Part III

Production

1 Introduction

The issues and results in abstract production theory are similar to those in demand theory: justifying from first principles the problem we solve, namely profit maximization; finding properties of the solutions to the profit-maximization problem and the related cost-minimization problem; understanding how to solve specific examples of these problems and comparative statics (price effects); aggregation of firms - when is there a representative aggregate firm whose solution is the same as the sum of solutions to individual firm problems, and to what extent is this efficient (Paretian).

There are some important differences: production is simpler in that there are no wealth effects, so that aggregation will work better; production is cardinal while consumption utility is ordinal; abstract production is slightly more difficult as in it allows for any good to be an input or output (we can produce tables from chairs or chairs from tables) while in consumption the output is "utility" and the inputs are the goods.

2 Production sets and production functions

There are two ways used to model production: *production sets* and *production functions*. While the first is more general, it requires quite different methods to deal with and some of the analogies between demand and production are clearer when we look at production functions.

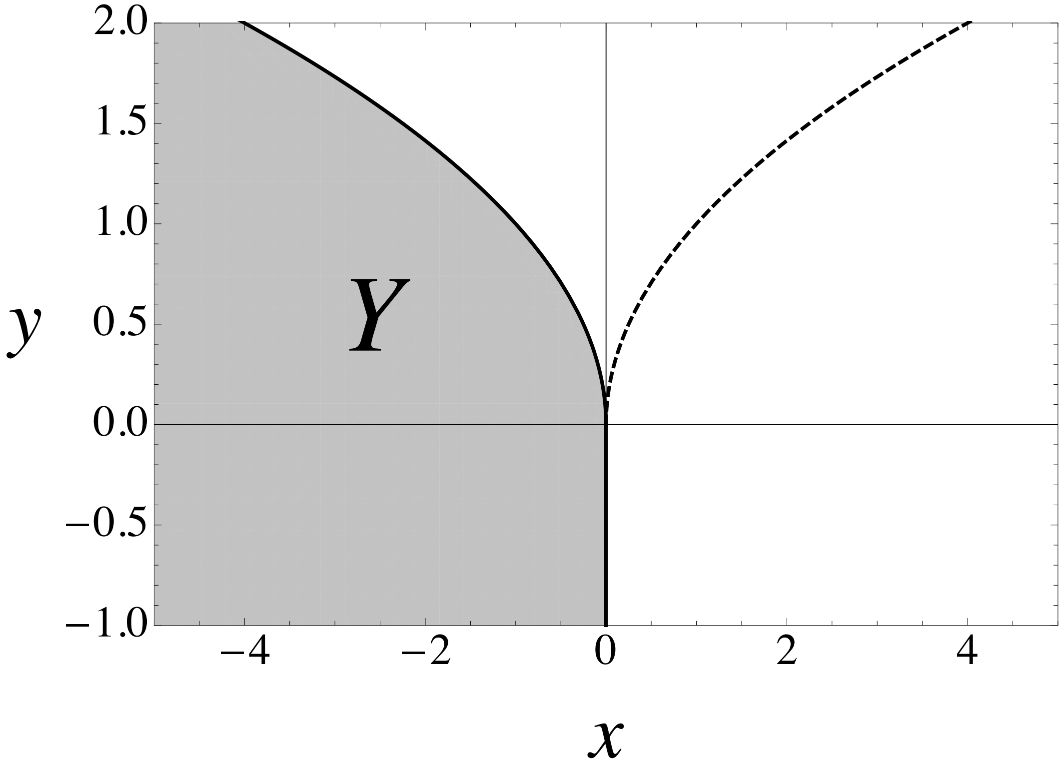
A *production function* is a function $f : \mathbb{R}^n_+ \to \mathbb{R}$[27], to be interpreted as the quantity of a single output, denoted q, that is created by a non-negative vector of inputs, denoted z. If there is more than one input, i.e., $n > 1$, then we have an analog to indifference curves, called isoquants. These are the set of inputs that produce a given level of output: $\{z : f(z) = q\}$. The analog to the MRS is the marginal rate of technological substitution between the inputs, $MRTS_{lk}$, which is the slope of the isoquant; i.e., $f_l(z)/f_k(z)$.

A *production set* is denoted by $Y \subseteq \mathbb{R}^m$. It is interpreted as all the *net* bundles that can result from production — sometimes it is possible to use the same good as an input or an output, and maybe multiple goods can result as net outputs. Given a vector $y \in \mathbb{R}^m$ if $y_i < 0$ then a negative quantity of i is produced at that vector, i.e., good i is an input. Conversely $y_j > 0$ iff y is an output.

When there is only one input Y is the mirror image (along the vertical axis) of the area under the production function, as in the figure below. The production function f is the dashed function on

[27] Typically we take the domain to be all non-negative combinations of inputs, so $f : \mathbb{R}^n_+ \to \mathbb{R}$, although to allow for sunk-costs, which are inputs that have already been purchased, we may define production only for inputs greater than those that have already been sunk to indicate that $z = 0$ is not feasible.

the right-hand side (RHS); the area *on and under* the solid line (that continues down to $(0, -\infty)$) on the LHS is the set Y.



Given a production set Y we can define functions, called *transformation functions*, that satisfy $F(y) \leqslant 0$ iff $y \in Y$, and $F(y) = 0$ iff y is on the boundary of Y. For instance, with single-good production, for $y = (-z, q)$, it is natural to define $F(y) = q - f(z)$, which is non-positive iff $q \leqslant f(z)$ (i.e., q is feasible given z) and the boundary is then $F(y) = 0$ so the boundary is equivalent to the production function $q = f(z)$, as in the figure above. Assuming F is differentiable then the slope of F at $\bar{y}$ when we consider changing only two goods l and k (holding the level of all others fixed) is $MRT_{lk} = \frac{\partial F(\bar{y})}{\partial y_l} / \frac{\partial F(\bar{y})}{\partial y_k} = \frac{F_l}{F_k}$.[28] In the case of two inputs l and k, and one output, where f describes the boundary of Y, this is the same as the MRTS.

Exercise. *In the case where l and k are two outputs what is the interpretation of MRT_{lk}?*

Solution: *This is the slope of the production-possibility frontier from your undergraduate (guns-and-butter) days.* ■

Exercise. *What is the interpretation of MRT_{lk} when l is an input and k is an output?*

[28] With single-output production, when F is differentiable so is f, but the converse is false. If f has a finite slope at 0, then the boundary of Y has a kink at 0 and therefore F cannot be differentiable. So assuming differentiability of F at zero is a significant economic (and not only a technical) restriction resulting from the use of the Y, rather than the f, formalism. However, except at 0 this issue does not arise, and it is clear that we could easily extend the FOC to Fs that are differentiable everywhere except at 0.

As a first glance at how these two formulations are used consider the profit functions. When dealing with production sets, we consider a price vector $\mathbf{p} \in \mathbb{R}^m_+$, which specifies a price for each good. The profit from choosing bundle $y \in Y$ is then $\mathbf{p} \cdot y$. With production functions, we let $p \in \mathbb{R}_+$ be the price of the output and $\mathbf{w} \in \mathbb{R}^n_+$ be the vector of prices for the inputs. Then the profit from choosing the vector of inputs z is $pf(z) - \mathbf{w} \cdot z$.

To see that production sets are more general than production functions, given a production function f we can define $Y = \{(-z, q) : q \leqslant f(z)\}$ to be the set of feasible input-output vectors. When dealing with production functions, we may refer to the set Y, to be understood as coming from this definition. Clearly Y is more general: Y may allow for two outputs, it allows for both inputs and outputs, it specifies explicitly if there is free disposal, and – more technically – it may be open so that we can produce as close as we want to some $q = f(z)$ but not quite get there.

3 Properties of production functions and sets

3.a Returns to scale and concavity/convexity

Property 3.1. *Non-increasing Returns to Scale (NIRTS):*

1. *The production function f satisfies $f(\alpha x) \geqslant \alpha f(x)$ for $\alpha \in [0,1]$ and all x.*

2. *Any feasible production can be proportionally shrunk (but not necessarily expanded): $y \in Y \Rightarrow \alpha y \in Y$ for $\alpha \in [0,1]$.*

Property 3.2. *Constant RTS (CRTS):*

1. *The production function is homogeneous of degree 1: $f(\alpha x) = \alpha f(x)$.*

2. *Any feasible production can be proportionally changed: $y \in Y \Rightarrow \alpha y \in Y$ for all $\alpha \geqslant 0$.*

Property 3.3. *Non-decreasing RTS:*

1. *The production function f satisfies $f(\alpha x) \geqslant \alpha f(x)$ for $\alpha \geqslant 1$ and all x.*

2. *Any feasible production can be proportionally expanded (but not necessarily shrunk): $y \in Y \Rightarrow \alpha y \in Y$ for all $\alpha \geqslant 1$.*

Property 3.4. *Convexity:*

1. *Y is convex.*

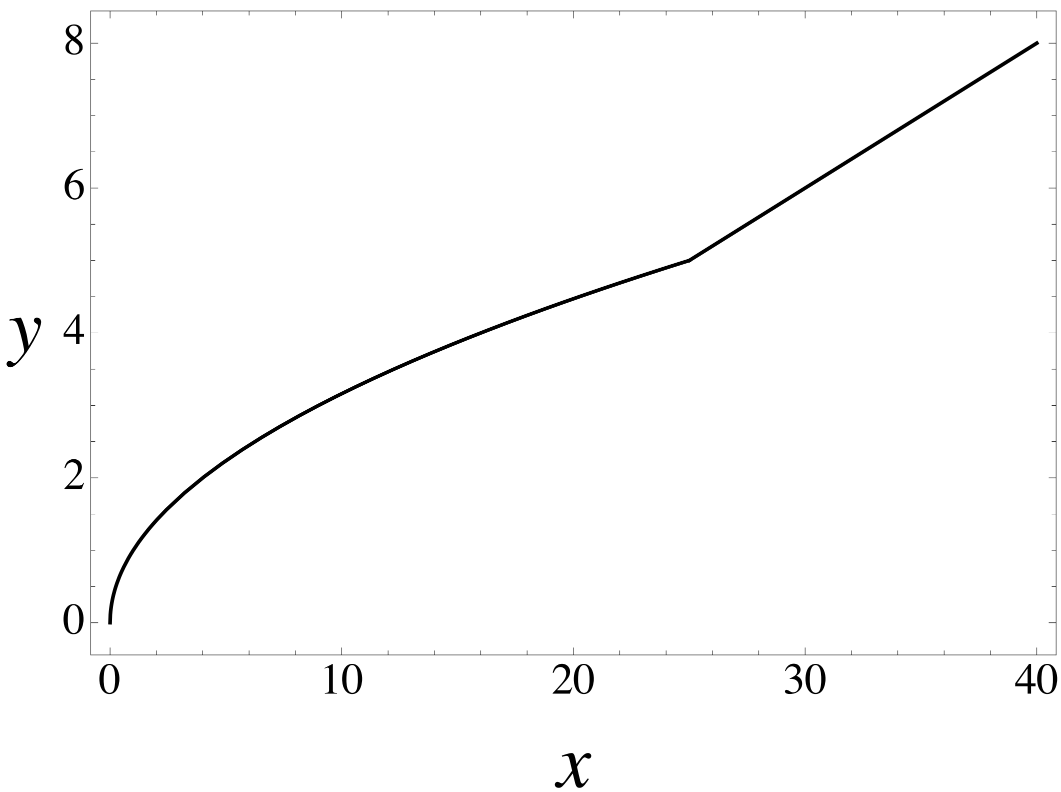
2. *f is concave.*

Exercise. *For each property above, prove the equivalency of the two statements when* $Y = \{(-z, q) : q \leqslant f(z)\}$.

Remark 3.1. *Observe that if we have* f *strictly increasing, strictly quasiconcave, CRTS, and* $f(0) = 0$*, then* f *is concave. As this is usually a problem set in macro we won't be covering it, but if you have not seen it you should view it as a problem set or at least read the proof in Reny's book.*

Convexity of Y (concavity of f) is, in general, a more restrictive assumption than NIRTS in two senses.

First, it is more restrictive in terms of each individual input. RTS is about how $f(\alpha x)/\alpha$ relates to $f(x)$. Consider the one-input case and assume $f(0) = 0$. Then NIRTS requires that any line between the origin and a point on the function $(x, f(x))$ lies below the graph of the function. Concavity requires that the line between *any* two points (not only $(0, f(0))$ and $(x, f(x))$, but also between $(\hat{x}, f(\hat{x}))$ and $(x, f(x))$) lies below the graph of the function. Section 3.a shows a non-concave function that is NIRTS: $f(x) = x^{1/2}$ for $0 \leqslant x \leqslant 25$, and $f(x) = x/5$ for $x > 25$.



Concavity also implies quasiconcavity which says that averaging inputs is good. This has nothing to do with RTS which only considers changes keeping the proportion of inputs fixed. For instance, if mixing two inputs destroys the output, so that $f(z_1, z_2) = 0$ whenever $z_1 \times z_2 \neq 0$, while $f(z, 0) = f(0, z) = z$ we have non-increasing (in fact constant) returns to scale with a (discontinuous and very) non-concave (and non-linear) f.

Observe that CRTS is equivalent to f being homogeneous of degree 1. With one input CRTS is equivalent to linearity. But, as the preceding example indicates, with multiple inputs it is implied by, but does not imply, linearity of f.

3.b Other properties

Property 3.5. *Y is non-empty and closed.*

Property 3.6. *No free lunch: $y \in Y$ and $y \geqslant 0$ implies $y = 0$.*

Property 3.7. *The firm can be shut down: $0 \in Y$.*[29]

Exercise. *What properties of f correspond to properties 3.5-3.7 for the Y defined from f as above?*

Property 3.8. *Free disposal: $y \in Y$ and $y' \leqslant y \Rightarrow y' \in Y$.*

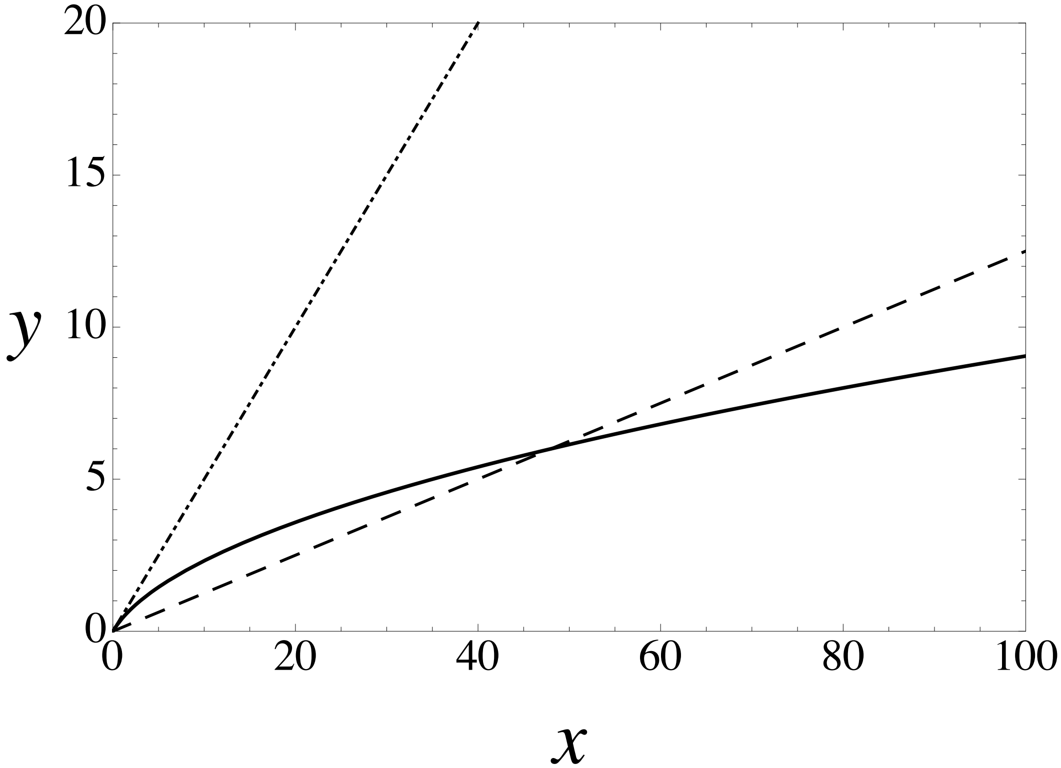
Property 3.9. *Irreversibility: $y \in Y$ and $y \neq 0 \Rightarrow -y \notin Y$.*

The following property is satisfied by non-decreasing RTS but not by DRTS technologies.

Property 3.10. *Additivity (free entry): $y, y' \in Y \Rightarrow y + y' \in Y$.*

This property helps us understand that the Y formalism is more general than the f formalism. Consider the additive closure of a DRTS technology. That is, find the smallest set $\hat{Y}$ such that it is additive and contains a Y that has DRTS. As in the figure below this set contains the expansion of any feasible production, such as any point on the dashed line under the solid line to any point on the dashed line above the solid line, and thus all points along the dashed line, and therefore all those "under" the tangent to the production function at 0 (the dot-and-dash line), but the tangent is not included. So there is no f that generates this $\hat{Y}$ as $\hat{Y}$ is open.

[29] But see footnote 28.



Remark 3.2. *You may want to see why additivity and non-increasing RTS is equivalent to convexity and CRTS (see MWG).*

3.c Cost minimization

While the end objective is to maximize profits, it is clear that to do so, whatever level of output is chosen, it is necessary to minimize costs in production of that output. However, this notion is relevant only for given inputs and output(s). Nevertheless, while not applicable to general Y functions, it is still of significant interest and as this is most similar to what we have studied we begin our study with the cost-minimization problem.

Let $w \in \mathbb{R}^n_+$ be the price vector of the inputs. Then the problem is

$$c(w,q) = \min_{z \geqslant 0} w \cdot z \text{ s. t. } f(z) \geqslant q$$

Let $z(w,q)$, called the conditional factor demands, be the solution to this problem.

Notice the similarity to the expenditure problem: $e(p,\bar{u}) = \min_{x \geqslant 0} p \cdot x$ s.t. $u(x) \geqslant \bar{u}$, with solution $h(p,\bar{u})$. Therefore all the properties that we proved for e and h holds for c and z. The equivalence to compensated demand follows since there are no wealth effects

Proposition 3.1. *1. (Assuming suitable differentiability...) FOC:* $w_l \geqslant \lambda \frac{\partial f(z^*)}{\partial z_l}$, *with equality if* $z_l^* > 0$; *i.e.,* $w \geqslant \lambda \nabla f(z^*)$, *and* $(w - \lambda \nabla f(z^*)) \cdot z^* = 0$.[30] *This implies that for any pair of used inputs the price ratio equals the marginal rate of technological substitution between the inputs,* $MRTS_{lk}$, *which is the slope of the isoquant; i.e.,* $\frac{f_l(z^*)}{f_k(z^*)} = \frac{w_l}{w_k}$.

2. *If* f *is (strictly) quasiconcave, so that the upper contour set* $\{z \geqslant 0 : f(z) \geqslant q\}$ *is (strictly) convex, then* z *is a (singleton) convex set.*

3. c *is homogeneous degree* 1 *in* w *and increasing in* q.

4. z *is homogeneous degree* 0 *in* w.

5. c *is concave in* w.

6. *If* $z(\bar{w}, q)$ *is a singleton then* c *is differentiable in* w *at* $(\bar{w}, q)$ *and* $\nabla_w c(\bar{w}, q) = z(\bar{w}, q)$; *[Shephard's Lemma]*

7. *If* z *is differentiable at* $\bar{w}$ *then* $D_w z(\bar{w}, q) = D_w^2 c(\bar{w}, q)$ *is symmetric negative semidefinite with* $D_w z(\bar{w}, q) \cdot \bar{w} = 0$.

8. *If* f *is homogeneous degree* 1 *(equivalently CRTS) then* c *and* z *are homogeneous degree* 1 *in* q.

The FOC relates the price of an input l to its marginal productivity, MP, f_l. Since λ is the value of relaxing the constraint, it equals the marginal cost of producing another unit. Below we will see that the FOC for maximizing profit sets this equal to the price of the output. So the FOC is to equate (modulus corner solutions) the value of the MP to the price of the output.

There is an additional property that is useful for production, for which no analogy was stated earlier.

Proposition 3.2. *If* f *is (strictly) concave then* c *is a (strictly) convex function of* q *(i.e. marginal costs are (increasing) nondecreasing in* q*).*

Exercise. *Prove the preceding proposition. State the corresponding result for the consumption model studied; why did we not state this result?*

3.d Profit maximization

Let the price of the output be p. Then the vector of output-input prices will be denoted by $\mathbf{p} = (p, w)$. The profit function is

[30] See Figure 5.C.2. in MWG

$$\begin{aligned}\pi(p,w) &= \max_{y\in Y} \mathbf{p}\cdot y \\ &= \max_{q,z} pq - w\cdot z \text{ s. t. } q \leqslant f(z) \\ &= \max_{z} pf(z) - w\cdot z \\ &= \max_{q} pq - c(w,q)\end{aligned}$$

where the last three version apply if we have a production-function specification, and not only a production set. Denote the solutions by $(q(p,w), -z(p,w)) = y(\mathbf{p})$.

The next result summarizes many useful properties of the solution to the profit-maximization problem. Before stating the result it is helpful to relate profit maximization to utility maximization. Specifically, the profit-maximization problem is the same as a utility-maximization problem with quasilinear utility, $u(x) = v(x_{-1}) + x_1$, since then by substituting in the budget constraint, utility-maximization becomes $\max_{\hat{x}} \hat{v}(\hat{x}) - \hat{p}\cdot\hat{x}$, where $\hat{x} = x_{-1}$, $\hat{p} = \frac{p_{-1}}{p_1}$, $\hat{v}(\cdot) = v(\cdot) + \frac{I}{p_1}$. This is because when utility is quasilinear, say in good j, then there are no wealth effects for other goods, i.e., $x_i(p,w)$ does not depend on w for all $i \neq j$. The absence of wealth effects (as we saw) are what causes the law of demand to hold for compensated demand functions (and fail to hold for demand functions) and more generally underlie the results we had on the relationship between e and h.

Proposition 3.3. *1. FOC (Under suitable differentiability assumptions...) Given a production function f, the FOC is that, in addition to the cost-minimization FOC, we have $p - c_q(w,q^*) \leqslant 0$, and $=$ if $q^* > 0$. That is, if the firm produces,then price equals marginal cost.*

Given a transformation function F the FOC are that $\mathbf{p} = \lambda\nabla F(y^*)$, i.e., $\mathbf{p}_l = \lambda F_l$ for all l.

2. *$\pi(\mathbf{p}) = \pi(p,w)$ is homogeneous of degree 1; $y(\mathbf{p}) = (q(p,w), z(p,w))$ is homogeneous of degree 0.*

3. *π is **convex**.*

4. *If f is (strictly) concave, i.e. Y (strictly) convex, then $y(\mathbf{p})$ is a (singleton) convex set.*

5. *If $y(\bar{\mathbf{p}})$ is a singleton then π is differentiable at $\bar{\mathbf{p}}$ and $\nabla_p\pi(\bar{\mathbf{p}}) = y(\bar{\mathbf{p}})$; [Hotelling's lemma].*

6. *If $y(\bar{\mathbf{p}})$ is differentiable at $\bar{\mathbf{p}}$ then $D_p y(\bar{\mathbf{p}}) = D^2_{p^2}\pi(\bar{\mathbf{p}})$ is a symmetric and **positive** semidefinite matrix with $D^2_{p^2}\pi(\bar{\mathbf{p}})\bar{\mathbf{p}} = 0$.*

Proof of item 3. Let $\tilde{\mathbf{p}} := \alpha\mathbf{p}' + (1-\alpha)\mathbf{p}''$.

$$\pi(\tilde{\mathbf{p}}) = \alpha\mathbf{p}'y(\tilde{\mathbf{p}}) + (1-\alpha)\mathbf{p}''y(\tilde{\mathbf{p}}) \leqslant \alpha\mathbf{p}'y(\mathbf{p}') + (1-\alpha)\mathbf{p}''y(\mathbf{p}'')$$

■

You might wonder why when using F we get a simpler FOC than with f – no non-negativity conditions. This is all buried in the differentiability assumption about F (see also footnote 28). For example, consider the one input and one output case, where the input cannot be produced from the output: if f fails to satisfy the Inada condition that $\lim_{x \to 0} f'(x) = \infty$, then for sufficiently high price ratio of w/p, optimal production will be zero and the FOC $p = c_q$ cannot hold; instead we will have $p < c_q(0)$. Analogously, in this case F is not differentiable at $(0, 0)$.

Exercise. *Use the envelope theorem to prove Hotelling's lemma (Proposition 3.3 item 5) when production is specified with a production function. Also prove Hotelling's lemma, using whatever method you like, for the case of a general production set Y.*

Remark 3.3. *Why do we study profit maximization? An owner managed firm (in an environment without uncertainty) who takes prices as given would maximize profits as described. If prices can be influenced, then profit maximization would still occur, but would not necessarily satisfy all the properties above. If there is more than one owner, and prices cannot be influenced, and there is no uncertainty, then they would all want the manager to maximize profits. But then the problem is how to incentivize the manager. With more than one owner, and if prices can be influenced, the owners might disagree about the best strategy. Finally, if there is uncertainty, then owners will, in general, disagree about the appropriate degree of risk taking.*

3.d.1 Comparative statics

Convexity of π implies the differential version of the law of supply; the discrete form is $(\mathbf{p}' - \mathbf{p}'') \cdot (y(\mathbf{p}') - y(\mathbf{p}'')) \geqslant 0$. (To prove this open parenthesis and note that $y(\cdot)$ gives the highest profit, i.e., $y(\mathbf{p}) \cdot \mathbf{p} \geqslant \hat{y} \cdot \mathbf{p}$ for all $\hat{y}$.)

Exercise. *(Trivial.) Does the law of supply hold when $y(\mathbf{p})$ is not single-valued?*

Example 3.1. *(Comparative Statics) MWG 5.C.6.: Assume f is a twice differentiable, strictly concave ($D^2 f$ ND) production function of L inputs and that $D_z f(z) \geqslant 0$ and $D_z f(z) \neq 0$ for all z, i.e. that each input has non-negative marginal product and, at each z, at least one input has positive marginal product.*

1. *Prove that output is increasing in output price.*

   (a) *By strict concavity and part 4 in Proposition 3.3 the solution is unique. Let $\mathbf{p}' = \mathbf{p} + (\varepsilon, 0, ..., 0)$. The law of supply says $(\mathbf{p}' - \mathbf{p}) \cdot (y(\mathbf{p}') - y(\mathbf{p})) \geqslant 0$, so $\varepsilon \times (q(p', w) - q(p, w)) \geqslant 0$, so $\Delta q \equiv q(p', w) - q(p, w)$ is non-negative.*

   We need to prove that Δq is actually positive.

   Let z maximize profit at p, w. Suppose price increases to $p + \varepsilon$. Consider using inputs $z' = \alpha z$, and let $\pi(\alpha, p + \varepsilon)$ be profit from using inputs αz at price $p + \varepsilon$.

$$\pi(\alpha, p + \varepsilon) = (p + \varepsilon) f(\alpha z) - \alpha w z$$

Now at $\alpha = 1$ derivative equals

$$\pi'_\alpha(1, p+\varepsilon) = (p+\varepsilon)(D_z f(z))z - wz = \varepsilon(D_z f(z))z + z[p(D_z f(z)) - w]$$

But $(D_z f(z))z > 0$ since when $z_i > 0$, then $f_i = w/p > 0$ and $z_i > 0$ exists. Also $z[p(D_z f(z)) - w] = 0$ since either $z_i = 0$ or $pf_i = w_i$.

So there exists $\alpha > 1$ such that $q' = f(\alpha z) > f(z) = q$ delivers higher profit, and together with law of supply shows that output is increasing in price (here one only needs to assume that f is once continuously differentiable).

(b) *Using FOC.*

   i. *Using $p = c_q(w, q(p,w))$, so $1 = c_{qq} q_p(p,w)$, so (if q is differentiable) $q' = 1/c_{qq}$. By Proposition 3.2 this is non-negative. If c is twice continuously differentiable then by strict concavity of f, we have $c_{qq} > 0$ except at isolated points (i.e. given any (p,w) at which it is zero, it is strictly positive in a neighborhood around that point) and hence q is strictly increasing everywhere.*

   ii. *Using more elementary FOC.*
   First assume one input to see how the easy case works. FOC: $pf'(z(p)) = w$. Taking the derivative w.r.t. p and assuming f is differentiable, then if we knew that z is differentiable we would have $f'(z(p)) + pf''(z(p))z'(p) = 0$, so $z' = -f'/(pf'') > 0$ by strict concavity of f.
   In general $pD_z f(z(p)) = w$, where w is now a vector, as is $D_z f$. Similar to before, we get $D_z f + pD^2_{z^2} f \cdot D_p z = 0$. So $D_p z = \left(D^2_{z,z} f\right)^{-1} D_z f\left(\frac{-1}{p}\right)$. The change in q is $D_z f \cdot D_p z = D_z f\left(D^2_{z^2} f\right)^{-1} D_z f\left(\frac{-1}{p}\right) > 0$ since $\left(D^2_{z^2} f\right)$ is ND and hence so is its inverse.

2. *Prove that demand for some input increases due to an increase in output price.*

   (a) *Since output is strictly increasing and marginal productivities are non-negative, some input has to increase.*

   (b) *Equivalently, consider step b.ii. above and note that since $D_z f \cdot D_p z > 0$ and $D_z f > 0$ it cannot be that $D_p z \leqslant 0$.*

3. *Prove that an increase in input price leads to decreased demand for that input.*

   (a) *By the law of supply this is immediate, as in 1.a.*

   (b) *By FOC of part 1.b.ii above we get $pD_z f = w$. Differentiating system wrt w_k gives $p(D_{zz} f)(D_{w_k} z) = e_k$ where e_k is vector of zeroes except with a one on k-th coordinate. So $(D_{w_k} z) = p^{-1}(D_{zz} f)^{-1} e_k$.*
   Since $D_{zz} f$ is ND, its inverse is as well, and in the k-th row which corresponds to $\frac{\partial z_k}{\partial w_k}$, e_k picks up a diagonal element and so $(D_{zz} f)^{-1} e_k$ is negative.

Exercise. *After we study Topkis' theorem come back to this question and apply that method. Hint: Use aggregation as we will do in the notes on Topkis' theorem. The trick is how to aggregate out all the other inputs, say $2, \ldots, n$. Further hint: Define a function that holds fixed the amount of, say, z_1, and finds the profit maximizing solution for that amount of z_1. Now use that to derive a two-part optimization problem: for each z_1 maximize according to that function; then choose z_1 optimally. This will enable you to use the univariate version of Topkis' theorem.*

Remark 3.4. *Convexity of π implies that firms maximizing expected profits will like price uncertainty.*

Exercise. *In MWG, read through example 5.C.1 and solve Exercises 5.C.9.(c) and 5.C.10.(c). (If you find this difficult you may want to make sure you can solve parts a and b in these questions as well; this is not required and should not be handed in.)*

3.e Some pictures: total, average and marginal cost, and resulting supply

You should read Section 5.D in MWG - I will only be going over one example of the many for which they present figures describing the relationship between technology, f (or Y) and the total, average and marginal cost functions at fixed input prices.

Holding the input price vector fixed at $\bar{w}$, abuse notation and write $c(q) = c(q, \bar{w})$ as the minimal cost of producing output level q. The FOC of profit maximization (see FOC in Proposition 3.3) imply that the firm produces either $0, \infty$ (i.e., there is no solution) or q^* where marginal cost equals price: $MC(q^*) = c'(q^*) = p$ (recall that p is the price of the output and is not being held fixed). Whether or not the firm produces is determined by whether the (total and hence also per-unit) revenue covers the cost, so the firm will produce a non-zero quantity q only if the average cost is no more than the price: $AC(q) = \frac{c(q)}{q} \leqslant p$. For this reason it is useful to understand the MC and AC functions.

We will focus on the one-input case, where the relationship between f and Y is as described above. We will assume the price of the one input is 1 so the costs to produce $q = f(z)$ is $z = f^{-1}(q)$. The inverse of a function in a diagram can be seen by switching the axes (or turning the mirror-image 90 degrees around the origin). Then $AC(q)$ is the slope of the line from the origin to the cost function evaluated at q, and $MC(q)$ is the slope of the cost function at q. So consider the case (Figure 5.D.4 in MWG) of a convex technology with an initial fixed cost. Zero production costs zero, but any arbitrarily small positive amount requires the fixed cost and hence the cost function is discontinuous at zero. The marginal cost is not well defined ('infinite') at 0, but right after 0 starts at the slope of the cost function (which may or may not be zero) and increases (convexity implies increasing MC). The average cost is undefined ('infinity') at 0, but converges to infinity as production converges to zero, decreases until its minimum which is reached as it crosses the MC curve.

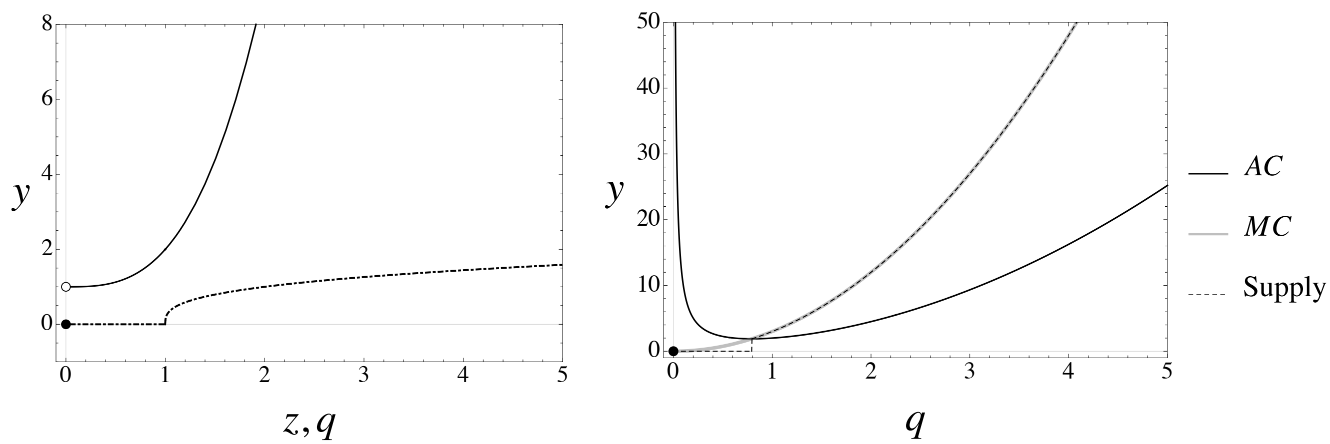
Remark 3.5. *These more general features are immediate:* $\min \frac{c(q)}{q}$ *satisfies* $\frac{c'(q)}{q} - \frac{c(q)}{q^2} = 0$ *so* $c'(q) = \frac{c(q)}{q}$.

So the firm's optimal decision is to supply according to MC if price is higher than AC. If the fixed costs are sunk we consider the average variable cost: the average cost ignoring the fixed cost; sunk costs are irrelevant to the firm's decision, and the supply becomes MC again.

If f is $z^{\frac{1}{3}}$ with a fixed cost of 1 unit then in the LHS figure the dotted-and-dashed function is f; $f(z) = 0$ if $z < 1$ and $(z-1)^{\frac{1}{3}}$ for $z \geqslant 1$. The solid function (cut off for values above 8 for convenience) is c; $c(q) = q^3 + 1$ for $q > 0$, and $c(0) = 0$. In the RHS the U shaped curve (solid black, with the origin added) is the AC; $q^2 + \frac{1}{q}$ for $q > 0$ (and 0 otherwise) and the upward sloping line is MC, $3q^2$. The supply correspondence equals the MC wherever it is weakly above AC, and zero whenever it is weakly below (drawn as a dashed curve).

Exercise. *(Trivial) What values does the supply correspondence take when $q = 1$?*

If the cost is sunk then in the left graph the solid function, c, includes the point $(0, 1)$ instead of the origin. The average variable cost is $AVC = q^2$. As it is always below the MC we have that with sunk costs the supply function is the dashed line.



3.f Long-run and short-run costs

Recall the discussion when we studied the envelope theorem in consumption regarding the second-order envelope theorem and short-run vs long-run decisions. To remind you the former vs. the latter are optimization decisions made when some choice variables are held constants. when they are not.

When there is more than one input one can study the long-run cost and supply when all inputs can be freely changed, and the short-run cost and supply when one of the inputs, say z_n, is held fixed. The short- and long-run cost curves (and hence the average cost curves) will coincide at supply level q iff the fixed level of z_n, say $\bar{z}_n$, equals the long-run optimal level for producing q; otherwise the short-run curves (as it has less flexibility) will lie above the long-run curves.

To clarify this consider a production function $f : \mathbb{R}^n \to \mathbb{R}$, and let

$$z^{LR*}(w, q) = \arg\min_{z \in \mathbf{R}^n} z \cdot w \text{ s.t. } f(z) \geqslant q$$

denote the long-run cost-minimizing quantity of input i at input prices w and at quantity q. Let $\bar{z}_n = z_n^*(\bar{w}, \bar{q})$ be a fixed level of input 1, that would be optimal at $\bar{w}, \bar{q}$, and denote the short-run demands for inputs by

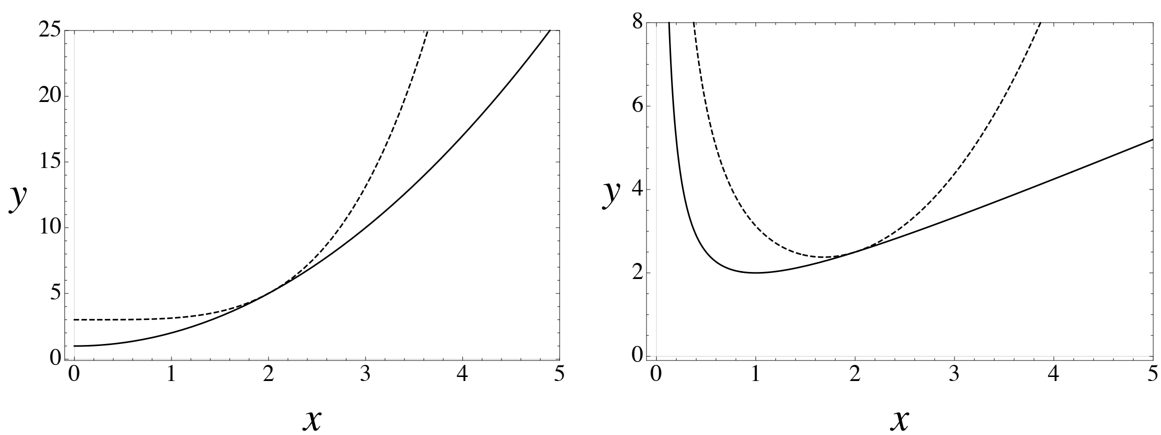
$$z^{SR*}(w, q, \bar{z}_1) = \arg \min_{\{z \in \mathbf{R}^n : z_n = \bar{z}_n\}} z \cdot w \text{ s.t. } f(z) \geqslant q$$

Finally, denote by $C^{LR}(w, q)$ and $C^{SR}(w, q)$ the long-run and short-run total costs of producing q. That is, $C^{LR}(w, q) = z^{LR*}(w, q) \cdot w$ and similarly for SR $C^{SR}(w, q; \bar{z}_n) = z^{SR*}(w, q, \bar{z}_n) \cdot w$.

Then obviously $C^{LR}(\bar{w}, \bar{q}) = z^*(\bar{w}, \bar{q}) \cdot \bar{w} = \bar{z}_n \bar{w}_n + \sum_{i=1}^{n-1} z_i^*(\bar{w}, \bar{q}) w_i = C^{SR}(\bar{w}, \bar{q})$. But at any other prices or quantities we will have $C^{LR}(w, q) = z^*(w, q) w \leqslant \bar{z}_n \bar{w}_n + \sum_{i=1}^{n-1} z_i^*(w, q) w_i = C^{SR}(w, q, \bar{z}_n)$. Thus the long-run cost curve is a lower envelope of the short-run cost curves, and they are tangent at the input-prices and quantity at which the constrained quantity $\bar{z}_n$ is optimal.

The solid curve on the LHS is total long-run costs and on the RHS is AC. The dashed curves are tangent at $\bar{q} = 2$ where $z_n(w, \bar{q}) = \bar{z}_n$ and on the LHS is the short-run total cost and on the RHS is the short-run AC.

The fact that the short-run curve is "more convex" than the long-run curve is again the second-order envelope theorem and gives us the result developed in the next subsection that short-run reactions are less pronounced than long-run reactions to changes in parameters.



3.f.1 The LeChatelier principle

The principle states that due to a change in parameters the long-run response is greater than the short-run response. Consider for example profit maximization with n inputs, and assume there is a unique solution at all prices. We will argue that z_i^{SR*} is less sensitive to changes in the parameter w_i than is z_i^{LR*}. We argue this for general production sets. Note that this just repeats what we did in the discussion of the second-order envelope theorem.

For some $\bar{\mathbf{p}} \in \mathbf{R}_+^n$ let $\bar{y}_n = \left(\arg\max_{y \in Y} \mathbf{p} \cdot y\right)_n$, i.e., $\bar{y}_n$ is the profit maximizing amount of good n at prices $\bar{p}$. We will consider the case where the level of y_n is fixed in the short run at $\bar{y}_n$ but changeable in the long run, for example n is the land used in production and in the short-run land cannot be bought or sold because its use needs to be changed in local regulations, or it is labor and due to certain regulations or union agreements labor cannot be changed without significant advance notice.

Assume that only one other price changes, wlog, p_i, and all others are fixed at $\bar{p}_{-i}$. Denote

$$\begin{aligned}
\pi^L(p_i) &= \max_{y\in Y}(p_i, \bar{p}_{-i}) \cdot y \\
y^L(p_i) &= y^L(p_i, \bar{p}_{-i}) \\
\pi^S(p_i; \bar{y}_n) &= \max_{(y_{-n}, \bar{y}_n)\in Y}(p_i, \bar{p}_{-i}) \cdot y \\
y^S(p_i; \bar{y}_n) &= \arg\max_{(y_{-n}, \bar{y}_n)\in Y}(p_i, \bar{p}_{-i}) \cdot y
\end{aligned}$$

Here π^L is the long-run profit function – when we can change all goods – as a function of p_i and y^L is the long-run profit-maximizing choice of goods – when we can change all goods – as a function of p_i. Similarly $\pi^S(p_i; \bar{y}_n)$ is the short-run profit maximizing function – when we can *not* change the amount of good n from $\bar{y}_n$ – as a function of p_i and $\bar{y}_n$, and $y^S(p_i; \bar{y}_n)$ is the short-run profit maximizing choice of goods – when we can*not* change the amount of good n from $\bar{y}_n$ – as a function of p_i and $\bar{y}_n$.

We argue that y_i^L changes (weakly) more than does y_i^S when p_i changes.

By convexity of π and the second-order envelope theorem, we know $0 \leqslant \frac{\partial^2 \pi^S}{\partial p_i^2} \leqslant \frac{\partial^2 \pi^L}{\partial p_i^2}$. Which we can rewrite as $\frac{\partial\left(\frac{\partial \pi^S}{\partial p_i}\right)}{\partial p_i} \leqslant \frac{\partial\left(\frac{\partial \pi^L}{\partial p_i}\right)}{\partial p_i}$. By Hotelling's lemma, we can conclude $\frac{\partial y_i^S}{\partial p_i} \leqslant \frac{\partial y_i^L}{\partial p_i}$.

4 Invertibility/integrability

We now turn to integrability results: how to define the production technology from cost or profit functions that satisfy the properties of the above propositions. As with utility, we can never deduce non-convex aspects of production. Mathematically, that is because these results are versions of the result that convex sets are defined (characterized) by their supporting hyperplanes. (The symbol $\gg$ means in all components when vectors are compared.)

Proposition 3.4. *1. If the upper contour sets of a production function are convex then*

(a) $Y = \{(-z, q) : w \cdot z \geqslant c(w, q) \forall w \gg 0\}$.

(b) Alternatively, $f(z) = \max\{q : w \cdot z \geqslant c(w, q) \text{ for all } w \gg 0\}$

2. If Y is closed, convex and satisfies free disposal then $Y = \{y \in \mathbf{R}^n : \mathbf{p} \cdot y \leqslant \pi(\mathbf{p}) \forall \mathbf{p} \geqslant 0, \mathbf{p} \neq 0\}$.

To see the intuition for 1b in Proposition 3.4 assume that for some $\hat{w}$ we have $\hat{w} \cdot z < c(\hat{w}, q)$. Then it cannot be that we can produce q with z since if we could $c(\hat{w}, q)$ would be lower than it is. So we ask what is the maximal q we can get.

Exercise. *Think through the intuition and proof for part 2 in Proposition 3.4.*

5 Aggregation (positive) and Efficiency (normative)

The summary statement is that as there are no wealth effects we get positive aggregation and that profit-maximizing production is efficient. To develop these results assume there are J firms. Each firm j has technology described by Y_j that determines a "supply" correspondence $y_j(\mathbf{p})$, the solution to the profit-maximization problem, and the resulting profit function $\pi_j(\mathbf{p})$. The aggregate supply is $y(\mathbf{p}) = \sum_{j \in J} y_j(\mathbf{p})$, and the aggregate profit is $\pi(\mathbf{p}) = \sum_{j \in J} \pi_j(\mathbf{p})$.

5.a Positive representative firms

We now show that this aggregate supply corresponds to a positive representative producer. This is plausible since part 6 of Proposition 3.3 implies that for each j, $Dy_j(\mathbf{p})$ is a symmetric and positive semidefinite matrix with $Dy_j(\mathbf{p})\mathbf{p} = 0$, and this carries over to $y = \sum y_j$.Hence this necessary condition is satisfied for aggregate supply.

Let $Y = \sum Y_j$, let $y^*(\mathbf{p})$ be the supply correspondence of the hypothetical representative producer with this technology, and $\pi^*(\mathbf{p})$ the profit obtained by this producer. Then I claim that $\pi^*(\mathbf{p}) = \pi(\mathbf{p})$ and $y^*(\mathbf{p}) = y(\mathbf{p})$. That the representative producer obtains at least the sum of the profits of the individuals is obvious: she can mimic the individual producers and chose $\sum y_j(\mathbf{p}) \in \sum Y_j$. But the opposite is also true: she cannot obtain strictly more, because if so it is with some $\hat{y} \in \sum Y_j$, which means that $\hat{y} = \sum \hat{y}_j$, but if $\hat{y}$ obtains more profit than $\sum \pi_j(\mathbf{p})$ then at least one producer does better choosing $\hat{y}_j(\mathbf{p})$ than $y_j(\mathbf{p})$ which is a contradiction. A similar argument shows the equivalence between the aggregate supply correspondence and the supply correspondence of the representative firm.

5.b Normative representative firms? No! But efficiency? yes

So if a price change increases the profits of all firms, it also increases the profits of the representative firm. But we do not really care in a normative sense about firm profits as they only determine consumer welfare indirectly, and therefore to analyze the welfare implication we should anyway study how the change in profit is distributed and how consumers are affected. So such a conclusion is not interesting per se, in general.

However, if consumers have preferences that are quasilinear in money, it is of interest. Then the allocation of money is irrelevant for efficiency. So, we would be interested in finding situations where consumer plus producer surplus is maximized, and producer surplus is profits. There are situations where maximizing the sum of producer profits is of interest. We do not focus on that, and study the more general case.

There is an important aspect of the representative firm that does have significant normative implications. In the single-output case it implies that the representative firm is cost minimizing, so that each firm acting on its own results in total production being cost minimizing – there is no reallocation of inputs among firms that could lower their costs even if they integrated. Thus, production resources are not being wasted, and production is "efficient." (If resources were wasted due to aggregation problems we could conclude that it is bad for consumers, since we could use the resources more efficiently to produce more goods that consumers liked.)

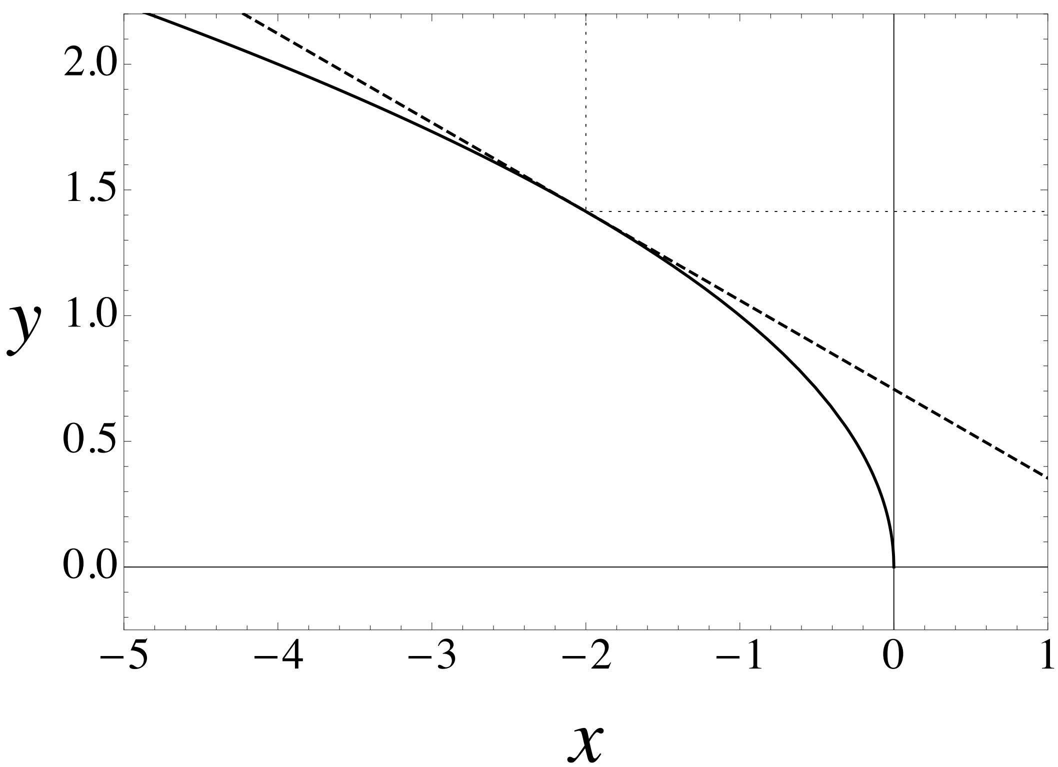
More generally, define production of y to be efficient if there is no feasible alternative that is better, that is, no $y' \in Y$ with $y' > y$ (strictly greater in some component). Then we can easily see the first welfare theorem in this context: if y is profit maximizing for some $\mathbf{p} \gg 0$ then it is efficient. (Proof by contradiction: otherwise y' would have higher profits; note the use in this argument of the fact that $\mathbf{p}_i > 0$ for all goods i.)

We can also see (albeit less easily) the second welfare theorem: If Y is closed and convex then for any efficient $y \in Y$ there exists a price $\mathbf{p} > 0$ (greater in some component) such that y is profit maximizing at p. The proof uses (once again) the separating hyperplane theorem.

Fix an efficient $y^* \in Y$ and consider the disjoint convex sets Y and $P_{y^*} = \left\{y \in \mathbb{R}^2 : y \gg y^*\right\}$. By the separating hyperplane theorem there exists $\mathbf{p} \neq 0$ s.t.

$$\mathbf{p}y' \geqslant \mathbf{p}y'' \qquad \forall y' \in P_{y^*}, \forall y' \in Y, \tag{17}$$

in particular for $y'' = y^*$. Consider a sequence y_n in P_{y^*} with $y_n \longrightarrow y^*$. Then we have $\mathbf{p}y_n \geqslant \mathbf{p}y''$ hence $\mathbf{p}y^* \geqslant \mathbf{p}y''$ for all $y'' \in Y$ so y^* is profit maximizing at prices $\mathbf{p}$. It remains to show that $\mathbf{p} \geqslant \mathbf{0}$. But if $p_j < 0$ then let $\hat{y}_k = y^*_k + \varepsilon$ for ε small and $k \neq j$ and $\hat{y}_j = y^*_j + L$ for L large. Then $\mathbf{p}(\hat{y} - y^*) = \varepsilon \sum_{k \neq j} p_k + Lp_j$ which for L large enough is negative, contradicting eq. (17).



6 Monopoly

6.a FOC

We will not study much about monopoly, but as it is a very useful example which we will use for the comparative statics section coming up next we will review some basics. The difference between monopoly and the competitive firm we studied until now is that the monopolist knows that the price it sets determines the quantity it will sell (or the quantity it sells will determine the price at which it is sold).

So let $p(q)$ denote the inverse demand function (price as a function of quantity) resulting from some aggregate demand. Similarly – with some significant (but hopefully not confusing) abuse of notation – let $q(p)$ denote the demand function.

The monopolist solves

$$\max_q p(q)q - c(q)$$

with FOC

$$p'(q^*)q^* + p = c'(q^*)$$

where the LHS is the marginal revenue and the RHS is the marginal cost.

Equivalently the monopolist solves

$$\max_p q(p)\, p - c(q(p))$$

with FOC

$$p^* q'(p^*) + q(p^*) = c'(q(p^*)) q'(p^*)$$

Rewriting the last FOC we get

$$\frac{p - c'}{p} = -\frac{q}{q'}\frac{1}{p} = -\left(q'\frac{p}{q}\right)^{-1} = \frac{1}{\varepsilon}$$

where ε is the elasticity of the demand function.

Exercise (Truly trivial)**.** *Show that the two FOCs coincide (i.e., $q^* = q(p^*)$).*

6.b SOC

While so far we didn't focus on SOC, in the case of monopoly we see that conditions under which $q(p)p - c(q(p))$ is quasiconcave in p will depend on the demand function $q(p)$. We now ask, for constant marginal cost, what assumption on q (a non primitive!) is sufficient for $q(p)p - cq(p)$, to be quasiconcave in p.

Proposition 3.5. *$q(p)p - cq(p)$ is quasiconcave in p if $1/q(p)$ is convex (over the region where $q(p) > 0$).*

Proof. Let $f(p) = q(p)p - cq(p)$. Clearly f is increasing when $p < c$. So a failure of quasiconcavity implies there exists $p_1 > p_0 \geqslant c$ and $\lambda \in (0, 1)$ with $p_\lambda = \lambda p_0 + (1 - \lambda)p_1$ such that $f(p_\lambda) < \min\{f(p_1), f(p_0)\}$, that is, for $i = 0, 1$:

$$(p_i - c)q(p_i) > (p_\lambda - c)q(p_\lambda)$$

Rewriting it gives

$$\frac{p_i - c}{q(p_\lambda)} > \frac{p_\lambda - c}{q(p_i)}$$

So,

$$\begin{aligned}
\lambda\frac{p_0 - c}{q(p_\lambda)} + (1-\lambda)\frac{p_1 - c}{q(p_\lambda)} &> \lambda\frac{p_\lambda - c}{q(p_0)} + (1-\lambda)\frac{p_\lambda - c}{q(p_1)} \\
\iff \frac{p_\lambda - c}{q(p_\lambda)} &> \lambda\frac{p_\lambda - c}{q(p_0)} + (1-\lambda)\frac{p_\lambda - c}{q(p_1)} \\
\iff \frac{1}{q(p_\lambda)} &> \frac{\lambda}{q(p_0)} + \frac{(1-\lambda)}{q(p_1)}
\end{aligned}$$

Clearly if $1/q(p)$ is convex this cannot happen. ■

Proposition 3.6. *$1/f$ is convex if $\ln f$ is concave (i.e., f is log concave).*

Exercise. *Prove this result.*

Thus log concavity of demand is sufficient for profits to be quasiconcave in price. Log concavity of demand will arise if we have a population of consumers with quasilinear preferences and their willingness to pay is distributed according to a log-concave distribution. If willingness to pay is distributed F then demand at price p is just $1 - F(p)$. If the density f is log-concave and continuously differentiable then F and $1 - F$ are log-concave (see Theorems 1 and 3 in Bagnoli and Bergstrom (2005), "Log-concave probability and its applications", Economic Theory 26, 445–469). "Most" typical distributions are log-concave (see Table 1 in Bagnoli and Bergstrom (2005)).

6.c A comparative static

We now consider one comparative static (to which we return later).

Example 3.2. *Assume a monopolist's cost "decreases". In what sense must it decrease so that q will increase? Since what enters the FOC is the marginal cost a natural conjecture would be that if we compare the optimal quantity for two cost functions c_1 and c_2 where $c_2'(q) > c_1'(q)$ for all q we will get a lower optimal q with costs c_2 than with c_1. We will assume there is a unique solution to the firms cost-minimization problem with either cost function.*

Proof. Define q_i^* to be the optimal q with cost function c_i^*. By revealed preference, i.e., that q_i is optimal, we have

$$\begin{aligned} p\left(q_1^*\right) q_1^* - c_1\left(q_1^*\right) &\geqslant p\left(q_2^*\right) q_2^* - c_1\left(q_2^*\right) \\ p\left(q_2^*\right) q_2^* - c_2\left(q_2^*\right) &\geqslant p\left(q_1^*\right) q_1^* - c_2\left(q_1^*\right) \end{aligned}$$

so, subtracting,

$$\left(c_2\left(q_1^*\right) - c_2\left(q_2^*\right)\right) - \left(c_1\left(q_1^*\right) - c_1\left(q_2^*\right)\right) \geqslant 0.$$

Since the cost minimization problems have unique solutions each c_i is differentiable (see above) so $c_i\left(\hat{q}\right) - c_i\left(\bar{q}\right) = \int_{\bar{q}}^{\hat{q}} c_i'\left(q\right) \mathrm{d}q$. Applying this to the above equation we get

$$\int_{q_2^*}^{q_1^*} \left(c_2'\left(q\right) - c_1'\left(q\right)\right) \mathrm{d}q \geqslant 0$$

which, since $c_2' > c_1'$ implies $q_1^* \geqslant q_2^*$. ■