ENTRY, EXIT AND SCRAPPING DECISIONS WITH INVESTMENT LAGS: A SERIES OF INVESTMENT MODELS BASED ON A NEW APPROACH

Sigbjørn Sødal*

Agder University College, Kristiansand, Norway, and
University of California, Santa Barbara, USA

UCSB, Department of Economics, Working Paper in Economics #1-01
This version: March 15, 2001

Abstract: Several entry-exit models under price uncertainty are discussed by a new markup approach to investment, starting with the classical model by Dixit (1989). The markup approach, introduced by Dixit et al. (1999), enables us to state the expected value of the firm in the entry-exit model as a function of a chosen pair of entry and exit trigger prices. The optimal policy appears by maximizing the value function with respect to the trigger prices. Extensions being discussed include endogenous production costs, diminishing production capacity over time, limits to the number of available switches, and various models with scrapping decisions and investment lags. The main new extension allows for an investment lag in the entry-exit-scrapping model by Dixit (1988). Implications of the investment lag are investigated by use of experimental data and empirical data from shipping. We also correct some results on investment lags from Bar-Ilan and Strange (1996).

JEL classification: D92, D81, Q23.

Keywords: irreversible investment, uncertainty, markup pricing, investment lags.

1. Introduction

Models of entry and exit decisions under price uncertainty, pioneered by Mossin (1968) and generalized by Brennan and Schwarz (1985) and Dixit (1989), are main tools in real options theory. The last of these references, focused in this paper, solves the following problem: A firm can choose between using and mothballing a certain production capacity. The cost of producing is fixed, while the product price follows a geometric Brownian motion. When is it optimal to

---

* *Agder University College, Faculty of Economics and Social Sciences, Serviceboks 422, N-4604 Kristiansand, Norway. Telephone: +47 38 14 15 22; Telefax: +47 38 14 10 28; E-mail: Sigbjorn.Sodal@hia.no.* *This research has been supported by Agder Maritime Research Foundation. I am indebted to Avinash Dixit for helpful comments.*

produce, considering that switching is costly? The answer is familiar: the optimal policy for the idle firm is not to follow the Marshallian rule of starting production as soon as the expected net present revenue exceeds the costs of entry and production through infinity. One should wait for a fixed higher trigger price due to an option argument. Likewise, the active firm should wait for a fixed trigger price lower than the Marshallian one before leaving.

Dixit's model has been extended in several directions. Dixit (1988) includes (initial) construction costs and (final) scrapping costs. Brekke and Øksendal (1994) allow for diminishing production capacity over time. Ekern (1993) restricts the number of available switches, and Bar-Ilan and Strange (1996) include investment lags. Leahy (1993) showed more generally that the myopic behavior of the firm in these references is usually also optimal in partial equilibrium with a large number of firms acting independently.

The first objective of this paper is to show that a new approach to irreversible investment can be used to model all of the investment problems just mentioned. The basic idea of the approach stems from Dixit et al. (1999), which focuses on the analogy between the markup pricing rule of a static monopolist and the optimal timing decision for a firm with the option to invest. We find it natural to use the term "markup approach" for the new method as compared with the smooth pasting approach dominating the literature.[1] In this paper we expand the markup approach significantly. Our ambition is to show that the new approach simplifies both the analytical treatment and the numerical analysis of several entry-exit problems.[2]

The second objective of the paper is to use the markup approach to develop new extensions to the entry-exit model. More generally, however, we intend to show that the markup approach is suited for a wide spectrum of applications, so the list of extensions could easily be made longer. Section 8 contains the most thorough extension, including investment lags in the entry-exit-scrapping

---

[1] See Dixit and Pindyck (1994) for an introduction to real options theory based on smooth pasting. Sødal (1998) contains a brief discussion of the relationship between the markup approach and the smooth pasting approch.

[2] This does not remove the need for advanced stochastic analysis if one were to prove rigorously that the suggested solutions are indeed optimal. See Brekke and Øksendal (1994) for a discussion and a set of sufficiency conditions.

model by Dixit (1988). At that point the new approach has been developed so far that the model can be spelled out over a few paragraphs. A larger system of equations and tedious analysis seem to be necessary with the smooth pasting approach.

In the standard entry-exit setting, the core of the markup approach consists of stating the expected net present value of the firm as a function of an arbitrary pair of trigger prices, using a general, continuous and autonomous Ito price process. Thus we work with a broad class of stochastic processes, although the geometric Brownian motion is used for numerical examples. The optimal policy is found by maximizing the value function with respect to the trigger prices. Two trigger prices leave us with two first-order conditions that must be solved for simultaneously. The solution by smooth pasting is given by a system of four equations, but the two solutions coincide when two option coefficients are eliminated from the latter system.

Several models will be discussed, and it is not necessary to go into the details for all of them. We pay most attention to models with investment lags, for which the markup approach seems to be especially convenient. As part of our discussion on investment lags, we point at an error that infected all numerical results in Bar-Ilan and Strange (1996). We argue that some of their conclusions must be modified because of the error.[3]

The rest of the paper is structured as follows: Section 2 develops the basic entry-exit model of Dixit (1989), using the new approach and illustrating it with numerical examples. Section 3 contains a new extension, endogenizing the production cost. Section 4 restricts the number of switches as in Ekern (1993). Section 5 discusses diminishing production capacity as in Brekke and Øksendal (1994). Section 6 allows for investment lags as in Bar-Ilan and Strange (1996). Section 7 expands to construction and scrapping as in Dixit (1988), and section 8 includes investment lags in the latter model. This also allows for a study of investment lags in practice, based on empirical shipping data. Section 9 discusses equilibrium and section 10 sums up.

---

[3] The authors knew about the error when I contacted them; they had also suggested an erratum some time ago.

2. The entry-exit model

As in Dixit (1989), we consider a firm with the option to switch between using and mothballing its production capacity. The potential revenue from being an active producer is stochastic. The expected and discounted revenue from continuous production into infinity, starting now and discounting at a constant rate $\rho>0$, follows an autonomous and continuous Ito process

(1) $$dP = f(P)dt + g(P)dz \, .$$

Here dz is the Wiener increment, dt is the time increment, and f(P) and g(P) are continuous and differentiable functions representing drift and volatility of the total potential revenue process. The firm is not producing initially, and the process starts from a low initial value, P_0.[4]

The process (1) usually originates from a more basic price process. For example, if the firm produces continuously at rate unity, and the current price of the good, p, is geometric Brownian with drift μ ($<\rho$) and volatility σ, then $P=p/(\rho-\mu)$. Thus P is geometric Brownian with the same drift and volatility as the underlying process. For notational convenience we state all variables in expected net present terms right from the beginning. Having just noted that P is really an expected net present price, we mainly refer to it as a price. When presenting numerical results, this and other net present values are transformed back to more common flow variables.

The firm can start (or restart) production by the investment A, and stop production (temporarily or permanent) by the investment B. Both of these switching costs are constant, and they are referred to as the entry cost and the exit cost, respectively. The annual cost of production is a constant c, so the net present cost of producing forever is $c/\rho\equiv C$. Below we call C a production cost even if it is also a net present value like the price P. Costs and revenue from production only apply when the firm is active, so the objective of making entry is to gain from production when prices exceed marginal costs; the objective of exit is to avoid production when they do not.

The firm has two options: either to remain in the current state (idle or active) or to switch. Since investments are irreversible and the only dynamic component of the model is the price P, a given

[4] Alternatively, we could start with a high initial value and an active firm. That leads to the same optimal policy.

combination of state and price leads to the same decision at all points in time. Moreover, the perpetual character of the model ensures that the value of the firm coincides in such cases.

As argued intuitively by Dixit (1989) and proven rigorously by Brekke and Øksendal (1994), the optimal policy for the idle firm is to enter when a certain price is reached from below, and for the active firm to exit when a certain lower price is reached from above. The firm remains in the current state whenever the price is between these two trigger prices. Suppose therefore that the firm enters when the arbitrary price R is reached from below, and exits when the arbitrary price L (<R) is reached from above. For the moment we set P_0=L; i.e., we assume that the initial price coincides with the exit trigger price. Remember, however, that the firm is idle at first, waiting for R to be reached. The expected net present value of the firm, $W_L(L,R)$, can be written as

$$W_L(L,R) = Q(L,R)\big(R - A - C + Q(R,L)\big(C - B - L + W_L(L,R)\big)\big) \,, \tag{2}$$

where $Q(P_1,P_2) \equiv E[e^{-\rho T}]$ is the expected discount factor when P is to move from the current price P_1 to another price P_2 for the first time, and T is the first hitting time from P_1 to P_2.[5] Eq. (2) can be explained as follows: No revenue or cost applies until the entry trigger price R is reached, and all gains from then on must be discounted. In this case the start price is L and the end price is R, so the discount factor is Q(L,R). The net present revenue from remaining active through infinity is R, the firm has to invest A to become active, and the production cost is C. This explains the first part of the parenthesis, R–A–C. However, the firm does not have to remain active forever. Gains arising from the exit option are discounted further by Q(R,L), since the price at time of entry is R, and the firm will exit when the price gets down to L. At time of exit, the production cost C is saved if the firm should remain idle forever. The exit cost B must be paid, and the net present loss

---

[5] See Dixit et al. (1999) for a discussion, including several examples, of the discount factor function $Q(P_1,P_2)$. This function, which depends on the parameters of the price process (1), has the following general properties: It obviously satisfies $0 \le Q(P_1,P_2) \le 1$ for all P_1 and P_2. The maximum value, $Q(P_1,P_2)=1$, is reached when $P_1=P_2$ as there will be no discounting when the start point equals the end point. Moreover, $\partial Q/\partial P_1>0$ and $\partial Q/\partial P_2<0$ when $P_1<P_2$, since the distance to move (and the time needed, since the process is continuous) is decreasing in the first case and increasing in the other. For similar reasons, $\partial Q/\partial P_1<0$ and $\partial Q/\partial P_2>0$ when $P_1>P_2$. Finally, $Q(P_1,P_2)\to 0$ as P_1 and P_2 get far apart, since that implies a long way to move, and thus heavy discounting.

of revenue from production is L. This explains the first part of the inner parenthesis, C–B–L. Right after exit, the initial combination of price and state has been reached again. As discussed above, the value of the firm coincides in these situations. Therefore we add the recursive element $W_L(L,R)$ to find the expected net present value of the firm over infinite time.

By re-arranging Eq. (2), $W_L(L,R)$ can be written as a function of the two trigger prices and exogenous parameters of the model. Recall, however, that the true initial price is not L, but a fixed lower price, P_0. Therefore the true value function, or profit function, is not $W_L(L,R)$, but $Q(P_0,L)W_L(L,R)\equiv W_0(L,R)$. Using that $Q(P_0,L)Q(L,R)=Q(P_0,R)$, this translates to:

(3) $$W_0(L,R)=\frac{Q(P_0,R)\big(R-A-C+Q(R,L)(C-B-L)\big)}{1-Q(L,R)Q(R,L)}.$$

Eq. (3) is the key of the model, showing the expected net present value of the firm at the initial point, P_0, as a function of the chosen trigger prices. The numerator contains the expected value of the first production period. The denominator, which is close to zero if the production cycle is short when measured by expected discounting, accounts for the infinite time horizon. The optimal policy is found by maximizing W_0 with respect to L and R. Before showing the result of this exercise, we note some general properties of the model.

The optimal policy does not depend on the initial price, P_0. To see this, write the profit function as $W_0(L,R)=Q(P_0,R)F(R,L)$, where the definition of F(R,L) follows from Eq. (3). As long as W_0 is smooth, the optimal solution requires $\partial W_0/\partial L=0$ and $\partial W_0/\partial R=0$. Since $Q(P_0,R)$ does not depend on L, the first of these first-order conditions does not depend on P_0. The other one becomes $F\partial Q/\partial R+Q\partial F/\partial R=0$. Defining $\varepsilon=-(\partial Q/\partial R)/(Q/R)$ and $\gamma=(\partial F/\partial R)/(F/R)$, this simplifies to $\varepsilon=\gamma$. As shown by Dixit et al. (1999), ε is independent of P_0. Nor can γ depend on P_0, so the second first-order condition is also independent of P_0.

For the model to make sense, we must have $A+B\geq 0$. In the opposite case the firm could make infinite profit by switching back and forth continuously. To see this formally, rewrite W_0 as

(3') $$W_0(L,R)=Q(P_0,R)\frac{(R-L)-(1-Q(R,L))(C-B-L)-(A+B)}{1-Q(L,R)Q(R,L)}.$$

The properties of the discount factor function imply that the denominator in Eq. (3') decreases strictly in L, approaching zero as $L \to R$. The first parts of the numerator also approach zero as $L \to R$, so unless $A+B \geq 0$, profits could be increased to infinity by letting $L \to R$.

The two first-order conditions for a maximum of W_0, determining the optimal trigger prices, can be written as follows (see Appendix A):

$$\text{(4a)} \qquad \frac{R-A-C}{R} = \frac{1}{\varepsilon^u}\left(1 - Q(R,L)\varepsilon_R^n\left(\frac{C-B-L}{R}\right)\right)$$

$$\text{(4b)} \qquad \frac{C-B-L}{L} = \frac{1}{\varepsilon^d}\left(1 - Q(L,R)\varepsilon_L^n\left(\frac{R-A-C}{L}\right)\right)$$

where $\varepsilon^u = -(\partial Q^u / \partial R)/(Q^u / R)$, $\qquad Q^u = Q(L,R)/(1-Q(L,R)Q(R,L))$,

$\varepsilon^d = (\partial Q^d / \partial L)/(Q^d / L)$, $\qquad Q^d = Q(R,L)/(1-Q(L,R)Q(R,L))$,

$\varepsilon_R^n = -(\partial Q^n / \partial R)/(Q^n / R)$, $\qquad Q^n = Q(L,R)Q(R,L)/(1-Q(L,R)Q(R,L))$,

$\varepsilon_L^n = (\partial Q^n / \partial L)/(Q^n / L)$.

Here Q^u, Q^d and Q^n are aggregate discount factors applicable to certain initial situations and future decisions. Q^u applies to all future entry decisions, starting with a firm at time of exit. Likewise, Q^d applies to future exit decisions, starting with a firm at entry. Q^n applies to future entry decisions for a firm right after entry, as well as to exit decisions for a firm right after exit.

Eq. (4a) resembles the investment rule of Dixit et al. (1999). That paper discusses the standard problem of when to invest a constant C to obtain a stochastic benefit V. The optimal decision is to invest when a certain markup $(V-C)/V=1/\varepsilon$ is reached, where $\varepsilon=-(\partial D/\partial V)/(D/V)$, and $D=D(V_0,V)$ is the expected discount factor when the benefit is to move from V_0 to a fixed $V \geq V_0$.

Returning to our model, and disregarding the large parenthesis on the right-hand side, Eq. (4a) shows how much the trigger price R should exceed A+C, the total net present cost if entry was not affected by exit. While Dixit et al. (1999) consider a single investment and use the elasticity ε of the ordinary discount factor, we need the elasticity ε^u of an aggregate discount factor to account for the infinite time horizon. The parenthesis is caused by the exit option, and reduces the markup. The adjustment is small if L is far away at time of entry – i.e., if Q(R,L) is close to zero. The adjustment also depends on the aggregate elasticity ε_R^n and the relative size of the exit markup. The gain from the exit option vanishes if the firm could behave optimally at entry but for some reason must follow a Marshallian exit rule, leaving when L=C–B. Then Eq. (4a) simplifies to a perpetual version of the investment rule in Dixit et al. (1999).

The interpretation of Eq. (4b) is similar: the firm does not exit when L=C–B, but waits for a lower price according to a markup rule. Here the parenthesis vanishes if the firm could behave optimally at exit but had to follow a Marshallian entry rule, investing when R=A+C.

The model we have developed applies to a broad class of Ito processes, not only the geometric Brownian motion. We just need to find the discount factors and elasticities associated with the chosen stochastic process, and then solve Eq. (4a,b). Nonetheless, the geometric Brownian motion is of special interest due to tractability and widespread use. A geometric Brownian price with drift μ ($<\rho$) and volatility σ implies the following discount factor function:[6]

(5) $$Q(P_1, P_2) = \begin{cases} (P_1 / P_2)^{\beta}, & P_1 \le P_2 \\ (P_2 / P_1)^{\alpha}, & P_1 \ge P_2 \end{cases} .$$

Here β (>1) is the positive root and α (>0) is the magnitude of the negative root of the equation

(6) $$\tfrac{1}{2}\sigma^2 x^2 + (\mu - \tfrac{1}{2}\sigma^2)x - \rho = 0 .$$

We will use several discount factors based on Eq. (5), but in this section the only interesting ones are $Q(L,R)=(L/R)^{\beta}$ and $Q(R,L)=(L/R)^{\alpha}$. By inserting these formulas into Eq. (4a,b) we get two equations that can also be derived from Dixit (1989, eqs. 12-15). The aggregate elasticities ε^u, ε^d,

[6] As shown by Dixit et al. (1999), the expected discount factor going from P_1 to $P_2 \ge P_1$ can be found by solving the differential equation $\tfrac{1}{2}g^2(P_1)Q''(P_1,P_2)+f(P_1)Q'(P_1,P_2)-\rho Q(P_1,P_2)=0$, where derivatives are with respect to the first argument. Two boundary conditions apply. First, $Q(P_2,P_2)=1$, and, second, $Q(P_1,P_2)\to 0$ when P_2-P_1 becomes large. Setting $f(P_1)=\mu P_1$ and $g(P_1)=\sigma P_1$, the upper part of Eq. (5) follows easily. By replacing P_2-P_1 with P_1-P_2 in the second boundary condition, the lower part follows similarly.

ε_R^n and ε_L^n are stated in Appendix A, which also shows that the optimal R is lower than the corresponding markup price of the standard investment model studied by Dixit et al. (1999).

It is well known how L and R depend on exogenous variables, so that issue does not have to be pursued here. We shall rather demonstrate the sensitivity of the decision policy, an issue the smooth pasting approach is not as well suited to address. W_0 is the expected profit arising from any feasible pair of trigger prices, so it also shows the implications of missing the right ones (– a likely outcome in practice!). Fig. 1 plots W_0 as a function of L and R, suppressing regions where $W_0<0$ or L>R. The numbers are taken from Dixit (1988), which again draws on Mossin (1968).

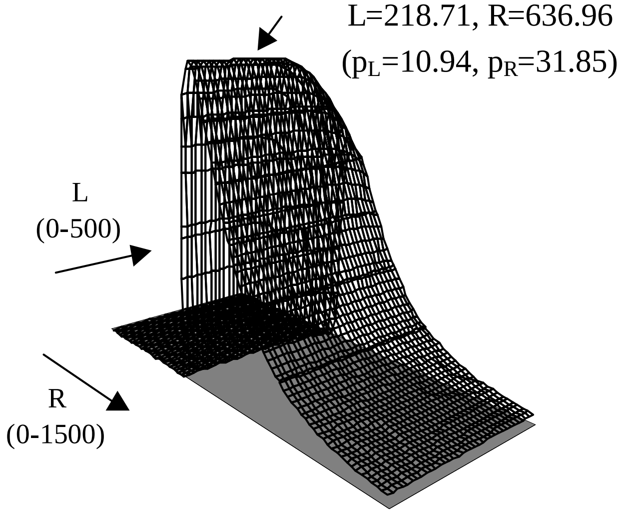

Fig. 1. The profit function W_0 with low volatility (ρ=0.05, μ=0, σ=0.05, A=144, B=144, c=20).

In this case W_0 is very sensitive to the entry trigger price. The exit trigger price does not matter much as long as it is fairly low. The reason is that exit rarely applies with such low uncertainty and high switching costs. Reducing the exit trigger price below the optimal level is therefore not very harmful. (W_0 is proportional to P_0^β, so the initial price determines the vertical scale. The closer P_0 is to R, the higher is the value of the firm; i.e., the maximum of W_0.)

The importance of finding the correct trigger prices changes by increasing the volatility to σ=0.2 as in Fig. 2. Now the expected profit is less sensitive to deviations from the optimal R and L. This is reasonable, since the cost of missing optimum is essentially a time cost. When the volatility is high and the firm picks its trigger prices fairly well, it usually does not take long to reach the optimal prices anyway. Then the cost of failing cannot be high. (The vertical scale is different in Fig. 1 and Fig. 2. The value of the firm is much higher in Fig. 2, where the volatility is higher.)

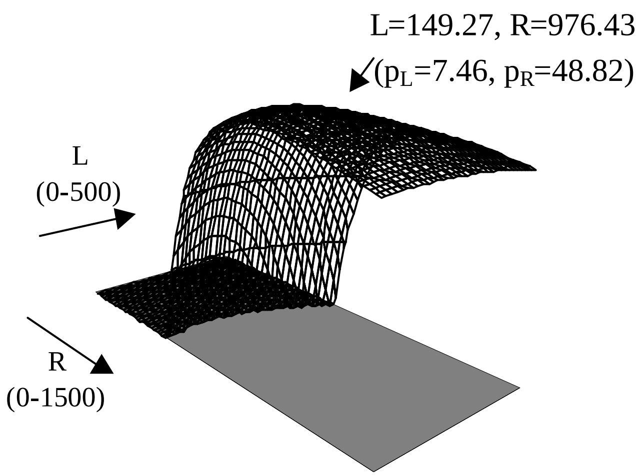

Fig. 2. The profit function W_0 with high volatility (ρ=0.05, μ=0, σ=0.20, A=144, B=144, c=20).

Thinking of many practical cases, the switching costs are high in Fig. 1 and Fig. 2, and switching does not occur frequently. In Fig. 3, both switching costs have been reduced from 144 to 1. That brings L and R closer, and W_0 becomes more sensitive to changes in both trigger prices close to optimum. Loosely speaking, however, the entry trigger price still appears to be more important.

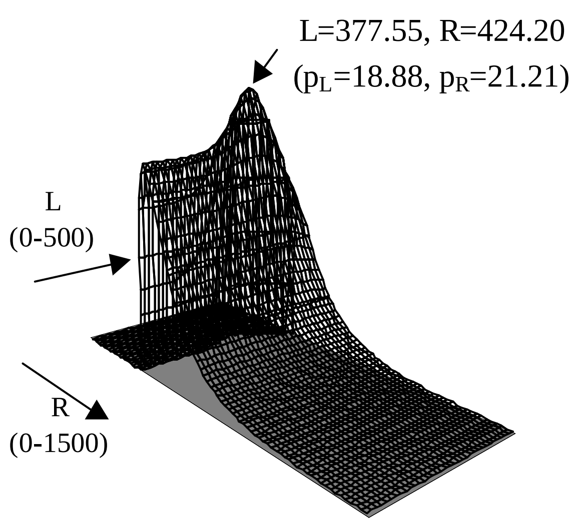

Fig. 3. The profit function W_0 with low switching costs (ρ=0.05, μ=0, σ=0.05, A=1, B=1, c=20).

The three figures above illustrate that W_0 is well defined and yields a unique maximum. This could also be shown by a formal analysis, which we omit. Note also that the profit function is typically quite smooth, and it may often be easier to solve models of this kind numerically by maximizing W_0 directly instead of going via first-order conditions like Eq. (4a,b). For most of the models below we do not discuss the first-order conditions. They are usually lengthy, and not needed for the numerical analysis. Instead we focus on deriving the profit functions.

3. Endogenous production costs

One common trade-off in economic decision-making is the one between fixed and marginal costs. By long-term capital investments, labor costs and other short-term variable costs can be reduced. Suppose that our firm is faced with such a trade-off. If the entry cost A is increased, the production cost C will decrease. When implementing this assumption, we must decide whether the effect of increasing the entry cost lasts for more than one production period. This could be the case, but we do not consider it here. Instead we assume that the firm each time entry takes place must decide the size of the entry cost, regardless of previous investments. Then the only candidate for an optimal decision is the usual one: the firm will enter and exit at two certain trigger prices. Hence, the firm will have to optimize L, R and A.

To formalize the model, we assume $C=C(A)$, where $dC/dA<0$ and $d^2C/dA^2>0$. Furthermore, set $W_0=W_0(L,R,A)$ in Eq. (3), and maximize W_0 with respect to L, R and A. The two first-order conditions in Eq. (4a,b) apply once again, except that C is now a function of A. The third condition, $\partial W_0/\partial A=0$, implies

(7) $$dC/dA = \frac{-1}{1-Q(R,L)} .$$

As long as $0<Q(R,L)<1$, we have $dC/dA<-1$. Thus the first investment should not be pushed so far that the last dollar saves one dollar of production costs over infinite time. The reason is that exit may be optimal some day in the future, and then the gain from the investment will be lost.

Other aspects of this model can be addressed by studying how R and L are influenced by C(A) in specific cases. Table 1 shows one set of results based on the cost function $C(A)=(1+h/A)/\rho$, where h>0. The price is assumed to be geometric Brownian with 20 percent volatility and no drift, the discount rate is 5 percent, and there is no exit cost.

h	p_L	p_R	A	c	A/A_0
0.00	1.000	1.000	0.000	1.000	(0.000)
0.01	0.859	1.283	0.273	1.037	0.610
0.10	0.836	1.548	1.010	1.099	0.714
1.00	0.884	2.194	3.609	1.277	0.807
2.00	0.930	2.552	5.251	1.381	0.830
5.00	1.033	3.238	8.565	1.584	0.857
10.00	1.157	3.996	12.345	1.810	0.873
20.00	1.337	5.054	17.728	2.128	0.886
50.00	1.701	7.136	28.461	2.757	0.900
100.00	2.116	9.471	40.591	3.464	0.908

Table 1. Optimal investment with endogenous costs (ρ=0.05, μ=0, σ=0.2, B=0, c=1+h/A).

The limiting case h=0 implies c=1 and A=0. Then no investment is needed, and $p_L=p_R=1$. The larger h is, the more investment is needed to reduce the production cost. This makes the entry trigger price increasing in h. The exit trigger price decreases in h when h is low, because the irreversibility effect of increased entry costs dominates the marginal effect of increased production costs. This is reversed when h is sufficiently high, but the amount of hysteresis, measured by the ratio p_R/p_L, increases in h all the way. The Marshallian investment A_0 is given by dC/dA=−1, so the last column shows how much the optimal investment decreases due to the exit option. The lower h is, the closer are the trigger prices, and the more switching there is. That makes the error from not considering the exit option decreasing in h.

4. Restricted switching

The standard entry-exit model allows for an infinite number of switches. This may be unrealistic, at least without increasing the switching costs. For example, frequent switching could harm to the equipment in use. Ekern (1993) therefore restricts to a fixed maximum number of switches. Let us show briefly how this modification can be handled by the markup approach.

The stationary character of the model disappears by restricting the switching options, so now R and L will depend on the number of switches still available. If entry and exit can take place N times each, we get a 2N-dimensional profit function

(8) $$W_0(L_1,...,L_N,R_1,...,R_N)=\sum_{k=1}^{N}\left\{Q_k^{entry}\left(R_k-A-C\right)+Q_k^{exit}\left(C-B-L_k\right)\right\},$$

where

$$Q_k^{entry}=\prod_{i=1}^{k}Q(R_{i-1},L_{i-1})Q(L_{i-1},R_i)\,,\qquad Q_k^{exit}=\prod_{i=1}^{k}Q(L_{i-1},R_i)Q(R_i,L_i),\qquad R_0\equiv L_0\equiv P_0\,.$$

The aggregate discount factors Q_k^{entry} and Q_k^{exit}, k=1,…,N, are products of ordinary discount factors up until the decision to enter or exit for the k'th time. The argument for this description of W_0 follows from section 2. The optimal trigger prices are given by 2N first-order conditions. They are usually complicated, but recursive formulas can be worked out. In the geometric Brownian case, Ekern (1993) finds that the entry trigger price decreases and the exit trigger price increases as functions of the number of switching options still available. This is in line with our markup discussion in section 2.

5. Diminishing production capacity

Brekke and Øksendal (1994) assume that the production capacity is decreasing in the accumulated time of production. As an example, the resource stock of a mine is limited, and the revenue from exploiting the mine typically decreases over time. Then the flow of revenue from production is $qp\equiv p'$, where q is the current capacity and p is the current price. The price is assumed to geometric Brownian. Since the capacity is fixed when the firm is not producing, the initial revenue from starting production will be geometric Brownian with the same drift (μ) and volatility (σ) as the price process. The price process also applies during production, but now the capacity shrinks. Assuming a constant rate of capacity reduction, $\lambda>0$, we have $dq=-\lambda dt$. From Ito's lemma the actual revenue flow during production is then still geometric Brownian with volatility σ, but the drift is $\mu-\lambda\equiv\mu'$, which is lower than during lay-up. Since all income shows up in the flow of revenue, it follows that the income of the firm depends on two state-dependent geometric Brownian motions. During lay-up, the drift of the potential revenue process is μ, while the drift of the actual revenue process during production is μ'. Moreover, the future of the firm in a given state looks the same for any fixed p'. Thus the optimal policy will be to enter when p' hits a certain value from below, and to exit when a certain lower value is hit from above.

To describe this model by the markup approach we must first adjust the net present price for diminishing production capacity. This yields $p'/(\rho-\mu')$ for any net present price $p/(\rho-\mu)$ in the standard model. Hence, $p_L/(\rho-\mu)$ and $p_R/(\rho-\mu)$ are replaced by $p'_L/(\rho-\mu')\equiv L'$ and $p'_R/(\rho-\mu')\equiv R'$, where p'_L and p'_R are investment triggers in terms of p'. The discount factor that applies during production must also account for capacity reductions. This yields $Q(R',L')=(L'/R')^{\alpha'}$, where α' is the magnitude of the negative root of Eq. (6) after replacing μ with μ' in that equation. The potential revenue process during lay-up coincides with the one in section 2, so in the idle state the usual discount factor applies, $Q(L,R)=Q(L',R')=(L'/R')^{\beta}$.

With the revised interpretations, Eq. (3) still contains the expected value of the firm, and by replacing α, L and R with α', L' and R' in the geometric Brownian version of Eq. (4a,b), we end up with the solution in Brekke and Øksendal (1994, p. 1032).

6. Investment lags

Investment has taken place instantaneously in all models up until now. The real world is different. Bar-Ilan and Strange (1996) include a deterministic lag between the decision to enter is made and revenues and costs from production start to flow. The entry decision is irreversible, but to avoid the investment lag from affecting net present costs when studying its implications, one can assume the entry cost to be incurred at the end of the lag.

Since no revenue or cost applies during the investment lag, there is no cost of postponing exit decisions until the end of the lag even if the price gets very low. However, there is a gain from waiting to see whether the price rises again. This implies that exit during the lag cannot be optimal. If the firm decides to exit later on, and enters once again, a new investment lag applies.

The investment lag complicates the expected profit function of the markup approach slightly, but it complicates the smooth pasting approach much more. Bar-Ilan and Strange need an intricate procedure, and must even solve a partial differential equation to derive their model. We can set up the expected profit function of the firm directly, accounting for the delay of revenues and costs, and the decision at the end of the investment lag. Having chosen L and R in advance, the decision of the firm will be to continue production if the price at the end of the investment lag exceeds L, and otherwise to exit immediately.

As in section 2, suppose for a moment that the current price is L, and that the firm is idle. The profit function, explained below, becomes:

$$(9) \qquad W_L(L,R) = Q(L,R)e^{-\rho\tau} \times \left(R_\tau - A - C + Q^+(R,L)(C - B - L + W_L) + p^-(R,L)(C-B) - L^-(R,L) + Q^-(R,L)W_L\right)$$

where

$$Q^+(R,L) = \int_L^\infty Q(x,L)d\Omega(x), \qquad R_\tau = E\left[P_\tau \middle| P_0 = R\right], \qquad p^-(R,L) = p(P_\tau < L \middle| P_0 = R),$$

$$Q^-(R,L) = \int_{-\infty}^L Q(x,L)d\Omega(x), \qquad L^-(R,L) = \int_{-\infty}^L x d\Omega(x) .$$

The firm will not enter before the trigger price R is reached, so all future benefits and costs are discounted by the factor Q(L,R). Then the decision is made, but no revenue or cost shows up until τ years later. The deterministic discount factor $e^{-\rho\tau}$ takes care of the delay. The expected net present revenue from infinite production from then on is R_τ, and the entry cost A and the production cost C apply as usual. This explains the terms R_τ–A–C. The rest of the parenthesis contains the expected value of future decisions. The decisions will depend on the price movements during the investment lag. If the price at the end of the lag exceeds L, the firm will remain active and wait for L to be reached from above before exit takes place. When the latter event occurs, the net gain is C–B–L+W_L as in section 2. However, the (downward) discount factor must be modified, since the price could start from any level exceeding L. The adjusted discount factor, Q^+(R,L), is found by integrating over the range of possible start values, using the distribution function Ω associated with the diffusion process. Thus, Ω is the distribution of the price at time τ given that the price at time zero is R.

The firm exits immediately if the price at the end of the investment lag is below L. That occurs with probability p^-(R,L), and in such cases the firm saves the production cost C but has to pay the exit cost B. This explains the term $p^{-}(R,L)(C-B)$. The expected loss of revenue, $L^{-}(R,L)$, is found by integrating over all prices lower than L, using the same distribution as above. Since the price at time of exit is some uncertain value lower than L, the recursive element W_L must also be discounted. To account for the motion up to L we use the discount factor $Q^{-}(R,L)$, which appears by integrating the (upward) discount factor Q(x,L) over applicable prices at the end of the lag. After re-arranging Eq. (9) and setting $Q(P_0,L)W_L(L,R) \equiv W_0(L,R)$, the value of the firm becomes:

$$(10) \qquad W_0(L,R) = \frac{Q(P_0,R)e^{-\rho\tau}\left(R_\tau - A - C + Q^+(R,L)(C-B-L) + p^-(R,L)(C-B) - L^-(R,L)\right)}{1 - Q(L,R)e^{-\rho\tau}\left(Q^+(R,L) + Q^-(R,L)\right)}$$

This function is more complicated than the one in Eq. (3), but the models merge as $\tau \to 0$, since then we have $e^{-\rho\tau} \to 1$, $R_\tau \to R$, $Q^+(R,L) \to Q(R,L)$, $p^-(R,L) \to 0$, $L^-(R,L) \to 0$ and $Q^-(R,L) \to 0$.

W_0 as given by Eq. (10) can be maximized with respect to L and R for a number of price processes, but closed-form solutions can only be obtained in simple cases. If the price is geometric Brownian, the following components can be derived (see Appendix B):

$$(11) \qquad Q^+(R,L) = (L/R)^\alpha e^{\rho\tau}\left[1 - \Phi(u + \alpha\sigma_\tau)\right], \qquad R_\tau = R e^{\mu\tau}, \qquad p^-(R,L) = \Phi(u),$$

$$Q^-(R,L) = (R/L)^\beta e^{\rho\tau}\Phi(u - \beta\sigma_\tau), \qquad L^-(R,L) = R e^{\mu\tau}\Phi(u - \sigma_\tau),$$

where $u = \dfrac{\ln(L/R) - (\mu - \frac{1}{2}\sigma^2)\tau}{\sigma_\tau}$ and $\sigma_\tau = \sigma\sqrt{\tau}$.

$\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Its derivative is the standard normal density function, so the partial derivatives of W_0 are well defined, and closed-form first-order conditions like Eq. (4a,b) can be worked out.

As argued by Bar-Ilan and Strange (1996, 1998), investment lags can lead to preemptive investment. Early investment acts like an insurance against a profit loss that could otherwise arise while the firm is waiting for the investment to be completed. In this model the investment lag can have great implications, sometimes implying that uncertainty hastens investment. Bar-Ilan and Strange claim that this unconventional result can occur even for moderate levels of uncertainty. Unfortunately, a technical error infects all of their numerical results, and although their qualitative reasoning is correct, the numerical results are misleading.[7] Table 2 gives correct results under the same assumptions as in Bar-Ilan and Strange (1996, p. 616, Table 2).

	$\tau = 0$		$\tau = 6$	
σ^2	P_L	P_R	P_L	P_R
0.00	1.000	1.025	1.000	1.025
0.01	0.834	1.243	0.793	1.146
0.02	0.795	1.312	0.736	1.151
0.03	0.770	1.362	0.697	1.149
0.04	0.751	1.405	0.666	1.145
0.05	0.735	1.442	0.640	1.140
0.10	0.682	1.586	0.551	1.112
0.20	0.623	1.791	0.450	1.072
0.30	0.587	1.953	0.388	1.048
0.40	0.560	2.094	0.342	1.036
0.50	0.539	2.221	0.308	1.031
0.60	0.522	2.338	0.280	1.033
0.80	0.495	2.554	0.237	1.049
1.00	0.474	2.753	0.206	1.078

Table 2. Trigger prices with investment lags (ρ=0.025, μ=0, A=1, B=0, c=1).

When τ=6, the entry trigger price as a function of σ first increases and then decreases, before eventually increasing steadily. Bar-Ilan and Strange found a similar pattern, but their minimum trigger price (0.960) was far below the Marshallian trigger price (1.025), and it occurred for a moderate level of uncertainty (σ^2=0.016). Table 2 shows that much more uncertainty, σ^2=0.5, is needed to reach the corresponding minimum (1.031). Moreover, the correct minimum is local, above the deterministic trigger price. Thus any stochastic model based on these data yields a higher entry trigger price than the deterministic model. Other experiments show that the investment lag must be more than 12 years to yield a lower entry trigger price than the Marshallian one for σ^2=0.016. The entry cost is already low, equal to the annual production cost, and investment lags exceeding 12 years are not likely in such cases. The lag would have to be

[7] The authors mix up the volatility of the geometric Brownian motion (σ) with the standard deviation of the price after the lag (σ_τ). The error can be observed by comparing arguments of $\Phi(\cdot)$ in Eq. (11) with similar arguments in Bar-Ilan and Strange (1996, p. 614, eqs. 22-25).

increased even more to obtain similar results with a higher entry cost. Nor does reasonable changes in other variables seem to help much. For example, a positive exit cost would also increase the entry trigger price.

To conclude, increased uncertainty may hasten investment when there is a lag. A lower entry trigger price under uncertainty than under certainty is also possible, but combinations of parameters implying such outcomes are less realistic than argued by Bar-Ilan and Strange (1996).

7. Construction and scrapping

Dixit (1988) incorporates the entry-exit model in a larger model where mothballing is preceded by construction and succeeded by scrapping. Dixit and Pindyck (1994, pp. 229-241) present the same model and apply it to shipping. In section 8 we allow for investment lags in this so-called entry-exit-scrapping model. To avoid too many new issues at the same time, we show here how the markup approach can be used to derive the model without lags.

The model yields four trigger prices (in descending order): H (construction), R (entry), L (mothballing) and S (scrapping). There is no reason to build before the price is high enough for immediate entry, so the first entry will occur right after construction at price H, while the next ones occur at price R. The exit policy is more complicated as mothballing may be irrelevant. For scrapping to be of interest, we now need a positive mothballing cost, $m>0$. The cost of mothballing into infinity becomes $m/\rho \equiv M$, and $M<C$ is required for mothballing to be of interest. However, mothballing may be too costly even if $M<C$ because of the switching costs. If that is the case, the model simplifies to the standard model in section 2; the sum of the construction cost and the entry cost will make up the entry cost of the standard model, and the scrapping cost will make up the exit cost. We return to such outcomes in section 8, but in this section we proceed under the assumption that mothballing does apply. In that case we have $S<L$.

The value of the firm can be expressed by the following sequence of functions:

(12a) $W_0 = Q(P_0, S)W_S$

(12b) $W_S = Q(S, H)\left(H - I - A - C + W_H\right)$

(12c) $W_H = Q(H, L)(C - B - L - M + W_L)$

(12d) $W_L = Q_c(L, R, S)(R - A - C + M + W_R) + Q_c(L, S, R)(M - K + W_S)$

(12e) $W_R = Q(R, L)(C - B - L - M + W_L)$

The cost components and discount factors in Eq. (12a-e) are illustrated in Fig 4.

Decision	Price	Investment	Value	Price movements and discount factors
Construction	H	I	W_H	
Entry	R	A	W_R	
Exit	L	B	W_L	
Scrapping	S	K	W_S	
Initial situation	P_0	-	W_0	

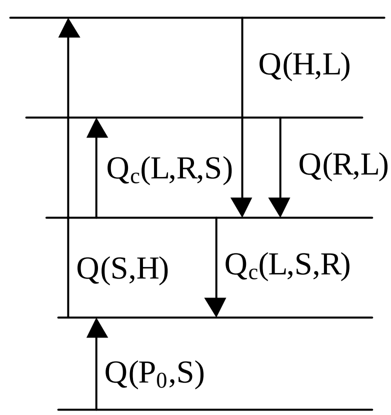

Fig. 4. The entry-exit-scrapping model.

Eq. (12a) says that the initial value of the firm, W_0, equals the discounted value of the firm after the scrapping trigger price S has been reached, using the expected discount factor going from P_0 to S. This is true as no investment occurs between P_0 and S.

Actually, no investment occurs construction is triggered at price H. Therefore W_S in Eq. (12b) follows from discounting further benefits and costs by Q(S,H). At that time the expected revenue from production through infinity is H, and three costs apply: the construction cost I, the entry cost A, and the production cost C. The value of future switching options, W_H, must also be included.[8]

[8] This setup deviates slightly from Dixit (1988), where the construction cost includes the cost of first entry. We define the construction cost differently to make the model consistent with the one in section 8.

Exit at price L is the next event to take place, so the value of the firm after construction and first entry, W_H in Eq. (12c), is found by discounting future benefits and costs by the factor Q(H,L). When exit occurs, the production cost C is saved, the exit cost B is incurred, and the revenue L would also be lost if mothballing should go on forever. In addition, we must subtract the mothballing cost M, and include the value of further options, W_L.

Like the previous value functions, W_L in Eq. (12d) is determined by future expected events. The firm will re-enter if the entry trigger price R is reached before the scrapping trigger price S. Scrapping takes place in the opposite case. If the entry trigger price is hit first, the revenue R and the costs A and C apply as usual. The mothballing cost M is saved, and the term W_R account for the value of remaining options. The net gain, $R–A–C+M+W_R$, is multiplied by the conditional discount factor $Q_c(L,R,S)$, which applies to cases where the price moves from L to R without hitting S. In the opposite case, when S is hit first, the mothballing cost M is saved, the scrapping cost K is incurred, and the value of the firm right afterwards is W_S. The net gain, $M–K+W_S$, is multiplied by the conditional discount factor going from L to S without hitting R; i.e., $Q_c(L,S,R)$.

W_R as given by Eq. (12e) equals the value of the firm right after entry in the section 2 model, except for the mothballing cost that needs to be subtracted.

By re-arranging Eq. (12a-e), also using that $Q(P_0,S)Q(S,H)=Q(P_0,H)$, we find:

(13) $W_0(S,L,R,H) =$

$$\frac{Q(P_0,H)\begin{pmatrix} H-I-A-C+Q(H,L)\left(C-B-L-M\right)+Q(H,L)Q_c(L,R,S)(R-A-C+M) \\ +Q(R,L)Q_c(L,R,S)\left(C-H+I+A\right)+Q(H,L)Q_c(L,S,R)(M-K) \end{pmatrix}}{1-Q_c(L,R,S)Q(R,L)-Q_c(L,S,R)Q(S,H)Q(H,L)}$$

The optimal policy is found by maximizing W_0 with respect to S, L, R and H. In the geometric Brownian case, the conditional discount factor function included in W_0 becomes:[9]

[9] The conditional discount factor function is found by the same differential equation as the ordinary discount factor discussed earlier. Two boundary conditions apply: $Q_c(P_1,P_1,P_2)=1$ and $Q(P_2,P_1,P_2)=0$. The solution in the geometric Brownian case, given by Eq. (14), follows easily.

$$(14)\qquad Q_c(P_1,P_2,P_3)=\begin{cases}\dfrac{(P_1/P_2)^{\beta}-(P_3/P_1)^{\alpha}(P_3/P_2)^{\beta}}{1-(P_3/P_2)^{\alpha+\beta}}, & P_3\le P_1\le P_2\\[2ex] \dfrac{(P_2/P_1)^{\alpha}-(P_2/P_3)^{\alpha}(P_1/P_3)^{\beta}}{1-(P_2/P_3)^{\alpha+\beta}}, & P_2\le P_1\le P_3\end{cases}$$

The upper part of Eq. (14) implies $Q_c(L,R,S)=[(L/R)^{\beta}-(S/L)^{\alpha}(S/R)^{\beta}]/[1-(S/R)^{(\alpha+\beta)}]$. The lower part implies $Q_c(L,S,R)=[(S/L)^{\alpha}-(S/R)^{\alpha}(L/R)^{\beta}]/[1-(S/R)^{(\alpha+\beta)}]$. Inserting this into Eq. (13) and maximizing, it can be shown that the optimal policy coincides with the solution in Dixit (1988) or Dixit and Pindyck (1994, pp. 230-235). As shown in the latter references, the optimal entry and exit trigger prices, R and L, can be found from the model in section 2. We just have to replace the production cost C of that model with C–M; i.e., the difference between the production cost and the mothballing cost. Thus L and R do not depend on construction and scrapping costs. This simplifies the numerical analysis.

8. Construction, scrapping and investment lags

Dixit (1988) motivates the entry-exit-scrapping model with shipping. In that industry the most significant investment lags are related to newbuilding whereas operating ships can be mothballed quickly. The objective of this section is to model and discuss such investment lags.

We make the same assumptions about the investment lags as in section 6, except that the lags apply to construction instead of entry. Thus construction is irreversible and takes a fixed amount of time, τ. If scrapping occurs, and a decision to build again is made some time later, a new investment lag is required. The development of the price during construction will decide whether the new ship – or whatever production facility the model represents – should go into production, be laid-up or scrapped immediately. This means that the dynamics could differ highly from Dixit's model, where construction is always succeeded by immediate entry. As in the previous section, we find the expected value of the firm from a sequence of functions:

$$(15a)\qquad W_0=Q(P_0,S)W_S$$

$$(15b)\qquad W_S=Q(S,H)W_H$$

(15c) $W_H = e^{-\rho\tau} \times$

$$\begin{pmatrix} -I - M + R^{+}(R,H) + p^{+}(R,H)(M - A - C) + Q^{+}(H,R,L)(C - B - L - M + W_L) + \\ Q_c^{+}(H,R,S)(R - A - C + M + W_R) + Q_c^{-}(H,S,R)(M - K + W_S) + p^{-}(S,H)(M - K) + Q^{-}(H,S)W_S \end{pmatrix}$$

(15d) $W_L = Q_c(L,R,S)(R - A - C + M + W_R) + Q_c(L,S,R)(M - K + W_S)$

(15e) $W_R = Q(R,L)(C - B - L - M + W_L)$

where $Q^{+}(H,R,L) = \int_R^{\infty} Q(x,L)d\Omega(x)$, $\quad Q^{-}(H,S) = \int_{-\infty}^{S} Q(x,S)d\Omega(x)$,

$Q_c^{+}(H,R,S) = \int_S^{R} Q_c(x,R,S)d\Omega(x)$, $\quad Q_c^{-}(H,S,R) = \int_S^{R} Q(x,S,R)d\Omega(x)$,

$p^{+}(R,H) = p(P_\tau > R | P_0 = H)$, $\quad p^{-}(S,H) = p(P_\tau < S | P_0 = H)$,

$R^{+}(R,H) = \int_R^{\infty} x d\Omega(x)$.

The only part of Eq. (15a-e) that needs to be explained is (15c). The decision to build is made at price H, and all future costs and benefits are discounted further by $e^{-\rho\tau}$ due to the investment lag. Then the construction cost I is incurred. We assume that the ship continues in a laid-up position unless another decision is made, so the mothballing cost M is also subtracted. The rest of the first line in parenthesis captures cases for which the price after the investment lag ends up above R, leading to immediate entry. The expected net present revenue related to this event, $R^{+}(R,H)$, is found by integrating the price after the lag over the applicable range, R to ∞. In such cases the mothballing cost M is saved, and the entry cost A and the production cost C are incurred. All this happens with probability $p^{+}(R,H)$, so the expected cost gain is $p^{+}(R,H)(M–A–C)$. The firm continues production until the price gets down to L. Then the production cost C is saved, while the exit cost B, the mothballing cost M, and the loss of revenue L are incurred. The remaining option value W_L is added for the usual reasons. All this yields the net benefit $C–B–L–M+W_L$. The adjusted discount factor, $Q^{+}(H,R,L)$, is found by integrating the (downward) discount factor Q(x,L) over the same interval as above.

Mothballing continues if the price at the end of the lag is between S and R. Then entry occurs if R is hit before S, whereas scrapping occurs if S is hit first. The net benefit in the former case is $R-A-C+M+W_R$, similar to that of section 7. The discount factor, $Q_c^+(H,R,S)$, is found by integrating the (upward) conditional discount factor $Q_c(x,R,S)$ over the interval of start prices; i.e., from S to R. Likewise, the discount factor $Q_c^-(H,S,R)$ applying to cases for which S is hit first, is found by integrating the (downward) conditional discount factor $Q_c(x,S,R)$. The interpretation of the net benefit in the latter case, $M-K+W_S$, is familiar by now.

The remaining terms cover immediate scrapping, which occurs with probability $p^-(S,H)$. In such cases the mothballing cost M is saved and the scrapping cost K is paid, so the expected cost gain is $p^-(S,H)(M-K)$. The factor $Q^-(H,S)$ measures expected discounting up to price S and option value W_S. This is similar to the factor in the end of Eq. (9), and needs no further explanation.

Eq. (15a-e) determines W_0, W_S, W_L, W_R and W_H as functions of S, L, R, H. The formula for W_0 is lengthy and yields no further insight, so we omit it. More importantly, its components in the geometric Brownian version used below are as follows (see Appendix B):

(16) $$Q^+(H,R,L) = (L/H)^\alpha e^{\rho\tau}\left(1-\Phi(u_R+\alpha\sigma_\tau)\right)$$

$$Q^-(H,S) = (H/S)^\beta e^{\rho\tau}\Phi(u_S-\beta\sigma_\tau)$$

$$Q_c^+(H,R,S) = \frac{e^{\rho\tau}}{1-(S/R)^{\alpha+\beta}}\times \left\{\left(\frac{H}{R}\right)^\beta[\Phi(u_R-\beta\sigma_\tau)-\Phi(u_S-\beta\sigma_\tau)]-\left(\frac{S}{H}\right)^\alpha\left(\frac{S}{R}\right)^\beta[\Phi(u_R+\alpha\sigma_\tau)-\Phi(u_S+\alpha\sigma_\tau)]\right\}$$

$$Q_c^-(H,S,R) = \frac{e^{\rho\tau}}{1-(S/R)^{\alpha+\beta}}\times \left\{\left(\frac{S}{H}\right)^\alpha[\Phi(u_R+\alpha\sigma_\tau)-\Phi(u_S+\alpha\sigma_\tau)]-\left(\frac{S}{R}\right)^\alpha\left(\frac{H}{R}\right)^\beta[\Phi(u_R-\beta\sigma_\tau)-\Phi(u_S-\beta\sigma_\tau)]\right\}$$

$$p^+(R,H) = 1-\Phi(u_R), \qquad p^-(S,H) = \Phi(u_S), \qquad R^+(R,H) = He^{\mu\tau}\left(1-\Phi(u_R-\sigma_\tau)\right)$$

where

$$u_R = \frac{\ln(R/H) - (\mu - \frac{1}{2}\sigma^2)\tau}{\sigma_\tau}, \qquad u_S = \frac{\ln(S/H) - (\mu - \frac{1}{2}\sigma^2)\tau}{\sigma_\tau}, \qquad \sigma_\tau = \sigma\sqrt{\tau}.$$

We start the numerical analysis by discussing the timing of decisions to build. As in section 6 we set τ=6 when there is a lag. Table 3 contains results with and without the investment lag for several construction costs. The mothballing cost is 10 percent of the production cost, and there is no scrapping cost. Other numbers are inspired by section 2. The results show that construction should start earlier when there is a lag. For the lowest construction cost, I=1, the construction trigger price with the lag (1.259) is more than 20 percent lower than the trigger price without it (1.598). The construction trigger price is also lower than the entry trigger price (1.302) even though the total cost of construction and entry (I+A=2) is twice as high as the entry cost (A=1). This implies that initial mothballing will often occur, but preemptive investment is still optimal due to the insurance aspect.

	τ = 0		τ = 6	
I	p_S	p_H	p_S	p_H
1	0.714	1.598	0.633	1.259
2	0.688	1.742	0.579	1.484
3	0.661	1.873	0.548	1.659
4	0.621	1.996	0.526	1.811
5	0.592	2.112	0.510	1.950
10	0.513	2.655	0.465	2.552
15	0.477	3.166	0.443	3.090
20	0.456	3.662	0.430	3.601

Table 3. Trigger prices with various construction costs (ρ=0.05, μ=0, σ=0.2, A=1, B=0, K=0, c=1, m=0.1, p_L=0.675, p_R=1.302).

For higher construction costs it is optimal to wait longer before a decision to build, and the lag does not affect the decision very much. When I=20, the construction trigger price with the lag (3.601) is less than 2 percent lower than the trigger price without the lag (3.662).

Observe also from Table 3 that the investment lag makes mothballing more interesting. When excluding the investment lag, mothballing does not apply for I=1 and I=2. When including the lag, mothballing applies with a good margin for I=2, and with a smaller margin for I=1.

The length of the investment lag affects the timing decision significantly. This is shown in Table 4, reporting results with similar assumptions as in Table 3 for various lags. The longer the lag is, the earlier a decision to build should be made. The reason is the same as discussed by Bar-Ilan and Strange (1996): the exit option makes the upper part of the future price distribution more important when there is a lag. Since this effect arises because early construction acts like insurance, it tends to be suppressed by increasing the construction cost. As in Table 3, a longer investment lag therefore has more implications for the timing of investment with low construction costs (I=2) than with high construction costs (I=20). Note also that investment lags of more than one year are needed for mothballing to apply when I=2.

	I=2		I=20	
τ	p_S	p_H	p_S	p_H
0	0.688	1.742	0.456	3.662
1	0.677	1.719	0.451	3.654
2	0.667	1.688	0.446	3.645
3	0.630	1.617	0.441	3.635
4	0.610	1.571	0.437	3.625
5	0.593	1.526	0.434	3.614
6	0.579	1.484	0.430	3.601
7	0.566	1.444	0.427	3.588
8	0.554	1.406	0.424	3.574
9	0.544	1.371	0.421	3.559
10	0.534	1.337	0.419	3.544

Table 4. Trigger prices with various investment lags (ρ=0.05, μ=0, σ=0.2, A=1, B=0, K=0, c=1, m=0.1, p_L=0.675, p_R=1.302).

We conclude this section with an empirical example. Dixit and Pindyck (1994, pp. 237-241) apply the entry-exit-scrapping model without lags to shipping, focusing on a medium-sized oil tanker. Mothballing has been used several times in this industry, especially during the oil crisis in the 1970s, and the analysis confirms that this can be optimal. However, it also shows that mothballing would not be viable if the mothballing cost were 40 percent higher than estimated.

Capital costs are high in shipping. The price of the oil tanker studied by Dixit and Pindyck is 9 times as high as the annual operating cost, which again is almost 9 times as high as the mothballing cost. Referring to the results above, one should not expect realistic investment lags to have much impact on decisions to build in such cases even if the estimated price volatility is quite high (σ=0.15). Based on Dixit and Pindyck's data, we find that a realistic lag, τ=3, decreases the entry trigger price with 1.2 percent. The scrapping trigger price decreases with 2.8 percent, while the critical cost where mothballing becomes irrelevant increases significantly. With the investment lag, the mothballing cost can be 50 percent higher than estimated before mothballing becomes irrelevant. Hence, mothballing oil tankers can indeed be optimal in times of crisis.

9. Equilibrium

This paper focuses on the firm level, and a formal analysis of equilibrium is beyond its scope. Nevertheless, it is appropriate to comment on how the models fit into a broader perspective. More generally, Leahy (1993) found that myopic behavior as in the entry-exit model is also optimal for a firm in an industry exposed to aggregate uncertainty. In other words, the optimal trigger prices coincide in partial equilibrium and monopoly. In equilibrium, the trigger prices represent lower and upper barriers for the price process. When R is reached, a sufficiently high rate of entry will prevent the price from rising further. Similarly, exit will prevent the price from decreasing below L. Construction and scrapping make this picture more complicated, but we do not pursue such issues. Instead we comment on another type of equilibrium of interest.

For all models it has been assumed that the price starts from a low fixed level, P_0. We have not considered how this situation arose; i.e., how the firm received the initial option to invest. A setup cost or patent cost is usually needed to establish the option. Dixit and Pindyck (1994, pp. 267-277) discuss such option investments in partial equilibrium with two-step investment – patenting and production – and a set of firm-specific demand processes. Sødal (2001) contains a related one-sector general equilibrium model based on the markup approach. Both models assume homogenous goods, but with each firm facing an independent demand process that takes off as soon as the firm has acquired its patent. The number of active firms is assumed to be so large that the law of large numbers applies. Then the equilibrium setting does not affect the optimal behavior of individual firms, but the price level will settle where the expected profit from acquiring a patent is zero. Steady state is characterized by a constant flow of new firms acquiring patents. A finite equilibrium number of active firms can be imposed by some death mechanism.

It seems easy to expand the models of this paper into equilibrium models of the kind just described. The expected profit function ahead of the patent investment, conditional on the initial price P_0, could be written as

$$W = -G + W_0 , \tag{17}$$

where G is a fixed patent cost, and W_0 is given by Eq. (3) or one of its extensions. The patent investment is irreversible and does not affect subsequent decisions. If immediate activation of the patent can be ruled out, the trigger prices can therefore be found as in previous sections.[10] The value of a patent is increasing in P_0, and as long as new firms are established when there is net profit from acquiring a patent, the rate of newcomers will move P_0 to a level where W=0.

As an example, the car industry is highly competitive even if various cars are facing specific demand curves. The Dixit-Stiglitz (1977) model of monopolistic competition, based on the combination of static preferences for variation and scale economies, is often applied to such industries. However, different cars could also be perfect substitutes but exposed to firm-specific demand shocks. In the real world it can be hard to say from where a markup originates; it could be from static preferences for variation, but it could also be from dynamic demand uncertainty leading to value from waiting. Assuming the latter for now, the firm needs to undertake research to develop a new car (patenting). A production facility must be set up at the optimal time (construction), and initiating production requires additional efforts (entry). Such investments take time, so investment lags may apply. Production costs are necessary to keep the production line open, and in times of low demand it may be necessary to mothball or switch to alternative uses (exit), until demand one day gets so low that the production line is abolished (scrapping).

Fluctuating demand is usually not entirely specific to the firm, but arises from combined aggregate and firm-specific shocks. Thus aggregate uncertainty and other important issues related to inventories etc., should be included to bring more realism to all this. Nevertheless, the

[10] Typically, it will not be optimal to activate patents right away in equilibrium if the patent cost is low, and the initial price is fixed as assumed here and in Sødal (2001). More realistically, the initial price could be uncertain as in Dixit and Pindyck (1994). In the latter case, immediate patent activation cannot be ruled out quite as easily.

framework we have described seems to capture many important aspects of investment both at the firm level and the industry level.

10. Final remarks

We have shown how to develop the standard entry-exit model and a number of extensions of this model, using a new markup approach to investment. The significance of investment lags in practice has also been discussed. Most extensions to the entry-exit model were treated separately, but many of them can be combined. Several other extensions could have been added. For example, more costs could be made endogenous by the same approach as in section 3. The price process could be made endogenous as in Dixit et al. (1999), and new types of state-dependent price processes could be included in the entry-exit-scrapping model based on ideas from section 5.

It remains to be seen how far the markup approach can be developed. All of our models share some basic properties: the life cycle of the firm can be described by a sequence of possible events and associated discount factors, and investment is to take place as soon as some trigger prices are hit within the various states. These properties apply in several contexts including phenomena like multiple investment lags, flexible scale of production and, more generally, settings with more compound options than the ones studied here.

References

Aitchison, J. and J.A.C. Brown, 1957, *The Lognormal Distribution*, Cambridge University Press.

Bar-Ilan, Avner and William C. Strange, 1996, “Investment Lags”, *American Economic Review* 86:3, 610-622.

Bar-Ilan, Avner and William C. Strange, 1998, “A Model of Sequential Investment”, *Journal of Economic Dynamics and Control* 22, 437-463.

Brekke, Kjell Arne and Bernt Øksendal, 1994, “Optimal Switching in an Economic Activity under Uncertainty”, *SIAM Journal on Control and Optimization* 32:4, 1021-1036.

Brennan, Michael J. and Eduardo S. Schwarz, 1985, “Evaluating Natural Resource Investments”, *Journal of Business* 58, 135-157.

Dixit, Avinash, 1988, "Optimal Lay-up and Scrapping Decisions", Unpublished, Princeton University, July.

Dixit, Avinash, 1989, "Entry and Exit Decisions under Uncertainty", *Journal of Political Economy* 97, 620-638.

Dixit, Avinash K. and Robert S. Pindyck, 1994, *Investment Under Uncertainty*, Princeton University Press.

Dixit, Avinash, Robert S. Pindyck and Sigbjørn Sødal, 1999, "A Markup Interpretation of Optimal Investment Rules", *The Economic Journal* 109:455, 179-189.

Dixit, Avinash and Joseph E. Stiglitz, 1977, "Monopolistic Competition and Optimum Product Diversity", *American Economic Review* 67, 297-308.

Ekern, Steinar, 1993, "Entry and Exit Decisions with Restricted Switching", *Working Paper 8/93*, The Norwegian School of Economics and Business Administration, Department of Finance and Management Science.

Leahy, John, 1993, "Investment in Competitive Equilibrium: The Optimality of Myopic Behaviour", *Quarterly Journal of Economics* 108, 1105-1133.

Mossin, Jan, 1968, "An Optimal Policy for Lay-up Decisions", *Swedish Journal of Economics* 70, 170-177.

Sødal, Sigbjørn, 1998, "A Simplified Exposition of Smooth Pasting", *Economics Letters* 58, 217-223.

Sødal, Sigbjørn, 2001, "Does Waiting Matter? A General Equilibrium Model with Sleeping Patents", forthcoming in Trigeorgis (ed.): (*title to be decided*), Oxford University Press.

Appendix A. The first-order conditions in the entry-exit model

Since $Q(P_0,R)=Q(P_0,L)Q(L,R)$, we can write W_0 in Eq. (3) as

$$\text{(A.1)} \quad W_0 = Q(P_0,L)\left(Q^u(R-A-C)+Q^n(C-B-L)\right),$$

where Q^u and Q^n are defined in Eq. (4a,b). The first-order condition $\partial W_0/\partial R=0$ implies

$$\text{(A.2)} \quad \frac{\partial Q^u}{\partial R}\left(R-A-C\right)+Q^u+\frac{\partial Q^n}{\partial R}\left(C-B-L\right)=0\,.$$

Eq. (4a) is obtained by re-arranging Eq. (A.2), using that $Q^n=Q(R,L)Q^u$, and inserting for ε^u and $\varepsilon_R{}^n$.

From definitions we have $1/(1–Q(L,R)Q(R,L))=Q^n+1$, so W_0 can also be written as

(A.3) $$W_0 = Q(P_0, R)\big((Q^n+1)(R-A-C)+Q^d(C-B-L)\big).$$

The first-order condition $\partial W_0/\partial L=0$ implies

(A.4) $$\frac{\partial Q^n}{\partial L}(R-A-C)+\frac{\partial Q^d}{\partial L}(C-B-L)-Q^d=0.$$

Eq. (4b) is obtained by re-arranging Eq. (A.4), using that $Q^n=Q(L,R)Q^d$, and inserting for ε^d and $\varepsilon_L{}^n$.

The geometric Brownian case yields the following elasticities:

(A.5) $$\varepsilon^u=\frac{\alpha(L/R)^{\alpha+\beta}+\beta}{1-(L/R)^{\alpha+\beta}},\qquad \varepsilon^d=\frac{\alpha+\beta(L/R)^{\alpha+\beta}}{1-(L/R)^{\alpha+\beta}},\qquad \varepsilon_R^n=\varepsilon_L^n=\frac{\alpha+\beta}{1-(L/R)^{\alpha+\beta}}\ .$$

After inserting this into Eq. (4a,b), it can be shown by some algebra that the optimal trigger prices coincide with those obtained by smooth pasting as in Dixit (1989, eqs. 12-15). Note also that $\varepsilon^u>\beta$, while the corresponding elasticity ε of the ordinary discount factor Q(L,R) equals β. Thus the optimal R is lower than the corresponding markup price of the standard investment model studied by Dixit et al. (1999).

Appendix B. Components of W_0 in models with investment lags

Below we determine the profit function in the section 6 model when the price is geometric Brownian; i.e., we find the components of W_0 in Eq. (11). We also comment on the components of W_0 in the section 8 model as given by Eq. (16); they can be determined in similar ways.

In section 6, the decision to enter is made when the price equals R. From the properties of the geometric Brownian motion it follows that the price after the investment lag, x, is distributed lognormally $\Lambda(x\,|\,\mu_\tau,\sigma_\tau^2)$, where $\mu_\tau=\ln R+(\mu-1/2\sigma^2)\tau$ and $\sigma_\tau^2=\sigma^2\tau$. Thus the logarithm of the price is distributed normally $N(\ln x\,|\,\mu_\tau,\sigma_\tau^2)$. For the lognormal distribution the following relations hold for any $\xi>0$ and j (see Aitchison and Brown, 1957):

(B.1) $$\int_0^{\xi} x^j d\Lambda(x|\mu_\tau,\sigma_\tau^2) = e^{j\mu_\tau+\frac{1}{2}j^2\sigma_\tau^2}\Lambda(\xi|\mu_\tau + j\sigma_\tau^2,\sigma_\tau^2)$$

(B.2) $$\int_{\xi}^{\infty} x^j d\Lambda(x|\mu_\tau,\sigma_\tau^2) = e^{j\mu_\tau+\frac{1}{2}j^2\sigma_\tau^2}\left(1-\Lambda(\xi|\mu_\tau + j\sigma_\tau^2,\sigma_\tau^2)\right).$$

Eq. (B.2) determines the first adjusted discount factor, $Q^+(R,L)$:

(B.3) $$\begin{aligned} Q^+(R,L) &= \int_L^{\infty} Q(x,L)d\Lambda(x|\mu_\tau,\sigma_\tau^2) = L^{\alpha}\int_L^{\infty} x^{-\alpha}d\Lambda(x|\mu_\tau,\sigma_\tau^2) \\ &= L^{\alpha}e^{-\alpha\mu_\tau+\frac{1}{2}\alpha^2\sigma_\tau^2}\left(1-\Lambda(L|\mu_\tau-\alpha\sigma_\tau^2,\sigma_\tau^2)\right) = (L/R)^{\alpha}e^{\rho\tau}\left(1-N(\ln L|\mu_\tau-\alpha\sigma_\tau^2,\sigma_\tau^2)\right) \end{aligned}$$

Here it has been used that $\frac{1}{2}\sigma^2\alpha^2-(\mu-\frac{1}{2}\sigma^2)\alpha=\rho$. Then $Q^+(R,L)$ in Eq. (11) follows from standard properties of the normal distribution. The relation $R_\tau=Re^{\mu\tau}$ is trivial from the properties of the geometric Brownian motion, and the probability $p^-(R,L)$ follows from Eq. (B.1). The next adjusted discount factor, $Q^-(R,L)$, is found as follows:

(B.4) $$\begin{aligned} Q^-(R,L) &= \int_0^{L} Q(x,L)d\Lambda(x|\mu_\tau,\sigma_\tau^2) = L^{-\beta}\int_0^{L} x^{\beta}d\Lambda(x|\mu_\tau,\sigma_\tau^2) \\ &= L^{\beta}e^{\beta\mu_\tau+\frac{1}{2}\beta^2\sigma_\tau^2}\Lambda(L|\mu_\tau+\beta\sigma_\tau^2,\sigma_\tau^2) = (R/L)^{\beta}e^{\rho\tau}N(\ln L|\mu_\tau+\beta\sigma_\tau^2,\sigma_\tau^2) \end{aligned}$$

Here we have used that $\frac{1}{2}\sigma^2\beta^2+(\mu-\frac{1}{2}\sigma^2)\beta=\rho$. Then the standard relationship between $N(\cdot)$ and $\Phi(\cdot)$ brings about $Q^-(R,L)$ in Eq. (11). Finally, we derive $L^-(R,L)$:

(B.5) $$L^-(R,L) = \int_0^{L} x d\Lambda(x|\mu_\tau,\sigma_\tau^2) = e^{\mu_\tau+\frac{1}{2}\sigma_\tau^2}\Lambda(L|\mu_\tau+\sigma_\tau^2,\sigma_\tau^2) = \mathrm{Re}^{\mu\tau}\,\Phi(u-\sigma_\tau)\,.$$

In section 8, the price after investment lag is lognormal $\Lambda(x|\mu_\tau,\sigma_\tau^2)$, where $\mu_\tau=\ln H+(\mu-1/2\sigma^2)\tau$ and $\sigma_\tau^2=\sigma^2\tau$. Using the appropriate discount factors and intervals of integration, the components of the profit function in Eq. (16) follow as above. The integrand of $Q^+(H,R,L)$ is $(L/x)^{\alpha}$, and the integrand of $Q^-(H,S)$ is $(x/S)^{\beta}$. The similar integrands of $Q_c^+(H,R,S)$ and $Q_c^-(H,S,R)$ are $[(x/R)^{\beta}-(S/R)^{\alpha}(S/x)^{\beta}]/[1-(S/R)^{(\alpha+\beta)}]$ and $[(S/x)^{\alpha}-(S/R)^{\alpha}(x/R)^{\beta}]/[1-(S/R)^{(\alpha+\beta)}]$. After completing all of these calculations, it is not hard to determine $p^+(R,H)$, $p^-(S,H)$ and $R^+(R,H)$.