Chapter 10: The Envelope Theorem

Elements of Decision: Lecture Notes of Intermediate Microeconomics 1

Charles Z. Zheng

Department of Economics, University of Western Ontario

Last update: January 1, 2019

1 Quantitative comparative statics

All the decision problems studied in this course are special cases of the following form:

$$\Phi(t) := \max_{x \in X(t)} \varphi(x,t), \tag{1}$$

where $\varphi(x,t)$ denotes the objective, x the choice variable, t the parameter, $X(t)$ the choice set from which x has to be chosen, and $\Phi(t)$ the maximand, the maximum value of the objective among all x given parameter t. For example, in a firm's output quantity decision studied in Chapter 1, $\varphi(x,t)$ corresponds to the firm's profit, with x being the output quantity, t the output price, and $X(t)$ the set $\mathbb{R}_+$ of output quantities. (In this example, the choice set $X(t)$ is independent of the parameter t.) For another example, in a consumer's decision (Chapters 6–9), $\varphi(x,t)$ corresponds to the consumer's utility function, x the consumption bundle (x_1, x_2), t the configuration (p_1, p_2, m) of prices and income, and $X(t)$ the budget set $\{(x_1, x_2) \in \mathbb{R}^2_+ : p_1 x_1 + p_2 x_2 \leq m\}$. (In this example, the choice set $X(t)$ does depend on the parameter t.)

Previous chapters have explained techniques of directional comparative statics on various cases of the decision problem (1), "directional" in the sense of giving a directional answer to the question whether the optimal choice of x increases or decreases given a change in the parameter. This chapter, to offer a general perspective of previous chapters and to end this course with a glimpse of modern microeconomics, introduces a technique for quantitative comparative statics, "quantitative" in the sense of giving a quantitative answer to the question how much the maximand $\Phi(t)$ changes given a change in the parameter t. That is,

$$\frac{d}{dt}\Phi(t) = ? \tag{2}$$

The technique to answer this question is the modern envelope theorem, modern in the sense that it does not rely on much about the optimal choice of x, and that the technique is applicable not only to the decision problems studied here but also to the economics of incentives and information.

2 The envelope theorem when $X(t)$ is constant to t

To answer Question (2), look at Figure 1. There, regardless of the value of the paramter t, the choice set $X(t)$ consists of three elements, x', x'' and x'''. Comparing the vertical positions of $\varphi(x',t)$, $\varphi(x'',t)$ and $\varphi(x''',t)$ at each value of t, we see that the optimal x is x' when $t \in [0, t_*)$, x'' when $t \in (t_*, t^*)$, and x''' when $t > t^*$. Thus, if we denote $\tilde{x}(t)$ for the element in the choice set that maximizes the objective when the parameter is equal to t, then

$$\tilde{x}(t) = \begin{cases} x' & \text{if } t < t_* \\ x'' & \text{if } t_* < t < t^* \\ x''' & \text{if } t > t^*. \end{cases}$$

That is, the thickly drawn curve in Figure 1 is the graph of the maximand Φ. Note that the curve of Φ coincides with the curve of $\varphi(x', t)$ when x' is the optimum, coincides with the curve of $\varphi(x'', t)$ when x'' is the optimum, and coincides with the curve of $\varphi(x''', t)$ when x''' is the optimum. In other words, at any value of t except t_* or t^*, the slope of the curve of Φ is equal to the slope of the curve of $\varphi(x, t)$ such that x is the optimum $\tilde{x}(t)$ at the particular value of t. Since the slope of the curve

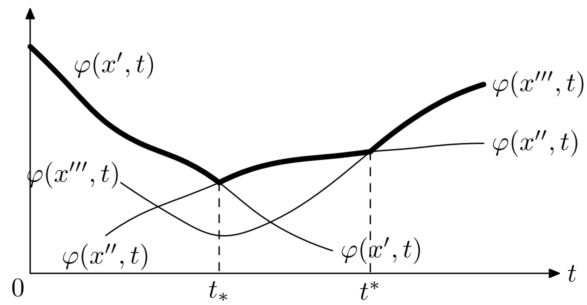

Figure 1: The envelope theorem

of Φ at t is equal to $\frac{d}{dt}\Phi(t)$, and the slope of the curve of $\varphi(x, t)$ is equal to the partial derivative $\frac{\partial}{\partial t}\varphi(x, t)$, we have $\frac{d}{dt}\Phi(t) = \frac{\partial}{\partial t}\varphi(x, t)$. Keep in mind that this x is not an arbitrary element of the choice set, but the optimal choice $\tilde{x}(t)$ given t. Thus we obtain the answer to Question (2):

$$\frac{d}{dt}\Phi(t) \overset{\text{a.e.}}{=} \frac{\partial}{\partial t}\varphi(x, t)\bigg|_{x=\tilde{x}(t)}. \tag{3}$$

Here the notation $\overset{\text{a.e.}}{=}$ reads "almost everywhere equal to," meaning that the equation holds for all values of t but a set of *measure zero*, measure zero in the sense that if one were blindfolded while putting a finger of his on the axis for t, there is zero chance for the finger to land on a t that violates the equation. That is, the equation is true with probability one. In Figure 1, for example, Eq. (3) holds unless t is t_* or t^*, but if one were to randomly select a point in the real line, the chance for t_* or t^* to be selected is zero. One might think that such measure zero property relies on the assumption that we have only finitely many elements in the choice set. But No! It turns out that Eq. (3) holds not only when the choice set is finite but also when it is infinite, and not only when the choice set is countable but also when it is uncountable. The proof of Eq. (3), i.e., the envelope theorem, beyond the scope of this course, is provided by Milgrom and Segel [1].

To illustrate the envelope theorem, recall a firm's input-output decision studied in Chapter 2,

$$\pi(w) := \max_{(x,y)\in\mathbb{R}^2_+ : y\le f(x)} py - wx,$$

and suppose that we want to know the effect of the parameter w (hence denoting the maximand profit as a function $\pi(w)$ of w). That is, the symbol w corresponds to t in (1), and the choice variable (x, y) here corresponds to the x in (1). Note that the choice set $\{(x, y) \in \mathbb{R}^2_+ : y \le f(x)\}$ is independent of this parameter w, hence Eq. (3) applies. Take the partial derivative of the objective $py - wx$ with respect to this parameter w to obtain

$$\frac{\partial}{\partial w}(py - wx) = -x.$$

Thus, by Eq. (3), if the firm is employing $\tilde{x}(w)$ units of the input when the input price is w, then

$$\frac{d}{dw}\pi(w) \overset{\text{a.e.}}{=} -\tilde{x}(w). \tag{4}$$

To appreciate this equation, imagine a firm in the states that produces cookware with steel imported from China. The Trump administration's trade war with China, through raising the tariff of the steel import from China, increases the input price w for this firm. If the firm is currently employing, in a profit-maximizing manner, $\tilde{x}$ units of Chinese steel, the firm in lobbying against the trade war could have provided a quantitatively specific argument by pointing to Eq. (4) and saying: "The trade war is reducing my firm's profit by about $\tilde{x}$ dollars for each dollar increase in the tariff against Chinese steel." The beauty of this argument is that it remains valid regardless of how the other parameter, the output price p, might be changing and how the firm might be adjusting its production in this process, as long as the firm is maximizing its profits.

3 When the choice set depends on the parameter

For the envelope theorem to be valid when the choice set $X(t)$ depends on the parameter t, we need more stringent conditions. Here is a case where the theorem still works. It is the case where the Lagrange method, partially explained in Chapter 4, is valid: Suppose that the decision problem (1) is equivalent to

$$\begin{aligned} V(t) := \quad \max\nolimits_{x\in\mathbb{R}^m} \quad & u(x,t) \\ \text{s.t.} \quad & g_1(x,t) \geq 0, \\ & \vdots \\ & g_n(x,t) \geq 0, \end{aligned}$$

and that the Lagrange method is applicable so that the original problem can be viewed as finding a *saddle point* $(\tilde{x}, \tilde{\lambda})$ such that $\tilde{x}$ maximizes the Lagrangian

$$L(x, \lambda_1, \ldots, \lambda_n, t) := u(x,t) + \sum_{k=1}^{n} \lambda_k g_k(x,t)$$

among $x \in \mathbb{R}^m$ given $\tilde{\lambda} := (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n)$, and $\tilde{\lambda}$ minimizes the Lagrangian among $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n_+$ given $\tilde{x}$.[1] Thus, the role of the objective $\varphi(x,t)$ in (1) is played by the Lagrangian here, with the choice variable x in (1) corresponding to the x here in maximizing the Lagrangian, and corresponding to λ here in minimizing it. Now the choice set becomes $\mathbb{R}^m$ for x and $\mathbb{R}^n_+$ for λ, constant to the parameter t, so we are back to the previous section. Furthermore, the maximand of $L(x, \tilde{\lambda}, t)$ among all x is equal to the maximand $V(t)$ of $u(x,t)$ among all x, because the condition for $\tilde{\lambda}$ requires that $\tilde{\lambda} g_k(\tilde{x}, t) = 0$ for all k at the solution for $(x, \lambda_1, \ldots, \lambda_n)$ (c.f. Chapter 4), rendering the Lagrangian equal to the original objective $u(x,t)$. Thus, if we denote $\left(\tilde{x}(t), \tilde{\lambda}(t)\right)$ for the solution given any value t of the parameter, then Eq. (3) becomes

$$\frac{d}{dt}V(t) \overset{\text{a.e.}}{=} \frac{\partial}{\partial t} L(x, \lambda_1, \ldots, \lambda_n, t)\Big|_{x=\tilde{x}(t);\lambda_k=\tilde{\lambda}_k(t):k=1,\ldots,n}. \tag{5}$$

[1] The notion of saddle point, and the condition under which the original problem becomes a saddle point problem, are beyond the scope of this course.

To illustrate (5), consider a consumer's decision studied in Chapters 6–9 such that his preference relation is smooth and monotone, exhibits diminishing MRS, and is represented by a differentiable utility function $u(x_1, x_2)$ for any nonnegative consumption bundle (x_1, x_2), with prices p_1 and p_2 for the two goods and income m being positive parameters. Then as explained in Chapters 6–7 the Lagrange method with equality constraints applies, with the Lagrangian

$$L(x_1, x_2, \lambda, p_1, p_2, m) := u(x_1, x_2) + \lambda (p_1 x_1 + p_2 x_2 - m). \tag{6}$$

Suppose that we are interested in the effect of income m. Take the partial derivative of the Lagrangian with respect to m to obtain

$$\frac{\partial}{\partial m} L(x_1, x_2, \lambda, p_1, p_2, m) = -\lambda.$$

Then Eq. (5) implies

$$\frac{\partial}{\partial m} V(p_1, p_2, m) \overset{\text{a.e.}}{=} -\tilde{\lambda}(p_1, p_2, m). \tag{7}$$

The question is What is the solution $\tilde{\lambda}(p_1, p_2, m)$ for λ equal to? The answer one can find by recalling one of the first-order conditions for maximizing the Lagrangian:

$$\frac{\partial}{\partial x_1} L = \frac{\partial}{\partial x_1} u(x_1, x_2) + \lambda p_1 = 0$$

when $(x_1, x_2) = \tilde{x}(p_1, p_2, m)$ and $\lambda = \tilde{\lambda}(p_1, p_2, m)$. Solve the above equation for λ to obtain

$$\tilde{\lambda}(p_1, p_2, m) = -\frac{1}{p_1} \left. \frac{\partial}{\partial x_1} u(x_1, x_2) \right|_{(x_1, x_2) = \tilde{x}(p_1, p_2, m)}. \tag{8}$$

Plug this into Eq. (7) to get

$$\frac{\partial}{\partial m} V(p_1, p_2, m) \overset{\text{a.e.}}{=} \frac{1}{p_1} \left. \frac{\partial}{\partial x_1} u(x_1, x_2) \right|_{(x_1, x_2) = \tilde{x}(p_1, p_2, m)}. \tag{9}$$

What about the effect of the price p_1 of good 1? To answer that, note from (6) that

$$\frac{\partial}{\partial p_1} L(x_1, x_2, \lambda, p_1, p_2, m) = \lambda x_1.$$

Thus Eq. (5) implies

$$\begin{aligned} \frac{\partial}{\partial p_1} V(p_1, p_2, m) &\overset{\text{a.e.}}{=} \tilde{\lambda}(p_1, p_2, m) \tilde{x}_1(p_1, p_2, m) \\ &= -\frac{\tilde{x}_1(p_1, p_2, m)}{p_1} \left. \frac{\partial}{\partial x_1} u(x_1, x_2) \right|_{(x_1, x_2) = \tilde{x}(p_1, p_2, m)}, \end{aligned} \tag{10}$$

with the second line due to (8).

To appreciate (9) and (10), imagine a consumer working for a steel factory in the states. The Trumpian trade war, raising the price of imported steel, has recovered the steel factory thereby improving the employment opportunity for this consumer, giving him an additional income Δm. Thus, according to (9), the consumer's welfare improves by about

$$\frac{\Delta m}{p_1} \left. \frac{\partial}{\partial x_1} u(x_1, x_2) \right|_{(x_1, x_2) = \tilde{x}(p_1, p_2, m)}.$$

On the other hand, the tariff hike against Chinese imports has pumped up the price p_1 for consumption goods, in the category labeled as good 1, including all sorts of previously cheap daily needed stuff such as clothing, toys, electronics, and smartphones aka attention-span-collapsing autonomy-foregoing drugs. Say the average price for such consumption goods jumps up by Δp_1. Then according to (10) the consumer's welfare worsens by about

$$\Delta p_1 \cdot \frac{\tilde{x}_1(p_1, p_2, m)}{p_1} \frac{\partial}{\partial x_1} u(x_1, x_2)\bigg|_{(x_1,x_2)=\tilde{x}(p_1,p_2,m)}.$$

In sum, this consumer's net gain from Trump's trade war is approximately equal to

$$(\Delta m - \Delta p_1 \cdot \tilde{x}_1(p_1, p_2, m)) \frac{1}{p_1} \frac{\partial}{\partial x_1} u(x_1, x_2)\bigg|_{(x_1,x_2)=\tilde{x}(p_1,p_2,m)}.$$

Since the second factor, $\frac{1}{p_1} \frac{\partial}{\partial x_1} u(x_1, x_2)\Big|_{(x_1,x_2)=\tilde{x}(p_1,p_2,m)}$, is positive (as it is the marginal utility of good 1 divided by the price), whether the consumer benefits or suffers from the trade war boils down to whether the first factor $\Delta m - \Delta p_1 \cdot \tilde{x}_1(p_1, p_2, m)$ is positive or negative. Even if the boost of his income, Δm, may be substantial because he works for a factory that the trade war is meant to protect, if the consumer is a frequenter of Walmart for the consumption goods, that is, if $\tilde{x}_1(p_1, p_2, m)$ is large, chances are that $\Delta m - \Delta p_1 \cdot \tilde{x}_1(p_1, p_2, m)$ could be negative and so he could still be hurt by the trade war. The beauty of this calculation is that it does not require any specific information about the consumer's utility function u, nor much about how he is re-budgeting among the goods given the price and income changes due to the trade war. Isn't that amazing?

4 Exercises

1. Consider the input deployment problem, defined by Eq. (10) in Chapter 3, with production function $f(x_1, x_2) := x_1^3 x_2^{1/2}$ for all nonnegative input bundles (x_1, x_2) and positive parameters w_1, w_2 and y, representing the prices for inputs 1 and 2, and the output quantity to be delivered. Denote $C(w_1, w_2, y)$ for the left-hand side of Eq. (10) in Chapter 3, i.e., the minimum cost of delivering y given input prices (w_1, w_2).

   a. Write down the choice set of this decision problem. Is it constant to the parameter w_1?

   b. Rewrite this decision problem in the format such that the min operator on the objective is replaced by the max operator.

   c. Calculate the partial derivative of the objective in the rewritten problem in Step (b) with respect to w_1.

   d. Suppose, given the parameters (w_1, w_2, y), that the firm's optimal input bundle is

   $$(\tilde{x}_1(w_1, w_2, y), \tilde{x}_2(w_1, w_2, y)).$$

   Combine this with Step (c) and Eq. (3) to obtain $\frac{\partial}{\partial w_1} C(w_1, w_2, y)$.

   e. Suppose that $w_1 = 48$, $w_2 = 2$ and $y = 256$. Then:

      i. Calculate the input quantity $\tilde{x}_1(w_1, w_2, y)$ when $w_1 = 48$, $w_2 = 2$ and $y = 256$.

ii. Plug the result in (e.i) into that in (d) to obtain $\frac{\partial}{\partial w_1}C(w_1, w_2, y)$ when $w_1 = 48$, $w_2 = 2$ and $y = 256$. By approximately how much does the firm's cost $C(w_1, w_2, y)$ increase if the input price w_1 increases by a small quantity Δw_1?

2. Consider the same decision problem in the above exercise problem.

   a. Is the choice set constant to the parameter y?

   b. Explain why the Lagrange method with equality constraints is applicable to this problem.

   c. Write down the Lagrangian of this problem based on the rewritten format obtained in Step (b) of the previous exercise problem.

   d. Calculate the partial derivative of the Lagrangian with respect to y.

   e. Suppose, given the parameters (w_1, w_2, y), that the solution for the (x_1, x_2, λ) in the first-order conditions according to the Lagrange method is

$$\left(\tilde{x}_1(w_1, w_2, y), \tilde{x}_2(w_1, w_2, y), \tilde{\lambda}(w_1, w_2, y)\right).$$

   Combine this with Step (d) and Eq. (5), with the role of $V(t)$ there played by $-C(w_1, w_2, y)$ here, to obtain $\frac{\partial}{\partial y}C(w_1, w_2, y)$.

   f. Suppose that $w_1 = 48$, $w_2 = 2$ and $y = 256$. Then:

      i. Solve for (x_1, x_2, λ) through the first-order conditions in the Lagrange method.

      ii. Plug the result in (f.i) into that in (e) to obtain $\frac{\partial}{\partial y}C(w_1, w_2, y)$—the marginal cost according to the terminology in Chapter 1—when $w_1 = 48$, $w_2 = 2$ and $y = 256$.

3. Different from the modern envelope theorem introduced here, the "envelope theorem" in most textbooks requires the assumptions (A) that both the maximand $\Phi(t)$ and the optimum $\tilde{x}(t)$ are differentiable functions of the parameter t, (B) that the objective $\varphi(x, t)$ is a differentiable function of both x and t, and (C) that the optimum $\tilde{x}(t)$ satisfies the first-order condition

$$\left.\frac{\partial}{\partial x}\varphi(x, t)\right|_{x=\tilde{x}(t)} = 0.$$

   a. Note that $\Phi(t) = \varphi(\tilde{x}(t), t)$.

   b. Given the above-listed assumptions, explain why

$$\frac{d}{dt}\Phi(t) = \left(\left.\frac{\partial}{\partial x}\varphi(x, t)\right|_{x=\tilde{x}(t)}\right)\frac{d}{dt}\tilde{x}(t) + \left.\frac{\partial}{\partial t}\varphi(x, t)\right|_{x=\tilde{x}(t)} = \left.\frac{\partial}{\partial t}\varphi(x, t)\right|_{x=\tilde{x}(t)}.$$

   c. Does Figure 1 satisfy Assumption (A)? Is Assumption (C) warranted in general?

References

[1] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, March 2002. 2