Adjustment Costs and Optimal Firm Size

Gerard O. Gaudet

Follow this and additional works at: https://ir.lib.uwo.ca/economicsresrpt
Part of the Economics Commons

Citation of this paper:
Research Report 7309

ADJUSTMENT COSTS AND OPTIMAL FIRM SIZE

Gérard O. Gaudet*
University of Western Ontario

*I would like to thank Albert Ando, Walter P. Heller, and Peter Howitt for helpful comments. They of course bear no responsibility for any remaining errors.
The recent literature on investment behaviour of the firm with adjustment costs has glossed over the distinction between "internal" and "external" adjustment costs, mostly because the distinction was not crucial to the conclusions at hand. Most authors have as a result chosen to work with the somewhat neater internal cost specification of the adjustment cost function (see [3], [5], [9], [10], [11], [12]), or, where an external cost formulation was used, its implications were not fully exploited (see [2], [4], [6], [11]).

The internal cost approach makes the adjustment costs a function of the rate of net investment: expansion of the firm's stock of capital requires reorganization of the production lines, retraining of workers, temporary slowdowns in the production process, etc., all costs which will usually not be independent of the rate of expansion. These are not reflected in the market price of capital goods, but are considered purely internal to the firm. An external cost specification would on the other hand take account as well of costs which come entirely from the market side of the transaction, due to imperfections and frictions in the market for capital goods itself. The firm would face these costs whether it buys the capital goods for expansion or replacement purposes, and its adjustment costs would therefore be a function of its gross rather than net rate of investment. ¹ This seems like a more general specification of the problem in the sense that there is no reason why it could not be thought of as including internal costs as well.² Furthermore, most results obtained with the internal cost formulation can be obtained as a special case of the external cost approach by setting the rate of depreciation of capital equal to zero.
If that was all there is to it, then there would clearly be no loss in generality in either assuming the replacement rate equal to zero or in making adjustment costs a function of the rate of net investment only. One of the merits of accounting for adjustment costs in the decision process of the firm is that it can serve as one of the justifications for a limit to the size of the firm. The purpose of this paper is to show that the external cost approach is richer in its conclusions with respect to optimal firm size than is the internal cost specification. More specifically, we will show that if the firm's adjustment costs depend on its gross rate of investment, we may well find that its long-run equilibrium stock of capital lies in the increasing returns range of the production function even assuming a perfect output market. In the standard neoclassical theory of the firm such an equilibrium would of course be unstable, as it would if adjustment costs were a function of the net rate of investment only (see Treadway [10]). We will show that this is not necessarily the case if the firm faces external adjustment costs.
We describe the technology of the firm by the production function

\[ F(K(t), L(t)) \]

which we assume satisfies the following properties:

(i) \( F(K, L) \) is quasi-concave,

(ii) \( F_K > 0, \quad F_L > 0, \)

(iii) \( F_{KK} < 0, \quad F_{LL} < 0, \quad \text{for } K, L > 0, \)

(iv) \( \lim_{L \to \infty} F_L(K, L) = 0, \quad \lim_{L \to 0} F_L(K, L) > 0, \quad \text{for } K > 0, \)

\[ \lim_{K \to 0} F_K(K, L) = 0, \quad \lim_{K \to 0} F_K(K, L) > 0, \quad \text{for } L > 0. \quad \ldots (1) \]

The firm owns its stock of capital, \( K(t) \). It hires labour, \( L(t) \), at the wage rate \( w(t) \), and sells its output at price \( p(t) \). We assume that adjustment costs put a wedge between the market supply price of capital goods, \( q(t) \), and its actual acquisition cost to the firm, so that a gross rate of investment of \( I(t) \) requires an expenditure of \( q(t)I(t) + A(I(t)) \), where \( A(I(t)) \) represents adjustment costs as a function of the rate of gross investment.\(^3\)

We will assume that all prices holding at time zero are expected to hold forever. Dropping the explicit time notation, we can then write the net cash flow of the firm at any time \( t \) as

\[ R(t) = pF(K, L) - wL - qI - A(I). \quad \ldots (2) \]

The firm is searching for the adjustment path that will maximize its present value, which is given by

\[ V(0) = \int_0^\infty e^{-rt} R(t) \, dt, \quad \ldots (3) \]

where \( r \) is the continuous market rate of interest at which the firm can borrow or lend. The path that maximizes (3) is of course also subject to the constraint
\[ \dot{K} = I - \delta K \] ... (4)

where \( \delta \geq 0 \) is the constant rate of physical depreciation of the stock of capital, as well as

\[ K(0) = K_0, \] ... (5)

i.e., the given initial stock of capital.

We also assume the adjustment cost function to satisfy the following usual properties:

(i) \( A(I) \geq 0, \)
(ii) \( A'(I) \geq 0 \) as \( I \geq 0, \) ... (6)
(iii) \( A''(I) > 0. \)

We will speak of increasing, constant or decreasing returns to scale with respect to the spacing of isoquants along an isocline. Let

\[ Q = F(K,L) \] ... (7)
give the quantity of output, and let

\[ \rho = \alpha K + \beta L, \] ... (8)

where \( \rho \) is the scale coefficient and \( \alpha \) and \( \beta \) are arbitrary positive constants. Then isoquants become closer together, are evenly spaced, or become farther apart, and marginal returns to scale are increasing, constant, or decreasing, as

\[ \frac{d^2 \rho}{dq^2} \leq 0. \]

Along an isocline, i.e., points of tangency between (8) and isoquants of (7), \( K \) and \( L \) must also satisfy

\[ \alpha - \gamma F_K = 0 \] ... (9)
\[ \beta - \gamma F_L = 0 \] ... (10)
where \( \gamma > 0 \) is a lagrange multiplier. It is easily verified that under these conditions

\[
\frac{d^2 \rho}{dQ^2} = \frac{\gamma H(K,L)}{-\left[ F_{LL} F_{KK}^2 - 2F_{KL} F_{KL} F_{KL} + F_{KK}^2 F_{LL}^2 \right]}
\]

where \( H(K,L) = F_{KK}^2 F_{LL} - F_{KL}^2 \) is the Hessian determinant of the production function in (7). The denominator is positive from the quasi-concavity of the production function, and it follows that along an isocline marginal returns to scale are increasing, constant, or decreasing as \( H(K,L) \) is negative, zero or positive.

We might note that increasing, constant or decreasing marginal returns to scale as just defined correspond exactly to decreasing, constant and increasing marginal cost for given wage rate and rental rate of capital. All we have to do is let \( \alpha \) be the given rental rate of capital and \( \beta \) be the given wage rate. Then \( \rho \) is total cost, \( \gamma \) is marginal cost, and (9) and (10) yield the firm's cost minimizing expansion path.

Our inferences in what follows will be with respect to marginal returns to scale. However, from the relationship between marginal and average curves we know that if we have increasing marginal returns then average returns are increasing as well.

The firm's problem is then to maximize (3) subject to (4) and (5), and \( K, L > 0 \). We form the Hamiltonian

\[
e^{-rt} R(t) + \psi(t) [I(t) - \delta K(t)]. \tag{11}
\]

This expression gives us the present value, discounted to \( t=0 \), of a unit of capital acquired at time \( t \), and is to be maximized with respect to \( L \) and \( I \) at each \( t \). Letting \( \lambda(t) = \psi(t) e^{-rt} \), the necessary conditions can be written...
(i) \( \dot{K} = I - \delta K \), \( K(0) = K_0 \)

(ii) \( \dot{\lambda} = (r + \delta) \lambda - p F_K \)

(iii) \( \lambda = q + A'(I) \) \( ... \) (12)

(iv) \( F_L = w/p \)

(v) \( \lim_{t \to \infty} e^{-rt} \lambda(t) = 0 \)

It is convenient to reformulate conditions (12-i) to (12-iv), by differentiating (12-iii) with respect to time and substituting into (12-ii). We then get

(i) \( \dot{K} = I - \delta K \)

(ii) \( \dot{I} = \frac{(r+\delta) [q+A'(I)] - p F_K(K,L)}{A''(I)} \) \( ... \) (13)

where now \( L = 1(K,w/p) \) from (12-iv). These first order differential equations in \( K \) and \( I \), along with the initial condition (5) and the transversality condition determine the optimal investment path of the firm.
We now consider some of the possible optimal paths in phase space. If we let \((K^*, I^*)\) denote an arbitrary stationary point of system (13), then

\[(i) \quad I^* - \delta K^* = 0\]

\[(ii) \quad (r + \delta)[q + A'(I^*)] - pF_K (K^*, L^*) = 0 \quad \cdots(14)\]

where \(L^* = 1(K^*, w/p)\). The slope of the \(\dot{K} = 0\) locus in \((K, I)\) - space is \(\delta\).

The slope of the \(\dot{I} = 0\) locus is given by

\[
\frac{dI}{dK} \bigg|_{\dot{I} = 0} = \frac{p(F_{KK}F_{LL} - F_{KL}^2)}{(r+\delta) A'(I) F_{LL}} \quad \cdots(15)
\]

From assumptions (1-iii) and (6-iii), the denominator of the right-hand side of (15) is negative. The numerator takes on the sign of the Hessian determinant of the production function. We will assume that the Hessian determinant is negative for \(K < K'\) and positive for \(K > K'\), taking on the value zero at \(K = K'\). This corresponds to increasing marginal returns to scale (for \(K < K'\)), followed by constant returns (at \(K = K'\)) and decreasing returns (for \(K > K'\)), and

\[
\frac{dI}{dK} \bigg|_{\dot{I} = 0} \approx 0 \quad \text{as} \quad K \searrow K'.
\]

The \(\dot{I} = 0\) locus therefore takes the shape depicted in Figures 1 and 2.

From assumption (1-iv) and assumptions (6), we find that

\[
\lim_{K \to 0} I \bigg|_{\dot{I} = 0} = \text{some negative constant},
\]

and
\[
\lim_{K \to 0} \left. \frac{I}{I = 0} \right|_0 \geq 0 \text{ as } \lim_{K \to 0} \frac{F_K \geq (r+\delta)}{q/p}.
\]

Depending on whether \( \lim_{K \to 0} F_K \) is greater or smaller than \((r+\delta)q/p\), which is the rental price of capital that would prevail if there were no adjustment costs, we therefore have two possible situations. With \( \lim_{K \to 0} F_K < (r+\delta)q/p \), if the \( I = 0 \) locus cuts the \( K = 0 \) locus, it will cut it twice (and only twice).

We then have two stationary points, as depicted in Figure 1.

If we linearize system (13) around an arbitrary stationary \((K^*, I^*)\), we find that its characteristic roots are

\[
\theta = \frac{r}{2} \pm \left[ \left( \frac{r}{2} \right)^2 + (r+\delta) \delta - \frac{p}{H(K^*, w/p)} \right]^{1/2}
\]

where \( H(K, w/p) = F_{KK} F_{LL} - F_{KL}^2 \), the Hessian determinant of the production function.

If we denote the two stationaries by \((K_1^*, I_1^*)\) and \((K_2^*, I_2^*)\) respectively (with \( K_1^* < K_2^* \)), then at \((K_1^*, I_1^*)\) the \( I = 0 \) must necessarily cut the \( K = 0 \) locus from below and so at that point

\[
\left. \frac{dI}{dK} \right|_{I = 0} = \frac{p}{(r+\delta) \frac{H(K_1^*, w/p)}{A''(I_1^*) F_{LL}(K_1^*, L^*)}} > \delta = \left. \frac{dI}{dK} \right|_{K = 0}
\]

It follows that we always have

\[
(r+\delta) \delta - \frac{pH(K_1^*, w/p)}{A''(I_1^*) F_{LL}(K_1^*, L^*)} < 0, \quad \ldots (16)
\]

so that the roots at \((K_1^*, I_1^*)\) can be either real and both positive, or complex with positive real parts, and so \((K_1^*, I_1^*)\) is always unstable.

On the other hand, since at the second stationary, \((K_2^*, I_2^*)\), the
\( I = 0 \) locus always cuts the \( K = 0 \) locus from above, we necessarily have

\[
\frac{(r+\delta)\delta - pH(K_2^*, w/p)}{A''(I_2^*)F_{LL}(K_2^*, L^*)} > 0 \quad \ldots (17)
\]

and so its roots are necessarily real and of opposite signs. This means that \((K_2^*, I_2^*)\) is always a saddle-point.

The point that we wish to emphasize is that, whereas inequality (16) requires the Hessian determinant to be negative, inequality (17) does not require the Hessian determinant to be positive. In other words, we always have \(K_1^* < K'\), and so the first stationary always occurs in the increasing returns range of the production function, whereas the stable stationary \((K_2^*, I_2^*)\) can occur in any range of the production function.\(^7\)

In Figure 1, we assume that the roots at \((K_1^*, I_1^*)\) are real, in which case \((K_1^*, I_1^*)\) is an unstable node since both its roots are positive.\(^8\)

The optimal path is indicated by the heavy arrows, which are the stable arms of the saddle-point. If \(K(0) < K_1^*\), then the optimal path leads in finite time \(T\) to \(K(T) = 0\), i.e., the firm goes out of business. Note that it will not go out of business instantaneously because \(A' < 0\) and \(A'' > 0\) for \(I < 0\).\(^9\)

If \(K(0) > K_1^*\), then the optimal path leads asymptotically to \((K_2^*, I_2^*)\). It is clear from Figure 1 and from inequality (17) that the equilibrium \(K_2^*\) could be either greater or smaller than \(K'\), and thus could fall in either the decreasing or increasing returns range of the production function.

Figure 1 is drawn with \(K_2^* < K'\).

In the case where \(\lim_{K \to 0} F_K > (r+\delta)q/p\), then \(\lim_{K \to 0} I > 0\). This situation is somewhat simpler to depict since the \(I = 0\) locus then cuts the \(K = 0\) locus only once, and necessarily cuts \(K\) from above. The unique stationary is necessarily a saddle-point, with the optimal path being the stable arms of the saddle-point (see Figure 2). For any \(K(0)\), the optimal path leads
asymptotically to the unique $K^*$. Again, $K^*$ can lie in any range of the production function, and Figure 2 is drawn with $K^*$ in the increasing returns range.
IV

It is obvious that if there exists adjustment costs that are purely external to the firm and so affect the acquisition cost of capital goods whether these capital goods are bought for replacement or expansion purposes, then, when the rate of net investment is zero, and so gross investment equals replacement investment, these costs will not disappear. What we have just shown by examining the dynamic adjustment path of the firm is that if, ceteris paribus, the rate of replacement is relatively high, then the equilibrium adjustment costs may very well be high enough to make it efficient for the firm to operate in the increasing returns to scale range of its production. It is instructive to reconsider the problem in a more intuitive static profit maximization framework.

Consider the problem of the firm choosing K and L to maximize profits every period. Given this optimal K, and with all prices remaining the same, the firm will wish to maintain this stock of capital, and so its gross investment will equal replacement investment.

The firm's total costs of production are then given by

\[ wL + (r + \delta) S(K) \]  \hspace{1cm} (18)

where \( S(K) \) is the valuation of this firm's stock of capital. The per unit price of investment goods, at the rate of investment I, is

\[ q + A(I)/I \]

and with \( I = \delta K \), this becomes

\[ q + A(\delta K)/\delta K \]

dependant on K and \( \delta \). Therefore

\[ S(K) = (q + A(\delta K)/\delta K)K \]
The necessary conditions for profit maximization are

\[ (i) \quad p \frac{F_L}{F_K} = w \]
\[ (ii) \quad p \frac{F_K}{F_K} = v(K) \tag{19} \]

where \( v(K) = (r + \delta)(q + \Lambda'K) \). These are of course the same conditions that are satisfied at the stationaries of Figures 1 and 2. The expansion path of this firm is given by

\[ \frac{F_L}{F_K} = \frac{w}{v(K)} \tag{20} \]

and solving for \( K \) and \( L \) from (20) and the constraint

\[ F(K,L) = Q \tag{21} \]

we get \( L = L(Q) \) and \( K = K(Q) \). Substituting in (18) we get the firm's total cost function, \( C(Q) \). Marginal cost is given by

\[ C'(Q) = \frac{wdL + v(K) dK}{F_L dL + F_K dK} \tag{21} \]

Upon substitution of conditions (18) into (21) we find, as we would expect, that the profit maximizing level of output is given by

\[ C'(Q) = p \]

Except for the correction for adjustment costs, this is the standard profit maximizing problem. However, although \( C(Q) \) is the firm's true cost function, the corresponding average and marginal cost curves tell us little about returns to scale. As usually defined, and as we have defined it here, returns to scale is a purely technical parameter relating to the production function, and would not reflect that relative factor prices to the firm are in fact being affected by the existence of adjustment costs. It can of
course be related, as we have already pointed out, to the firm's cost structure given a constant factor price ratio. For this we consider costs net of the terms due solely to adjustment costs. This is given by

\[ C_N(Q) = C(Q) - (r + \delta) \frac{A(\delta K)}{\delta} \]

and so at the profit maximizing level of output we have

\[ C_N'(Q) = p - (r + \delta) \frac{A'(\delta K)}{F_K} . \]

The last term being positive, this means that marginal cost net of adjustment cost is smaller than price at equilibrium. The difference between the two marginal cost curves is a function, among other things, of \( \delta \). If for example replacement investment is relatively large, then the case shown in Figure 3 is quite conceivable. Here the minimum of \( C_N(Q)/Q \) occurs at a level of output higher than that at which \( C'(Q) = p \). The result is that although the equilibrium will of course always occur in the non-decreasing part of \( C(Q)/Q \), it may very well lie in the decreasing range of \( C_N(Q)/Q \) (and indeed even in the decreasing range of \( C_N'(Q) \)). This corresponds to the increasing returns range of the production function.

The explanation is straightforward: although the firm still finds itself technically facing increasing returns to scale, equilibrium adjustment costs may very well be large enough that its true average costs, inclusive of adjustment costs, are already rising.

This fact is also brought out by the second order conditions for profit maximization. Whereas the usual sufficiency condition, in the absence of adjustment costs, requires the Hessian determinant to be positive, \(^{10}\) when adjustment costs are accounted for we can verify that sufficiency simply requires inequality (17) to be satisfied. As we have already seen this
imposes no restrictions on the sign of the Hessian determinant, and therefore on the range of the production function as previously related to the Hessian determinant.
In conclusion, we should mention two points. First, that our definition of returns to scale does not directly relate to the firm's true expansion path and true cost structure is obvious. The firm will of course still operate in the non-decreasing range of its true cost curve. One might object that we should then redefine returns to scale accordingly. But discussions of returns to scale usually would treat it with reference purely to the technical production function relationship, and empirical studies dealing with returns to scale have certainly not defined it explicitly with the possible effect of adjustment costs on the expansion path in mind. Our point then is that a scale coefficient greater than unity could possibly be explained by the presence of external adjustment costs, and so does not necessarily have to be rejected on the grounds that it would be inconsistent with stability of the equilibrium, even under the assumption of a perfect output market.

Finally, it is worth pointing out that our results are not limited to an adjustment cost interpretation. The adjustment cost formulation used here could very well be reinterpreted as a particular form of an upward sloping supply curve for investment goods, the particularity being that it is strictly convex to the origin. This is of some relevance for empirical studies of the production function, which are usually carried out at a fairly high level of aggregation (i.e., at the industry level or higher), and usually assume that what holds at the level of the firm also holds at the higher level of aggregation. Although it may be correct to assume, for instance, that each firm in an industry can act as a price taker in the market for investment goods, it may not be correct to assume that the whole industry can.
Footnotes

1 The term "adjustment costs" may not be entirely appropriate here, since these costs will not disappear when the firm is simply replacing its existing stock and therefore not "adjusting" its stock.

2 As pointed out by Lucas [4], even if we consider only so-called internal costs, it is not clear that only net investment should enter the adjustment costs function: think of replacing an old machine by a newer model. This can very well involve retraining and other disruptions of the production process although no net expansion is taking place. We would also argue that the interpretation of external adjustment costs is not restricted to situations of monopsony or oligopsony in the investment goods market (see Treadway [10], p. 227), although this is clearly a possible interpretation. Indeed, transactions are not, as a rule, costless. Information flows are not perfect and instantaneous so that resources must be spent in searching the market, so to speak. Increasing the rate of acquisition of investment goods may also at times require shortening, at a cost, unexpected lags that may appear at various stages of the transaction. Increasing the rate of investment may well require speeding up of dismantling, installation, or construction, which may require overtime work and other charges of the kind. All of these would in effect introduce a wedge between the supply price of investment goods and its effective cost to the buyer, regardless of whether it is bought for replacement or expansion purposes. Most of these would be ignored by a net investment specification of the adjustment cost function.

3 We interpret \( A(I(t)) \) to be a monetary cost paid out by the firm. Another possibility would be to treat them as "output-reducing" costs, and value them at \( P(t) \). Our results lend themselves to either interpretation, and to that extent the reader can choose either interpretation.

4 See Rowe [8].

5 See Pontryagin, et al. [7]. These conditions are fairly standard by now and we will not dwell on an interpretation of each. The discussion in Treadway ([10],[11]) can be applied here, mutatis mutandis. For a good interpretation of these conditions for the problem without adjustment costs, see also Arrow [1].

6 We will not make any direct use of the transversality condition (12-V) in this paper.

7 Notice that if \( \delta = 0 \), or if the adjustment costs function depends only on the rate of net investment, then the second stationary has to lie in the decreasing returns range. This is the case in the model analyzed by Treadway [10].
8 The other possibility is that \((K^*_1, I^*_1)\) has complex roots, in which case it would be characterized locally by an unstable spiral, with the possibility of being contained in a limit cycle. The characteristics of the optimal path are of course different from those of Figure 1, but \((K^*_2, I^*_2)\) may still lie in the increasing returns range.

9 See Treadway [10].

10 To be more specific, the Hessian is required to be negative definite, which implies concavity. This is associated with non-increasing returns to scale in the relevant range, even by definitions of returns to scale other than the one used here. The definition used here has the merit of allowing us to directly relate returns to scale to the sign of the Hessian determinant in all three ranges of increasing, constant and decreasing returns. It is clear that this is not completely essential to our results: all we need is that a positive Hessian determinant is associated with non-increasing returns to scale.

11 This convexity can be shown to be a second-order necessary condition (Legendre condition) for a maximum in the variational problem at hand. We have guaranteed satisfaction of this condition in our problem by assuming \(A''(I) > 0\).
References


