Natural Monopoly Revisited

Oriol Carbonell-Nicolau*

September 2023

Abstract

We study the conditions under which production processes exhibit a decreasing average cost function in the absence of perfectly competitive input markets.

Keywords: natural monopoly, monopsony, returns to scale, average cost function.
JEL classifications: L12, L13.

Natural monopolies are typically defined as industries in which the average cost of production decreases with output. The property of increasing returns to scale fully characterizes natural monopolies in the sense that it is necessary and sufficient for the existence of decreasing average costs.[1] This result rests on the important assumption, often omitted, that input markets are perfectly competitive.

In the simplest case when there is only one input, say labor, denoted by l, the production function $f(l)$ exhibits increasing returns to scale if and only if

$$f(\lambda l) > \lambda f(l), \quad \text{for all } \lambda > 1 \text{ and all } l > 0,$$

which can be equivalently stated as

$$c(f(\lambda l)) > c(\lambda f(l)), \quad \text{for all } \lambda > 1 \text{ and all } l > 0, \tag{1}$$

where $c(\cdot)$ denotes the cost function. Letting $w > 0$ be the competitive wage rate, we have

$$c(f(\lambda l)) = w\lambda l = \lambda c(f(l)),$$

and so (1) can be restated as

$$\lambda c(f(l)) > c(\lambda f(l)), \quad \text{for all } \lambda > 1 \text{ and all } l > 0,$$

or, equivalently,

$$\lambda c(x) > c(\lambda x), \quad \text{for all } \lambda > 1 \text{ and all } x > 0,$$

which holds if, and only if,

$$\frac{c(x)}{x} > \frac{c(\lambda x)}{\lambda x}, \quad \text{for all } \lambda > 1 \text{ and all } x > 0,$$

*Department of Economics, Rutgers University, 75 Hamilton St., New Brunswick, NJ 08901. E-mail: carbonell-nicolau@rutgers.edu.

[1] See, e.g., Kreps (2004, pp. 219–220) and Serrano and Feldman (2013, pp. 152–155).

i.e., if, and only if, the average cost function is decreasing.

The purpose of this note is to understand the conditions under which production processes exhibit a decreasing average cost function in the absence of perfectly competitive input markets. The relevance of this line of inquiry rests on the observation that monopoly power tends to go hand-in-hand with monopsony power.[2]

It will be shown that increasing returns to scale are necessary but not sufficient for a decreasing average cost function, and that measures of the degree of oligopsony in input markets are important additional factors determining the shape of average costs.

The analysis is framed in terms of market power in upstream labor and capital markets. To keep the analysis as simple as possible, we confine attention to the one-input case.

To begin, consider a monopolist that produces a private good using labor, l, as its only input. The production function is denoted by $f(l)$. Suppose that $w(l)$ is an increasing function representing the *firm-level* inverse labor supply. This function describes how workers employed by the monopolist react to changes in wages. When $w(\cdot)$ is flat (i.e., completely elastic), workers quit (moving to other firms, for example) when the monopolist lowers their wages. This extreme case represents the textbook case of competitive input markets, i.e., a complete absence of the monopolist's labor market power. When $w(\cdot)$ is very steep, the monopolist's labor force tends to be unresponsive to even large wage declines. In this case, the monopolist holds significant labor market power.

The monopolist's cost function, expressed in terms of hours of hired labor, l, is given by $w(l)l$. The *elasticity of the cost function with respect to l* is expressible as

$$\frac{d(w(l)l)}{dl} \cdot \frac{l}{w(l)l} = \frac{w'(l)l + w(l)}{w(l)} = 1 + \frac{1}{\xi_l}, \tag{2}$$

where ξ_l denotes the wage elasticity of the labor supply facing the monopolist, i.e.,

$$\xi_l = \frac{1}{w'(l)} \cdot \frac{w(l)}{l}.$$

This elasticity is called "*residual labor supply elasticity*" in Naidu et al. (2018). It takes values in the range $[0, \infty)$. The interval's lower bound (resp., upper bound) represents the case of absolute labor market power (resp., the absence of labor market power).

The elasticity of the cost function with respect to l given in (2) measures the relative responsiveness of the monopolist's cost to a one-percent increase in the quantity of hired labor.

The *elasticity of scale*,

$$\xi_f(l) = f'(l) \cdot \frac{l}{f(l)}$$

[2] The case of Amazon is detailed in Khan (2016). See also the empirical analysis in Azar et al. (2022), which shows that "[g]iven high concentration, mergers of employers have the potential to significantly increase *labor* market power." Naidu et al. (2018, pp. 546–547) survey empirical literature suggesting that "industry consolidation has given employers greater bargaining power in labor markets."

takes values in the interval $(0, \infty)$ and measures the relative responsiveness of output to a one-percent increase in the quantity of hired labor. Note that

$\xi_f(l) > 1$ for all l if and only if the production function $f(\cdot)$ exhibits increasing returns to scale;
$\xi_f(l) = 1$ for all l if and only if the production function $f(\cdot)$ exhibits constant returns to scale;
$\xi_f(l) < 1$ for all l if and only if the production function $f(\cdot)$ exhibits decreasing returns to scale.

The ratio of the elasticity of the cost function with respect to l to the elasticity of scale,

$$\theta(l) = \frac{1 + (1/\xi_l)}{\xi_f(l)},$$

measures the relative responsiveness of the total cost to the relative responsiveness of output (with respect to a one-percent increase in the quantity of hired labor). Note that

$\theta(l) < 1$ for all l if and only if the monopolist's average cost function is decreasing;
$\theta(l) = 1$ for all l if and only if the monopolist's average cost function is constant;
$\theta(l) > 1$ for all l if and only if the monopolist's average cost function is increasing.

Thus, when the labor market is not competitive, the shape of the average cost function is determined by the elasticity of scale *and* the residual labor supply elasticity. Standard textbooks consider the extreme case of a competitive labor market, i.e., the case when $\xi_l = \infty$, which yields $\theta(l) = 1/\xi_f(l)$, and so $\theta(l)$ becomes the inverse of the elasticity of scale. In this case, increasing returns to scale are necessary and sufficient for a decreasing average cost curve.

In general, increasing returns to scale are necessary but not sufficient for a decreasing average cost curve. This is because $\theta(l) < 1$ requires $\xi_f(l) > 1$ (i.e., increasing returns to scale). However, if the monopolist holds significant labor market power, so that the residual labor supply elasticity is relatively low and the elasticity of the cost function relatively high, $\theta(l)$ tends to be large—in particular, greater than one. In fact, in the extreme case of absolute labor market power, i.e., the case when $\xi_l \approx 0$ for a range of hired labor l, we have $\theta(l) \approx \infty$, implying that the monopolist's average cost function is increasing (for said range of l).

Let us now consider the case of capital input markets.[3] Suppose that the monopolist owns the capital used in its production process and chooses how much of its capital stock, K, is used as an input in its own production. Alternatively, the monopolist can loan capital to other agents.

The production function is denoted by $f(k)$, where k measures 'capital.' Let $r(k)$ be a decreasing function representing the inverse demand for capital facing the monopolist.

[3] The essence of the argument that follows also applies to the case of 'land.'

The price elasticity of the firm-level demand for capital,

$$\epsilon_k = \frac{1}{r'(k)} \cdot \frac{r(k)}{k},$$

which ranges between 0 and ∞, can be taken as a measure of the monopolist's capital market power as a supplier of capital.[4] A perfectly inelastic (resp., elastic) demand represents the case of absolute market power (resp., the absence of market power).

If the monopolist loans κ units of capital and uses $K - \kappa$ units of capital in the production of its own final good, the cost of using an extra (marginal) unit of capital in the production of the final good is the opportunity cost of loaning that unit, i.e.,

$$MR(\kappa) = r(\kappa) + r'(\kappa)\kappa,$$

which represents the monopolist's marginal revenue evaluated at the quantity of capital loaned, κ. Intuitively, this opportunity cost consists of the forgone unit price of capital, $r(\kappa)$, minus the extra revenue from the increase in the price of capital paid for the inframarginal units, $r'(\kappa)\kappa$; the increase in the price of capital results from the marginal reduction in the supply of loanable funds.

Hence, the monopolist's marginal economic cost of using k units of capital in the production of the final good (which gives a corresponding loan size of $K - k$ units of capital) is given by

$$c'(k) = MR(K - k),$$

while the monopolist's total economic cost of using k units of capital in the production of the final good is given by

$$c(k) = \int_{K-k}^{K} MR(\kappa) d\kappa.$$

Note that the marginal cost is negative for those values of k for which $MR(K - k)$ is negative. However, any loan supply of size $K - k$, where $MR(K - k)$ is negative, is not profit maximizing, since, at those levels, reducing the loan supply brings about extra revenue.

The *elasticity of the cost function with respect to k*, which measures the relative responsiveness of the monopolist's cost to a one-percent increase in the quantity of capital, can be written as

$$c'(k) \cdot \frac{k}{c(k)} = \frac{MR(K-k)k}{c(k)}.$$

Note that if the $MR(\cdot)$ function is decreasing, then this elasticity is greater than 1, since, in this case, the marginal cost exceeds the average cost:

$$\frac{MR(K-k)k}{c(k)} > 1 \Leftrightarrow \frac{c'(k)}{c(k)/k} = \frac{MR(K-k)}{c(k)/k} > 1.$$[5]

[4]The Lerner index of a firm's market power (Lerner, 1934) can be expressed solely in terms of the firm-level price elasticity of demand.

[5]A downward sloping $MR(\cdot)$ function is sufficient but not necessary for the elasticity of the cost function with respect to k to be greater than one.

Note also that $MR(K-k)$ is expressible, in terms of the price elasticity of the firm-level demand for capital, as

$$MR(K-k) = \left(1 + \frac{1}{\epsilon_{K-k}}\right) r(K-k).$$

The ratio of the elasticity of the cost function with respect to k to the elasticity of scale,

$$\theta(k) = \frac{MR(K-k)k}{c(k)} \Big/ \xi_f(k) = \frac{\left(1+\frac{1}{\epsilon_{K-k}}\right) r(K-k)k}{c(k)} \Big/ \xi_f(k),$$

measures the relative responsiveness of the total cost to the relative responsiveness of output (with respect to a one-percent increase in the quantity of hired labor). Note that

$\theta(k) < 1$ for all k if and only if the monopolist's average cost function is decreasing;
$\theta(k) = 1$ for all k if and only if the monopolist's average cost function is constant;
$\theta(k) > 1$ for all k if and only if the monopolist's average cost function is increasing.

Thus, when the labor market is not competitive, the shape of the average cost function is determined by the elasticity of scale *and* the firm-level price elasticity of demand.

The extreme case of a perfectly competitive capital market, i.e., the case when $\epsilon_k = \infty$, corresponds to the case of a flat inverse demand function $r(\cdot)$ (hence a flat $MR(\cdot)$ function), implying that the elasticity of the cost function with respect to k is equal to one. In this case, $\theta(k) = 1/\xi_f(k)$, implying that increasing returns to scale are necessary and sufficient for a decreasing average cost curve.

Under complete capital market power, $\epsilon_k = 0$, which gives an infinite elasticity of the cost function with respect to k, implying that the monopolist's operating average costs are increasing.

If the elasticity of the cost function with respect to k is greater than 1 (which is true, for example, if the $MR(\cdot)$ function is decreasing), then increasing returns to scale are necessary, but not sufficient, for a decreasing average cost function.

References

Azar, José, Ioana Marinescu, and Marshall Steinbaum (2022) "Labor Market Concentration," *Journal of Human Resources*, Vol. 57, pp. S167–S199, DOI: `http://dx.doi.org/10.3368/jhr.monopsony.1218-9914R1`.

Khan, Lina H. (2016) "Amazon's antitrust paradox," *Yale Law Journal*, Vol. 126, URL: `https://digitalcommons.law.yale.edu/ylj/vol126/iss3/3`.

Kreps, D. M. (2004) *Microeconomics for Managers*: Norton, pp.1–652.

Lerner, A. P. (1934) "The Concept of Monopoly and the Measurement of Monopoly Power," *The Review of Economic Studies*, Vol. 1, pp. 157–175, DOI: `http://dx.doi.org/10.2307/2967480`.

Naidu, Suresh, Eric Posner, and Glen Weyl (2018) “Antitrust Remedies for Labor Market Power,” *Harvard Law Review*, Vol. 132, pp. 536–601.

Serrano, R. and A. M. Feldman (2013) *A Short Course in Intermediate Microeconomics with Calculus*: Cambridge University Press, pp.1–381.