options to invest are worth more. Most of the value of firms faced by high uncertainty consists of the opportunity to invest in the future ("growth options"). Firms with a low $|\beta|$ (indicating high monopoly power and/or strongly decreasing returns to scale) have higher value, given K, for the same value of Z and P.

By Ito's lemma, the firm's value has dynamics given by

$$
\begin{aligned}
\mathrm{d}V^*(K_t, Z_t, P_t) = &\frac{1}{1+\beta}\frac{K_t^{1+\beta}Z_t}{r+\delta-(\theta_z-\delta\beta)} \\
&\times[(-\delta(1+\beta)+\theta_z)\mathrm{d}t+\sigma_z\mathrm{d}W_{zt}] \\
&+\frac{1}{A-1}\frac{1}{-\beta A-1}\frac{K_t^{\beta A+1}Z_t^A}{c^{*A}P_t^{A-1}} \\
&\times\left[\left(-\delta(\beta A+1)+A\theta_z+(1-A)\theta_p+\frac{1}{2}A(A-1)\sigma^2\right)\mathrm{d}t\right. \\
&\left.+A\sigma_z\mathrm{d}W_{zt}+(1-A)\sigma_p\mathrm{d}W_{pt}\right].
\end{aligned}
\tag{3.6}
$$

Noting that

$$
-\delta(\beta A+1)+A\theta_z+(1-A)\theta_p+\frac{1}{2}A(A-1)\sigma^2=r
$$

by the definition of A, it is possible to verify that

$$
\begin{aligned}
E_t\mathrm{d}V^*(K_t, Z_t, P_t) = &\, rV^*(K_t, Z_t, P_t) \\
&-\left[\frac{1}{1+\beta}K_t^{1+\beta}Z_t\mathrm{d}t-P_t(\mathrm{d}K_t-\delta K_t\mathrm{d}t)\right]
\end{aligned}
$$

as implied by the definition of V^* in equation (2.8) above.

The fluctuations of the rate of return around its expected value are *not* normally distributed:

$$\begin{aligned}
&\mathrm{d}V^*(K_t, Z_t, P_t) - E_t V^*(K_t, Z_t, P_t) \\
&= \frac{1}{1+\beta} \frac{K_t^{1+\beta} Z_t}{r+\delta-(\theta_z-\delta\beta)} \sigma_z \mathrm{d}W_{zt} \\
&\quad + \frac{1}{A-1} \frac{1}{-\beta A-1} \frac{K_t^{\beta A+1} Z_t^A}{c^{*A} P_t^{A-1}} [A\sigma_z \mathrm{d}W_{zt} + (1-A)\sigma_p \mathrm{d}W_{pt}].
\end{aligned}$$

This expression is not proportional to V^*, i.e. the value of the firm does not follow a geometric Brownian motion process. The firm's value fails to have a conditionally log–normal distribution because the cash-flow process follows geometric Brownian motion almost always, but has a singular component at times of positive gross investment. Alternatively, the total firm value is given by the sum of discounted operating profits from currently installed capital and of the "growth options". Each of the components has normally distributed returns under the assumptions made above, but the relative importance of the two components varies as the firm finds itself closer or farther from the investment point. When far from the investment point, the options to invest in the future are less valuable and account for a smaller share of the firm's value. The options being more volatile than the profits from installed capital, the conditional variance of returns *decreases* after an increase in the price of capital or a deterioration of the business conditions. Since such events also decrease the total value of the firm, it appears that investment irreversibility should imply *lower* variability after abnormally low realizations of returns. Further research should attempt to reconcile this finding with empirical evidence on stock market returns.

3.4. AVERAGE Q AND MARGINAL Q

The firm's value is non-stationary, but is cointegrated with the current replacement value of the installed capital stock. Defining Tobin's *average* Q as $Q_A \equiv V^*(K_t, Z_t, P_t)/P_t K_t$, we find

$$\begin{aligned}
Q_A &= \frac{1}{1+\beta} \frac{\dfrac{K_t^\beta Z_t}{P_t}}{r+\delta-(\theta_z-\delta\beta)} + \frac{1}{A-1} \frac{1}{-\beta A-1} \left(\frac{K_t^\beta Z_t}{cP_t}\right)^A \\
&= \xi_t \left[\frac{1}{(1+\beta)(r+\delta-(\theta_z-\delta\beta))} + \frac{1}{A-1} \frac{1}{-\beta A-1} \left(\frac{\xi_t}{c}\right)^{A-1} \frac{1}{c}\right].
\end{aligned}$$

Average Q is a monotonic function of ξ_t (since $A>1$ and $-\beta A>1$), and therefore possesses an ergodic distribution as long as ξ_t does.

The inverse function $\xi_t = f(Q_A)$ does not have a closed form, making it necessary to use numerical methods to obtain the ergodic distribution of Q_A.

Since $\xi_t \leqslant c$ and $A-1>0$, there is an upper bound to average Q:

$$\mathrm{Max}(Q_A) = c\left[\frac{1}{(1+\beta)(r+\delta-(\theta_z-\delta\beta))} + \frac{1}{A-1}\frac{1}{-\beta A-1}\frac{1}{c}\right]$$

$$= \frac{A}{A-1} + \frac{1}{A-1}\frac{1}{-\beta A-1} = \frac{A(-\beta A-1)+1}{(A-1)(-\beta A-1)}.$$

Positive investment is only observed when average Q attains its maximum value. The formula for V^* could in principle be used to correct the specification of empirical investment equations that use stock market data to compute average Q: Hayashi (1982) shows in a convex-costs-of-adjustment model that average Q is the correct independent variable in an investment equation only under perfect competition and constant returns to scale. Here it is necessary to violate at least one of these conditions to obtain the investment rule, but the results still provide useful insights.

A more appropriate explanatory variable for investment is the so-called "Marginal Q", defined as the ratio of the shadow price of capital to the market price of uninstalled capital. Abel and Blanchard (1986) argue that not only should investment be related to this quantity, but that marginal Q should be the *only* determinant of investment decisions.

Under the assumption of investment irreversibility, marginal Q is easily computed from equation (2.15) above as

$$Q_m = \frac{K_t^\beta Z_t - \dfrac{1}{A}\left(\dfrac{K_t^\beta Z_t}{c^* P_t}\right)^{A-1} K_t^\beta Z_t}{r+\delta+\delta\beta-\theta_z}\,\frac{1}{P_t}.$$

Using the definition of ξ_t in equation (3.1),

$$Q_m = \frac{\xi_t - \dfrac{1}{A}\left(\dfrac{\xi_t}{c^*}\right)^{A-1}\xi_t}{r+\delta+\delta\beta-\theta_z}.$$

Under the investment rule (equations (2.14a) and (2.14b)) $Q_m \leqslant 1$ always, and $Q_m = 1$ when the firm is investing (this non-linear relationship between irreversible investment and Q is comparable to the one found by Sargent (1980) in his general equilibrium model). Q_m is monotonically increasing in ξ_t in the relevant range

$0<\xi_t \leqslant c^*$, and it is therefore possible to compute the ergodic distribution of Q_m using ξ's distribution derived above. Once again, the function $Q_m = Q_m(\xi)$ does not have an inverse in closed form, and it is necessary to invert it numerically to find $\xi = Q_m^{-1}(\mathrm{q})$.

4. Conclusions and directions for further research

This paper solves an irreversible sequential investment problem under uncertainty. Cobb–Douglas technology, constant elasticity demand and geometric Brownian motion stochastic processes allow the firm's investment rule to be found in closed form either by a dynamic programming argument, or by marginal option pricing techniques.

In the presence of uncertainty, to trigger investment the marginal revenue product of capital must be higher than the conventionally measured cost of capital, because of investment irreversibility, even though the firm's owners are assumed to be risk neutral. As noted by Pindyck (1988), there is informal evidence that managers often discount the expected revenues from an investment project at a rate far higher than the one implied by any reasonable risk premium. The model considered here shows that, under certain conditions, this may indeed be very close to the optimal investment rule.

Although a higher marginal profitability of capital is needed *ex ante* to trigger investment, study of the steady-state properties of the model finds that fluctuations in business conditions and in the price of capital will on average make the marginal revenue product of capital *lower, ex post,* than the conventionally measured rental cost of capital.

The results have important empirical implications, which should be explored in future research. Some degree of investment irreversibility is realistic at the individual firm's level, and idiosyncratic uncertainty is probably large enough to make the irreversibility constraint important. Irreversibility of capital accumulation is, however, most realistic at the aggregate level. To use the model proposed here in macroeconomic applications it will be necessary to solve complex aggregation and estimation problems, and to devise realistic and tractable assumptions about the degree of flexibility in the use of installed capital and about used capital markets.† Allowance for time-varying parameters (especially σ, the degree of uncertainty facing the firm, and r, the

† The model is readily modified—losing the closed-form solution—to allow for resale of used capital at a constant discount (see Bertola, 1988, Chapter 1; Bentolila and Bertola (1988) use a similar model to characterize demand for labour in the presence of firing restrictions).

required rate of return) would also probably be necessary before the model can be empirically implemented.

Besides its relevance to empirical study of investment, irreversibility has interesting implications about the cyclical behaviour of all economic variables. The installed capital stock will often be too large *ex post,* inducing hysteresis (path dependence) and affecting the responsiveness of prices, employment and production to adverse changes in a firm's business conditions;† even though the processes exogenous to the firm are modelled above as simple logarithmic random walks, investment occurs at distinct times in the model above, and this could possibly generate a fairly regular multiplier–accelerator cycle in a completely specified macroeconomic model.

References

Abel, A.B. (1983). Optimal investment under uncertainty. *American Economic Review,* **73**, 228–233.

Abel, A.B. (1985). A stochastic model of investment, marginal Q and the market value of the firm. *International Economic Review,* **26**, 305–322.

Abel, A.B. & Blanchard, O.J. (1986). The present value of profits and cyclical movements in investment. *Econometrica,* **54**, 249–273.

Ando, A., Modigliani, F., Rasche, O. & Turnovsky, S. (1974). On the role of expectations of price and technological change in an investment function. *International Economic Review,* **1974**, 384–414.

Arrow, K.J. (1968). Opimal capital policy with irreversible investment. In J.N. Wolfe, Ed. *Value, Capital and Growth. Papers in Honour of Sir John Hicks.* Edinburgh: Edinburgh University Press.

Benes, V.E., Shepp, L.A. & Witsenhausen, H.S. (1980). Some solvable stochastic control problems. *Stochastics,* **4**, 39–83.

Bentolila, S. & Bertola, G. (1988). *Firing costs and labor demand: how bad is eurosclerosis?* M.I.T.: mimeo.

Bernanke, B.S. (1983). Irreversibility, uncertainty and cyclical investment. *The Quarterly Journal of Economics*, **98**, 85–106.

Bertola, G. (1988). Adjustment costs and dynamic factor demands: investment and employment under uncertainty. M.I.T.: Ph.D. Thesis.

Chow, P.-L., Menaldi, J. & Robin, M. (1985). Additive control of stochastic flow systems with finite horizon. *SIAM Journal of Control and Optimization,* **23**, 858–899.

Cox, D.R. & Miller, H.D. (1965). *The Theory of Stochastic Processes.* London: Chapman & Hall.

Dietrich, J.K. & Heckerman, D.G. (1980). Uncertain inflation and the demand for capital. *Economic Inquiry,* **18**, 461–471.

Dixit, A. (1987). *Entry and exit decisions of firms under fluctuating real exchange rates.* Princeton University: mimeo.

Dixit, A. (1988). *An heuristic interpretation for the smooth pasting condition.* Princeton University: mimeo.

† This insight is most relevant to the analysis of the responsiveness of trade flows and intersectoral reallocations to changes in real exchange rates: see, for example, Dixit (1987).

Hall, R.E. (1987). *Investment under uncertainty: theory and tests using industry data.* N.B.E.R. Working Paper.
Harrison, J.M. (1985). *Brownian Motion and Stochastic Flow Systems.* New York: John Wiley & Sons.
Hayashi, F. (1982). Tobin's marginal q and average q: a neoclassical interpretation. *Econometrica,* **50**, 213–224.
Ingersoll, J.E. Jr & Ross, S.A. (1987). *Waiting to invest: investment and uncertainty.* Yale University: mimeo.
Jorgenson, D.W. (1963). Capital theory and investment behavior. *American Economic Review Papers and Proceedings,* **43**, 247–259.
Karatzas, I. & Shreve, S.E. (1984). Connections between optimal stopping and singular stochastic control I: monotone follower problems. *SIAM Journal of Control and Optimization,* **21**, 856–877.
Karatzas, I. & Shreve, S.E. (1985). Connections between optimal stopping and singular stochastic control II: reflected follower problems. *SIAM Journal of Control and Optimization,* **22**, 433–451.
Karatzas, I. & Shreve, S.E. (1988). *Brownian Motion and Stochastic Calculus.* New York: Springer-Verlag.
Lucas, R.E. Jr & Prescott, E.C. (1971). Investment under uncertainty. *Econometrica,* **39**, 659–681.
McDonald, R. & Siegel, D. (1985). Investment and the valuation of firms when there is an option to shut down. *International Economic Review,* **26**, 331–349.
McDonald, R. & Siegel, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics,* **101**, 707–727.
Merton, R.C. (1973). Theory of rational option pricing. *Bell Journal of Economics,* **41**, 141–183.
Nickell, S. (1974). On the role of expectations in the pure theory of investment. *Review of Economic Studies,* **41**, 1–19.
Pindyck, R.S. (1988). *Irreversible investment, capacity choice, and the value of the firm.* M.I.T.: mimeo.
Sargent, T.J. (1979). *Macroeconomic Theory.* London: Academic Press.
Sargent, T.J. (1980). Tobin's Q and the rate of investment in general equilibrium. In K. Brunner & A.H. Metzler, Eds. *On the State of Macroeconomics*. Amsterdam: North Holland.

Appendix 1—Derivation of equation (2.15)

Define $\eta_t \equiv \partial\Pi/\partial K = K_t^{\beta} Z_t$. Also define a functional expression for the conditional expectation defining the shadow value of capital appearing in equations (2.9) and (2.10):

$$f(\eta_t, P_t; c^*) \equiv E_t\left\{\int_t^{\infty} \eta_\tau e^{-(r+\delta)(\tau-t)} d\tau\right\}. \tag{A1.1}$$

When $dX_t = 0$ and $\eta_t < cP_t$ we have $dK_t = -\delta dt$; define $\zeta_t \equiv \{K_t^{\beta} Z_t; dX_t = 0\}$, and use Ito's lemma to find its stochastic differential:

$$\begin{aligned}
\mathrm{d}\zeta_t &= \frac{\partial \xi_t}{\partial K_t}\,\mathrm{d}K_t + \frac{\partial \xi_t}{\partial Z_t}\,\mathrm{d}Z_t + \frac{1}{2}\left[\frac{\partial^2 \xi_t}{\partial K_t^2}(\mathrm{d}K_t)^2\right.\\
&\quad \left. + \frac{\partial^2 \xi_t}{\partial Z_t^2}(\mathrm{d}Z_t)^2 + 2\,\frac{\partial^2 \xi_t}{\partial K_t \partial Z_t}(\mathrm{d}K_t \mathrm{d}Z_t)\right]\\
&= \beta K_t^{\beta-1} Z_t(-\delta K_t \mathrm{d}t) + K_t^{\beta}(\theta_z Z_t \mathrm{d}t + Z_t \sigma_Z \mathrm{d}W_{zt})\\
&\quad + \frac{1}{2}[\beta(\beta-1)K_t^{\beta-2}0 + \sigma^2 Z^2 \mathrm{d}t + 2\beta K_t^{\beta-1}0]\\
&= \zeta_t(-\delta\beta + \theta_z)\mathrm{d}t + \zeta_t(\sigma_z \mathrm{d}W_{zt})\\
&\equiv \zeta_t \mu \mathrm{d}t + \sigma_z \mathrm{d}W_{zt} \qquad \text{(A1.2)}
\end{aligned}$$

where $\mu \equiv -\delta\beta + \theta_z$ is defined.

Recall from the main text that

$$\mathrm{d}P_t = \theta_p P_t \mathrm{d}t + P_t \sigma_p \mathrm{d}W_{pt}, \quad \mathrm{d}W_{zt}\mathrm{d}W_{pt} = \rho\sigma_z\sigma_p.$$

Take $\zeta_0 = cP_0$ (to imply that time 0 is chosen to be a time of positive investment, with no loss of generality).

The investment policy followed by the firm is such that $\{\eta_t\}$ is obtained from regulating the geometric Brownian motion $\{\zeta_t\}$ with a *moving* control barrier at cP_t; equivalently, the firm regulates $\{\eta_t\}$ so that $\eta_t/P_t \leqslant c$ for all t. Adapting the approach in Harrison (1985), the stochastic process $\{\eta_t\}$ is defined as

$$\eta_t = \frac{\xi_t}{U_t} \qquad \text{(A1.3)}$$

or equivalently

$$\eta_t/P_t = \frac{\xi_t/P_t}{U_t}c \qquad \text{(A1.4)}$$

where the control process $\{U_t\}$ is the non-decreasing and continuous *running maximum* of $\{\zeta_t/P_t\}$:

$$U_t \equiv \sup_{t' < t}(\zeta_t/P_t);$$

$\{U_t\}$ only increases when $\eta_t = cP_t$. The non-decreasing process $\{U_t\}$ has finite variation, $(\mathrm{d}U_t)^2 = (\mathrm{d}U_t \mathrm{d}\zeta_t) = 0$. Applying Ito's lemma to η_t all the second-order terms vanish to yield:

$$\mathrm{d}\eta_t = \mathrm{d}\left(\frac{\xi_t L_t}{U_t}\right) = \frac{1}{U_t}\mathrm{d}\zeta_t - \frac{\xi_t L_t}{U_t^2}\mathrm{d}U_t$$

$$= \frac{1}{U_t}\zeta_t \mu \mathrm{d}t + \frac{1_t}{U_t}\sigma_z \mathrm{d}W_{zt} - \frac{\zeta_t}{U_t}\frac{\mathrm{d}U_t}{U_t}$$

$$= \zeta_t \mu \mathrm{d}t + \eta_t \sigma_z \mathrm{d}W_{zt} - \eta_t \frac{\mathrm{d}U_t}{U_t}. \qquad \text{(A1.5)}$$

Now consider $f(\eta_t, P_t)$, the conditional expectation defined in equation (A1.1), and apply Ito's lemma again to obtain (subscripts denote partial derivatives):

$$\begin{aligned}
\mathrm{d}f(\eta_t, P_t) = {} & f_1(\eta_t, P_t)\mathrm{d}\eta_t + f_2(\eta_t, P_t)\mathrm{d}P_t \\
& + \frac{1}{2} f_{11}(\eta_t, P_t)(\mathrm{d}\eta_t)^2 + \frac{1}{2} f_{22}(\eta_t, P_t)(\mathrm{d}P_t)^2 \\
& + f_{12}(\eta_t, P_t)(\mathrm{d}\eta_t P_t) \\
= {} & f_1(\eta_t, P_t)(\eta_t \mu \mathrm{d}t + \eta_t \sigma_z \mathrm{d}W_{zt}) - f_1(cP_t, P_t)cP_t \frac{\mathrm{d}U_t}{U_t} \\
& + f_2(\eta_t, P_t)(\theta_p P_t \mathrm{d}t + P_t \sigma_p \mathrm{d}W_{pt}) \\
& + \frac{1}{2} f_{11}(\eta_t, P_t)\sigma_z^2 \eta_t^2 \mathrm{d}t + \frac{1}{2} f_{22}(\eta_t, P_t)\sigma_p^2 P_t^2 \mathrm{d}t \\
& + f_{12}(\eta_t, P_t)\eta_t P_t \rho \sigma_z \sigma_p \mathrm{d}t
\end{aligned} \qquad \text{(A1.6)}$$

where the fact that $\mathrm{d}U_t \neq 0$ only if $\eta_t = cP_t$ is used in obtaining the last equality.

Recall now the Integration By Parts formula found in Harrison (1985, page 73): if $\{Y_t\}$ is an Ito process (i.e. the stochastic integral $\int \mathrm{d}Y_t$ is well defined) and $\{X_t\}$ is a continuous process with finite variation, then

$$Y_\nu X_\nu = Y_t X_t + \int_0^\nu Y_\tau \mathrm{d}X_\tau + \int_0^\nu X_\tau \mathrm{d}Y_\tau.$$

Apply the Integration by Parts formula to $Y_\nu = f(\eta_\nu, P_\nu)$ and $X_\nu = \mathrm{e}^{-(r+\delta)(\nu - t)}$; using

$$\mathrm{d}[\mathrm{e}^{-(r+\delta)(\tau - t)}] = -(r+\delta)\mathrm{e}^{(r+\delta)(\tau - t-}\mathrm{d}\tau$$

and $\mathrm{d}f(\eta_\tau, P_\tau)$ from equation (A1.6), one obtains, after rearranging terms,

$$
\begin{aligned}
e^{-(r+\delta)\nu} f(\eta_\nu, P_\nu) = f(\eta_t, P_t) \\
&+ \int_t^\nu e^{-(r+\delta)(\tau-t)} \Big(f_1(\eta_\tau, P_\tau)\eta_\tau \mu + f_2(\eta_\tau, P_\tau)\theta_p P_\tau \\
&+ \frac{1}{2} f_{11}(\eta_\tau, P_\tau)\sigma_z^2 \eta_\tau^2 + \frac{1}{2} f_{22}(\eta_\tau, P_\tau)\sigma_p^2 P_\tau^2 \\
&+ f_{12}(\eta_\tau, P_\tau)\eta_\tau P_\tau \rho \sigma_z \sigma_p - (r+\delta) f(\eta_\tau, P_\tau) \Big) d\tau \\
&+ \int_t^\nu f_1(\eta_\tau, P_\tau)\eta_\tau \sigma_z dW_{z\tau} + \int_\tau^\nu f_2(\eta_\tau, P_\tau) P_\tau \sigma_p dW_{p\tau} \\
&- \int_t^\nu f_1(\upsilon P_\tau, P_\tau)\upsilon P_\tau \frac{dU_\tau}{U_\tau}. \qquad \text{(A1.7)}
\end{aligned}
$$

Now let $\nu \to \infty$ and take the conditional expectation of equation (A1.7) at t to obtain

$$
\begin{aligned}
0 = f(\eta_t, P_t) + E_t \Big\{ \int_t^\infty e^{-(r+\delta)(\tau-t)} \Big(f_1(\eta_\tau, P_\tau)\eta_\tau \mu \\
&+ f_2(\eta_\tau, P_\tau)\theta_p P_\tau + \frac{1}{2} f_{11}(\eta_\tau, P_\tau)\sigma_z^2 \eta_\tau^2 \\
&+ \frac{1}{2} f_{22}(\eta_\tau, P_\tau)\sigma_p^2 P_\tau^2 + f_{12}(\eta_\tau, P_\tau)\eta_\tau P_\tau \rho \sigma_z \sigma_p \\
&- (r+\delta) f(\eta_\tau, P_\tau) \Big) d\tau \Big\} - E_t \Big\{ \int_t^\infty f_1(\upsilon P_\tau, P_\tau)\upsilon P_\tau \frac{dU_\tau}{U_\tau} \Big\} \qquad \text{(A1.8)}
\end{aligned}
$$

where we use

$$
\lim_{\nu \to \infty} ((e^{-(r+\delta)\nu} f(\eta_\nu, P_\nu)) = 0
$$

and, for all $\nu \geqslant t$,

$$
E_t \Big\{ \int_t^\nu f_1(\eta_\tau, P_\tau)\eta_\tau \sigma_z dW_{z\tau} \Big\} = E_t \Big\{ \int_t^\nu f_2(\eta_\tau, P_\tau) P_\tau \sigma_p dW_{t\tau} \Big\} = 0
$$

(by Proposition 4.3.7 in Harrison (1985), and boundedness of f_1 and f_2).

Equations (A1.1) and (A1.8) can only be simultaneously true if

$$\begin{aligned} & f_1(\eta, P)\eta\mu + f_2(\eta, P)\theta_p P + \frac{1}{2} f_{11}(\eta, P)\sigma_z^2\eta^2 \\ & + \frac{1}{2} f_{22}(\eta, P)\sigma_p^2 P^2 + f_{12}(\eta_t, P_t)\eta_t P_t \rho\sigma_z\sigma_p \\ & -(r+\delta)f(\eta, P) = -\eta \end{aligned} \tag{A1.9}$$

for all η and P such that $(\eta/P) \leqslant c$, *and* the last term in equation (A1.8) vanishes: since $\mathrm{d}U > 0$ only when $\eta = cP$, this implies

$$f_1(cP, P)cP = 0 \quad \text{for all } P. \tag{A1.10}$$

The general form of the solution to the non-homogeneous parabolic equation (A1.9) is a linear combination of power functions and of a linear term,

$$f(\eta, P) = B_1\eta + B_2\eta^A P^B. \tag{A1.11}$$

Impose the boundary condition (equation (A1.10)):

$$B_1 cP + B_2 A(cP)^A P^B = 0. \tag{A1.12}$$

This can be satisfied for all P only if $B = 1 - A$; rewrite then equation (A1.10) as

$$B_1 c + B_2 A c^A = 0 \tag{A1.13}$$

and equation (A1.12) as

$$f(\eta, P) = B_1\eta + B_2(\eta/P)^A P. \tag{A1.14}$$

Inserting equation (A1.14) in equation (A1.9), one finds that $B_1 = 1/(r+\delta-\mu)$ and that A must be a solution of

$$\frac{\sigma^2}{2}A^2 + \left(\mu - \theta_p - \frac{\sigma^2}{2}\right)A - (r+\delta-\theta_p) = 0$$

where $\sigma^2 \equiv \sigma_z^2 + \sigma_p^2 - 2\rho\sigma_z\sigma_p > 0$.

This quadratic equation has two real roots of opposite sign, provided that $r+\delta-\theta_p > 0$. Since η_t/P_t can be arbitrarily close to zero, the positive root must be chosen for A to obtain a bounded $f(\eta, P)$:

$$A \equiv \frac{-\left(\mu-\theta_p-\frac{\sigma^2}{2}\right)+\sqrt{\left(\mu-\theta_p-\frac{\sigma^2}{2}\right)^2+2(r+\delta-\theta_p)\sigma^2}}{\sigma^2}>0.$$

From equation (A1.13) we get

$$B_2=\frac{c^{1-A}/A}{r+\delta-\mu}. \tag{A1.15}$$

Inserting this expression in equation (A1.14) completes the derivation of equation (2.15).

Appendix 2

The value of the *marginal* unit available for installation at time t, given that the already installed capital stock K_t is depreciating at rate δ, has dynamics given by Ito's lemma as

$$\begin{aligned}
\mathrm{d}v(K_t, Z_t) &= \frac{\mathrm{d}(K_t^{\beta} Z_t)}{r+\delta-(\theta_z-\delta\beta)} \\
&= \frac{1}{r+\delta-(\theta_z-\delta\beta)}[\beta K_t^{\beta-1} Z_t(-\delta K_t \mathrm{d}t) \\
&\quad + K_t^{\beta}(\theta_z Z_t \mathrm{d}t + Z_t \sigma_z \mathrm{d}W_{zt})] \\
&= v(K_t, Z_t)(-\delta\beta+\theta_z)\mathrm{d}t + \mathrm{v}(K_t, Z_t)\sigma_z \mathrm{d}W_{zt} \\
&= v(K_t, Z_t)\mu \mathrm{d}t + v(K_t, Z_t)\sigma_z \mathrm{d}W_{zt}.
\end{aligned}$$

Recall that $\mathrm{d}P_t = P_t\theta_p \mathrm{d}t + P_t\sigma_p \mathrm{d}W_{pt}$, with $\mathrm{d}W_{pt}\mathrm{d}W_{zt}=\rho$. Denote $F(\eta_t, P_t)$ the value of the call option to purchase the unit of capital with value $v(K_t, Z_t)$. When the option is not yet exercised, the expected rate of return to its holder must equal the required rate of return, r, *plus* the depreciation rate δ: the optimal stopping problem is defined on the currently marginal unit of capital. Omitting time indexes for convenience, and using Ito's lemma, the unexercised option's value must satisfy

$$\begin{aligned}
F_1(\eta, P)\eta\mu + F_2(\eta, P)\theta_p P + \frac{1}{2}F_{11}(\eta, P)\sigma_z^2\eta^2 \\
+ \frac{1}{2}F_{22}(\eta, P)\sigma_p^2 P^2 = (r+\delta)F(\eta, P)
\end{aligned} \tag{A2.1}$$

where suscripts denote partial derivatives.

In addition, it must be that

$$F(0, P) = 0 \quad \text{for } P>0 \tag{A2.2}$$

i.e. the option is worthless if $\eta = v(K, Z) = 0$, an absorbing state.

The general solution to the homogeneous parabolic equation (A2.1) has the form

$$F(\eta, P) = B_1 \eta^{\alpha_1} P^{\beta_1} \tag{A2.3}$$

with $\alpha_1>0$ to satisfy equation (A2.2).

McDonald and Siegel (1986) prove that the exercise locus in (η, P) is homogeneous of degree zero; i.e. is defined by $\eta = cP$ for some c. On the locus the "Value Matching" condition must hold:

$$F(cP, P) + P = cP/(r + \delta(1+\beta) - \theta_z). \tag{A2.4}$$

Additional boundary equations are needed for determination of the boundary and of the option values: these are the so-called "smooth pasting" or "high impact" conditions (see Merton, 1973, footnote 60): since c is chosen to maximize the value of the option $F(\eta, P)$ among all possible $F'(\eta, P, c)$, it must be that

$$F(\eta, P) = \operatorname*{Max}_{c} F(\eta, P, c). \tag{A2.5}$$

Replacing F' for F in equation (A2.4), and totally differentiating with respect to c (Dixit, 1988 proves differentiability), one obtains

$$F_1'(cP, P, c)P + F_3'(cP, P, c) = P/(r + \delta(1+\beta) - \theta_z).$$

Note that $F_1' = F_1$ and $F_3' = 0$ by equation (A2.5), and obtain the smooth pasting condition

$$F_1(cP, P) = 1/(r + \delta(1+\beta) - \theta_z). \tag{A2.6}$$

The reader can verify, by inserting the functional form (equation (A2.3)) in equations (A2.1), (A2.4) and (A2.6), that the solution to differential equation (A2.3) and its boundary conditions is given by $F(\eta, P)$ defined in Proposition 5, and A and c defined in Proposition 4.