KC Border
Fall 2020

Lecture 2: Profit Maximization

2.1 Digression: Maximization

My on-line notes on optimization [1] cover the mathematics of optimization in one dimension, including the following topics.

- Local vs. global maxima and minima.
- Strict extrema.
- First order necessary conditions for interior extrema.
- First order necessary conditions at a boundary.
- Second order necessary conditions.
- Second order sufficient conditions.
- Taylor's Theorem and n^{th}-order conditions.
- Concave and convex functions.

2.2 Maximization and comparative statics

Just as above, our "equilibrium" conditions are often the results of some maximizing behavior. Consider this simple model of a firm. When the firm produces the level $y \geqslant 0$ of output, it receives revenue $R(y)$ and incurs cost $C(y)$. The profit is then $R(y) - C(y)$. In addition, it pays an *ad rem* tax ty. It seeks to maximize its after-tax profit:

$$\text{maximize } R(y) - C(y) - ty.$$

Let $y^*(t)$ solve this problem. What do we know?

$$R'(y^*) - C'(y^*) - t = 0$$

(or does it?). This does not tell us much about data that we might observe, but let's see how y^* changes with t:

$$R'\Big(y^*(t)\Big) - C'\Big(y^*(t)\Big) - t = 0 \quad \text{for all } t.$$

Therefore, by differentiating both sides with respect to t we get

$$\Big[R''\Big(y^*(t)\Big) - C''\Big(y^*(t)\Big)\Big] y^{*\prime}(t) - 1 = 0$$

or

$$y^{*\prime}(t) = \frac{1}{R''\big(y^*(t)\big) - C''\big(y^*(t)\big)}.$$

How can we sign this? The answer is, via the second order conditions. Namely,

$$R''(y^*) - C''(y^*) \leqslant 0.$$

This implies that

$$y^{*\prime}(t) < 0.$$

What if $R''(y^*) - C''(y^*) = 0$**?** Then the fraction has zero in the denominator, which means that $y^{*\prime}(t)$ does not exist!

The Implicit Function Theorem guarantees that if $R''(y^*) - C''(y^*) > 0$, then $y^*(t)$ is unique and differentiable. But more on that later.

2.2.1 Revenue maximization

What if the firm maximizes after-tax revenue $R(y) - ty$ instead of profit. The first order condition is

$$R'(y) - t = 0$$

and the second order condition is

$$R''(y) \leqslant 0.$$

(Note that I have used the economists' sloppy notation of omitting the *. I should actually use something different, since it is a different function.)

Differentiating the first order condition with respect to t yields

$$R''(y)y' - 1 = 0$$

or

$$y' = \frac{1}{R''} < 0,$$

where the inequality follows from the strict second order condition.

Thus a change in an *ad rem* tax gives us no leverage on deciding whether a firm after-tax maximizes revenue or after-tax profit.

2.2.2 Wages

Suppose that the firm's costs C are a function both of its level of output and a wage parameter, and assume that the partial derivative $D_y C > 0$ (which must be the case if the firm is minimizing costs).

For profit maximization,

$$\underset{y}{\text{maximize}}\, R(y) - C(y; w),$$

the first order condition is

$$R'(y) - D_y C(y; w) = 0$$

for an interior solution (marginal cost = marginal revenue), and the second order condition is

$$R''(y) - D_y^2 C(y; w) \leqslant 0.$$

Letting $y^*(w)$ be the maximizer we see that

$$h(w) = R'\Big(y^*(w)\Big) - D_y C\Big(y^*(w); w\Big) = 0$$

for all w. Thus h is constant so $h' = 0$. By the **chain rule**

$$h'(w) = R''\Big(y^*(w)\Big)y^{*\prime}(w) - D_y^2 C\Big(y^*(w); w\Big)y^{*\prime}(w) - D_{yw} C\Big(y^*(w); w\Big) = 0.$$

Solving for $y^{*\prime}$ gives

$$y^{*\prime}(w) = \frac{D_{yw} C\Big(y^*(w); w\Big)}{R''\Big(y^*(w)\Big) - D_y^2 C\Big(y^*(w); w\Big)}.$$

The denominator must be negative, so the sign of this is the opposite of the sign of the mixed partial $D_{yw} C$. Hmmm! We'll get back to this.

For revenue maximization,

$$\underset{y}{\text{maximize}}\, R(y),$$

the first order condition is

$$R'(y) = 0$$

for an interior solution, and the second order condition is

$$R''(y) \leqslant 0.$$

Letting $\hat{y}(w)$ be the maximizer we see that it is independent of w! Thus $\hat{y}'(w) = 0$!

An application to sports economics

What is the effect of player salaries on ticket prices?

For a profit-maximizing sports franchise (and one visit to Dodger Stadium ought to convince you that profits are being fiercely pursued), the revenue comes from ticket sales, parking, and concessions, but the costs are almost entirely determined by players' (and coaches' and groundskeepers') wages and utility bills for the lights, all of which do not depend on how many tickets are sold. The number of tickets sold will depend on the price charged, so the revenue is not going to be a linear function of the number of tickets sold.

A reasonable approximation to profit is

$$\text{profit} = R(y) - C(w),$$

where y is the number of tickets sold. (Parking and concessions tend to be proportional to the number of tickets.) There is also TV revenue, which does not depend on y, but can be treated as an additive constant. While some costs (free bobble heads, programs, etc.) are proportional to tickets they are small, and could also be netted out of the ticket revenues. We can see that w has no effect on y, so it cannot affect the ticket price.

What about second-order effects—higher wages attract better players, and so increase demand for tickets, enabling the franchise to sell the same number of tickets at a higher price. This works only if higher wages are limited to one team that is able to attract all the good players—if all teams' wages go up, there is no reason to expect any one team to get better.

What about third-order effects—if wages are too high the team will fold and then there will be no tickets available at any price. This might be more convincing if team prices were lower, but according to Forbes, as of September 2020, NFL franchises were worth on average over \$3 billion, ranging from from \$2 billion (Cincinatti Bengals) to \$5.7 billion (Dallas Cowboys). (See Forbes' list).

If, as the owners usually claim come time to negotiate with players, teams are such money losers, then why are team prices so high? For one thing, teams are a good tax shelter. A new owner can assign 80% of the value to player contracts and depreciate them over four or five years, then resell the team for largely capital gains. They are also frequently real cash cows. And then there are some special accounting practices that allow profits to be counted as expenses. (As when the the owner's son-in-law is given a fancy title without any particular value-producing responsibilities. Movie studios do this too.)

If you are interested in the economics of professional sports, I highly recommend *Pay Dirt* by my former colleagues Jim Quirk and Rod Fort [3].

2.3 After tax profit revisited

$$\underset{y}{\text{maximize}}\, R(y) - C(y) - ty.$$

Let y^* solve this problem. Last time we used the second order conditions to conclude

$$\frac{d}{dt} y^*(t) < 0,$$

provided the derivative exists.

But we got stuck when it came to dealing with wages. In that case

$$\operatorname{sgn} \frac{d}{dw} \hat{y}(w) = -\operatorname{sgn} D_{yw} C(y, w).$$

For this we can use another approach.

2.4 A lemma

2.4.1 Proposition *Let X and P be open intervals in $\boldsymbol{R}$, and let $f\colon X \times P \to \boldsymbol{R}$ be twice continuously differentiable. Assume that for all $x \in X$ and all $p \in P$,*

$$\frac{\partial^2 f(x,p)}{\partial p \partial x} \geqslant 0. \tag{1}$$

Let x^0 maximize $f(\cdot, p^0)$ over X and x^1 maximize $f(\cdot, p^1)$ over X. Then

$$(p^1 - p^0)(x^1 - x^0) \geqslant 0. \tag{2}$$

In other words, the sign of the change in the maximizing x is the same as the sign of the change in p.

If $\leqslant$ replaces $\geqslant$ in (1), then the sign of the change in x is the opposite of the sign of the change in p.

For minimization rather than maximization the sign of the effect is reversed.

Proof: By definition of maximization, we have

$$f(x^0, p^0) \geqslant f(x^1, p^0) \quad \text{and} \quad f(x^1, p^1) \geqslant f(x^0, p^1).$$

"Cross-subtracting" implies

$$f(x^1, p^1) - f(x^1, p^0) \geqslant f(x^0, p^1) - f(x^0, p^0). \tag{$*$}$$

But

$$f(x^1, p^1) - f(x^1, p^0) = \int_{p_0}^{p^1} \frac{\partial f}{\partial p}(x^1, \pi)\, d\pi$$

and

$$f(x^0, p^1) - f(x^0, p^0) = \int_{p^0}^{p^1} \frac{\partial f}{\partial p}(x^0, \pi)\, d\pi.$$

So $(*)$ becomes

$$\int_{p^0}^{p^1} \frac{\partial f}{\partial p}(x^1, \pi)\, d\pi \geqslant \int_{p^0}^{p^1} \frac{\partial f}{\partial p}(x^0, \pi)\, d\pi,$$

or

$$\int_{p^0}^{p^1} \left(\frac{\partial f}{\partial p}(x^1, \pi) - \frac{\partial f}{\partial p}(x^0, \pi) \right) d\pi \geqslant 0.$$

Now we use the same trick of writing a difference as an integral of the derivative to get

$$\int_{p^0}^{p^1} \left(\frac{\partial f}{\partial p}(x^1, \pi) - \frac{\partial f}{\partial p}(x^0, \pi) \right) d\pi = \int_{p^0}^{p^1} \left(\int_{x^0}^{x^1} \frac{\partial^2 f}{\partial p \partial x}(\xi, \pi)\, d\xi \right) d\pi \geqslant 0.$$

By assumption $\frac{\partial^2 f}{\partial p \partial x} \geqslant 0$, so by the convention that $\int_a^b = -\int_b^a$, we conclude that if $p_1 > p_0$, then $x_1 \geqslant x_0$, and the conclusion follows. ∎

Application

So consider
$$f(y,t) = R(y) - C(y) - ty.$$
Then
$$\frac{\partial f(y,t)}{\partial t} = -y$$
so
$$\frac{\partial^2 f(y,t)}{\partial y \partial t} = -1 < 0,$$
so
$$\frac{d}{dt} y^*(t) < 0.$$

Application

Apply to revenue maximization.

Application

Consider $y = f(x)$, where x is an input that gets paid wage w. The profit maximization problem is
$$\underset{x}{\text{maximize}}\, pf(x) - wx.$$
Letting $g(x,w) = pf(x) - wx$, we have $\frac{\partial^2 g}{\partial x \partial w} = -1$, so $dx^*/dw < 0$.

Letting $g(x,p) = pf(x) - wx$, we have $\frac{\partial^2 g}{\partial x \partial p} = f'(x)$, which is presumably positive, so $dx^*/dp > 0$.

2.5 ⋆ Supermodularity

If inequality $(*)$ holds whenever $x_1 > x_0$ and $p_1 > p_0$, we say that f exhibits **increasing differences**, a property related to what we now call **supermodularity**. To define this, we first need to define a **lattice**.

2.5.1 Definition *A* ***partial order*** $\succeq$ *on a set* X *is a binary relation that is transitive, reflexive, and antisymmetric. A* ***lattice*** *is a partially ordered set* $(X, \succeq)$ *with the property that every pair* $x, y \in X$*, has a least upper bound* $x \vee y$ *(also called the* ***join****) and a greatest lower bound* $x \wedge y$ *(also called the* ***meet****).*

For now, the most important example of a lattice is $\boldsymbol{R}^{\mathrm{n}}$ with the coordinatewise ordering $\geqq$, where $x \geqq y$ if $x_i \geqslant y_i$ for each $i = 1, \dots, n$.

2.5.2 Definition *A real-valued function* f *on a lattice is* ***supermodular*** *if*
$$f(x \wedge y) + f(x \vee y) \geqslant f(x) + f(y).$$

Proposition 2.4.1 can be restated as follows.

2.5.3 Proposition *If f is a twice differentiable function on $(\boldsymbol{R}^{\mathrm{n}}, \geqq)$, then f is supermodular if and only if for $i \neq j$*

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \geqslant 0.$$

That is, an increase in i increases the marginal increase of j. That is i and j are **complements**.

We might discuss this in greater depth later. See my on-line notes [2].

2.6 Appendix: A note on the chain rule

Above, I used the following version of the chain rule.

2.6.1 Proposition (An application of the chain rule) *Let $f_1, \ldots, f_m$ be differentiable functions on an open interval I of $\boldsymbol{R}$. These define a function $f\colon I \to \boldsymbol{R}^{\mathrm{m}}$ by $f(t) = \big(f_1(t), \ldots, f_m(t)\big)$. Let U be an open set in $\boldsymbol{R}^{\mathrm{m}}$ and let $g\colon U \to \boldsymbol{R}$ have partial derivatives and assume that $f(t)$ lies in U for each $t \in I$. Then the composition $h = g \circ f$ defined by*

$$h(t) = g\big(f(t)\big) = g\big(f_1(t), \ldots, f_m(t)\big)$$

is differentiable and

$$h'(t) = \sum_{i=1}^{m} D_i g\big(f(t)\big) f_i'(t).$$

If the functions are twice differentiable, then

$$\begin{aligned}
h''(t) &= \frac{d}{dt} \sum_{i=1}^{m} D_i g\big(f(t)\big) f_i'(t) \\
&= \sum_{i=1}^{m} \frac{d}{dt} D_i g\big(f(t)\big) f_i'(t) \\
&= \sum_{i=1}^{m} \left\{ \left(\sum_{j=1}^{m} D_{ij} g\big(f(t)\big) f_j(t) \right) f_i'(t) + D_i(g\big(f(t)\big) f_i''(t) \right\} \\
&= \left(\sum_{i=1}^{m} \sum_{j=1}^{m} D_{ij} g\big(f(t)\big) f_j(t) f_i(t) \right) + \left(\sum_{i=1}^{m} D_i(g\big(f(t)\big) f_i''(t) \right)
\end{aligned}$$

Here is an important special case.

2.6.2 Corollary *Let $g\colon U \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}$ and let $x \in U$. Pick some $v \in \boldsymbol{R}^{\mathrm{m}}$ and define*

$$h(t) = g(x + tv).$$

Then

$$h'(0) = \sum_{i=1}^{m} D_i g(x) v_i$$

and

$$h''(0) = \sum_{i=1}^{m} \sum_{j=1}^{m} D_{ij} g(x) v_i v_j$$

References

[1] K. C. Border. 2000. Notes on calculus and maximization I.
http://www.its.caltech.edu/~kcborder/Notes/Max2.pdf

[2] ——— . 2001. Preliminary notes on lattices.
http://www.its.caltech.edu/~kcborder/Notes/Lattice.pdf

[3] J. P. Quirk and R. D. Fort. 1992. *Pay dirt*. Princeton, New Jersey: Princeton University Press.