Comparative Statics

Motivation

When we build theories in political science, we are often particularly interested in how the behavior of a set of actors changes as some feature of the political environment changes. For instance, we may want to know how inequality affects marginal tax rates in a democracy or how alliance structures affect democratization. Before asserting "H1" and going to the data, it is necessary first to demonstrate that a hypothesized relationship between two variables follows logically from a set of explicit assumptions. This exercise is referred to as "comparative statics."

Parameterized Optimization

Let's suppose we want to build a theory of how much time a member of Congress spends fundraising. We will first make a set of assumptions about what the legislator want to achieve, for example, reelection. We then will make some additional assumptions about how time spent fundraising and maps onto these goals. In making these assumptions, we will presume that this mapping depends on a handful of parameters. We may think, for instance, that the legislator's popularity or their status in Congressional leadership will affect this mapping. These assumptions together will yield a parameterized objective function for the legislator,

$$f(x;\theta)$$

where x is time spent fundraising and θ is a parameter. To find out how much time a legislator spends fundraising given θ, we solve

$$\max_x f(x;\theta)$$

We are ultimately probably interested in how something in θ such as electoral danger affects time spent fundraising. This will allow us to derive a hypothesis that we then test empirically. Formally, we are interested in how

$$x^*(\theta) \equiv \arg\max_x f(x;\theta)$$

changes in θ. That is, we want to know the sign of $\frac{\partial}{\partial\theta}x^*(\theta)$. The set of maximizers, $x^*(\theta)$, is referred to as the **solution correspondence**. If we have a closed-form solution for $x^*(\theta)$, finding how $x^*(\theta)$ changes in θ is very straightforward. For example, if

$$f(x;\theta) = \theta\ln(x) - x^2$$

for $\theta \geq 0$, then by first order conditions,

$$\frac{\theta}{x} - 2x = 0$$

which yields

$$x^*(\theta) = \sqrt{\frac{\theta}{2}}$$

We see clearly then that $x^*(\theta)$ is increasing in θ. Very often, however, we will want to make our assumptions weak and general. This makes them more likely to hold in the real world and therefore make the results of the model more robust. Do we really believe that the legislator's utility function takes the very precise form as above? Probably not. It is more reasonable to think that the cost of fundraising in terms of time is increasing in x. We may also assume diminishing returns to fundraising. We may also assume that the marginal benefit of each hour spent fundraising in increasing in electoral danger. That is a more reasonable set of assumptions than a specific functional form. These assumptions gives us a general utility function

$$f(x;\theta) = v(x;\theta) - c(x)$$

where $v(x;\theta)$ is increasing in x for all θ and θ for all x and $c(x)$ is increasing in x. Now we cannot find a closed-form solution for $x^*(\theta)$ to help us. Fortunately, we have an extremely powerful tool to help us with this problem.

The Implicit Function Theorem

The **implicit function theorem** states that as long as a certain condition is satisfied, we can use differential calculus to characterize how $x^*(\theta)$ varies in response to small changes in θ. In low dimension, the implicit function theorem states:

Theorem 1 (Implicit Function Theorem (Low Dimension)) *Let $x^* \in \mathbb{R}$ solve $g(x,y) = 0$ at $y^* \in \mathbb{R}$. If $g(\cdot,\cdot)$ is continuously differentiable and $\frac{\partial}{\partial x} g(x^*, y^*) \neq 0$ then for some open set A containing x^* and an open set B containing y^*, there exists a continuously differentiable function $h : B \to A$ with $g(h(y), y) = 0$. The derivative of this function at y^* is given by*

$$\frac{\partial h(y^*)}{\partial y} = -(\frac{\partial g(x^*, y^*)}{\partial y})(\frac{\partial g(x^*, y^*)}{\partial x})^{-1}$$

To see how this applies to comparative statics, note that the first-order conditions of a parameterized optimization problem give us the condition $g(x, y) = 0$. The problem must, of course, have a solution to begin with. In the next lecture we will explore how verify that a general program has a solution. In this lecture we will assume that all problems have solutions, i.e. that $x^*(\theta)$ is always nonempty.

Returning to the example, we know that $\frac{\partial f(x^*;\theta)}{\partial x} = 0$. By the implicit function theorem

$$\frac{\partial x^*(\theta)}{\partial \theta} = -\frac{\partial^2 f(x;\theta)}{\partial x \partial \theta}(\frac{\partial^2 f(x;\theta)}{\partial x^2})^{-1} = -\frac{\partial^2 v(x;\theta)}{\partial x \partial \theta}(\frac{\partial^2 v(x;\theta)}{\partial x^2} - \frac{\partial^2 c(x)}{\partial x^2})^{-1}$$

We assumed that $v(x;\theta)$ is increasing and concave in x so $\frac{\partial^2 v(x;\theta)}{\partial x^2} < 0$. If we assume that marginal cost is non-decreasing in x, then $\frac{\partial^2 c(x)}{\partial x^2} \geq 0$. The denominator is therefore negative which cancels out the negative scalar. The sign of $\frac{\partial^2 v(x;\theta)}{\partial x \partial \theta}$ will therefore tell us how $x^*(\theta)$ changes with θ.

The implicit function theorem generalizes to higher dimensions with n choice variables and m parameters.

Theorem 2 (Implicit Function Theorem) *Let $x^* \in \mathbb{R}^n$ solve $f(x,y) = 0$ at $y \in \mathbb{R}^m$. If $f_1(\cdot)$ through $f_n(\cdot)$ are continuously differentiable in each coordinate of x and y and the Jacobian matrix with respect to the endogenous variables is nonsingular, then for some open set A containing x^* and an open set B containing y^*, there exists a continuously differentiable function $\phi : B \to A$ with $f(\phi(y), y) = 0$. The derivative of this function at y^* is given by the $n \times k$ matrix*

$$D_y\phi(y^*) = -[D_x f(x^*, y^*)]^{-1} D_y f(x^*, y^*)$$

Returning to our application, let $x = (x_1, x_2, ...x_n)$ and $\theta = (\theta_1, \theta_2, ..., \theta_m)$. Assume that a single-valued solution to

$$\max_x f(x;\theta)$$

exists for all θ, $x^*(\theta)$. Now say we want to know how $x^*(\theta)$ changes with θ. That is, we are interested in finding the $m \times n$ Jacobian matrix of $x^*(\theta)$, $D_\theta x^*(\theta)$.

The implicit function theorem tells us that

$$D_\theta x^*(\theta) = -[H_x f(x,\theta)]^{-1} D_\theta f(x,\theta)$$

where $H_x f(x^*,\theta)$ is the $n \times n$ Hessian matrix of $f(x^*,\theta)$ with respect to the choice variables and $D_\theta f(x^*, \theta)$ is the $n \times m$ Jacobian of $f(x^*, \theta)$ with respect to the parameters. Note that the implicit function theorem requires that H be nonsingular.

Example: Let

$$f(x, y; \theta) = v(x) + \theta w(y) - c(x, y)$$

where $v(x)$ and $w(y)$ are increasing an concave and $c(x, y)$ is increasing and convex in both arguments. By construction a solution will exist that is characterized by first-order conditions. We want to find $\begin{bmatrix} \frac{\partial x^*(\theta)}{\partial \theta} \\ \frac{\partial y^*(\theta)}{\partial \theta} \end{bmatrix}$. By the implicit function theorem,

$$\begin{bmatrix} \frac{\partial x^*(\theta)}{\partial \theta} \\ \frac{\partial y^*(\theta)}{\partial \theta} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 f(x,y;\theta)}{\partial x^2} & \frac{\partial^2 f(x,y;\theta)}{\partial x \partial y} \\ \frac{\partial^2 f(x,y;\theta)}{\partial x \partial y} & \frac{\partial^2 f(x,y;\theta)}{\partial y^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial^2 f(x,y;\theta)}{\partial x \partial \theta} \\ \frac{\partial^2 f(x,y;\theta)}{\partial y \partial \theta} \end{bmatrix}$$

$$= - \begin{bmatrix} \frac{\partial^2 v(x)}{\partial x^2} - \frac{\partial^2 c(x,y)}{\partial x^2} & -\frac{\partial^2 c(x,y)}{\partial x \partial y} \\ -\frac{\partial^2 c(x,y)}{\partial x \partial y} & \frac{\partial^2 \theta w(y)}{\partial y^2} - \frac{\partial^2 c(x,y)}{\partial y^2} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \frac{\partial w(y)}{\partial y} \end{bmatrix}$$

$$= -\frac{1}{\det(H)} \begin{bmatrix} \frac{\partial^2 \theta w(y)}{\partial y^2} - \frac{\partial^2 c(x,y)}{\partial y^2} & \frac{\partial^2 c(x,y)}{\partial x \partial y} \\ \frac{\partial^2 c(x,y)}{\partial x \partial y} & \frac{\partial^2 v(x)}{\partial x^2} - \frac{\partial^2 c(x,y)}{\partial x^2} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{\partial w(y)}{\partial y} \end{bmatrix}$$

$$= -\frac{1}{\det(H)} \begin{bmatrix} \frac{\partial^2 c(x,y)}{\partial x \partial y} \frac{\partial w(y)}{\partial y} \\ (\frac{\partial^2 v(x)}{\partial x^2} - \frac{\partial^2 c(x,y)}{\partial x^2}) \frac{\partial w(y)}{\partial y} \end{bmatrix}$$

By second order conditions, $\det(H) > 0$. By the concavity of $v(\cdot)$ and convexity of $c(\cdot, \cdot)$, $(\frac{\partial^2 v(x)}{\partial x^2} - \frac{\partial^2 c(x,y)}{\partial x^2}) < 0$. By assumption $\frac{\partial w(y)}{\partial y} > 0$. Therefore $y^*(\theta) > 0$. We have not yet made any assumptions about $\frac{\partial^2 c(x,y)}{\partial x \partial y}$. Since $-\frac{\partial w(y)}{\partial y} < 0$, if $\frac{\partial^2 c(x,y)}{\partial x \partial y} \geq 0$, then $x^*(\theta) \leq 0$. Otherwise, $x^*(\theta) > 0$. ■

A few caveats are in order before moving on. In all of our examples, $x^*(\theta)$ is nonempty and singleton. In general, $x^*(\theta)$ may be empty or multi-valued. It is your responsibility to first show that a solution exists before applying the implicit function theorem. When $x^*(\theta)$

is multi-valued, the implicit function works locally for every element of $x^*(\theta)$. However, the sign of the derivative may be different at different optima. Additionally, FOCs require the differentiability of f. In practice, we may deal with functions that are discontinuous or not differentiable. For these cases, we will need a different tool for comparative statics. In practice you will also need to check that H is non-singular. Finally, we have ignored constraints in this treatment of the implicit function theorem. The implicit function theorem is sufficiently general to be used in equality and inequality constrained problems. Simply apply the theorem to the first-order conditions of a Lagrangian or to the problem's KKT conditions.

The Envelope Theorem

So far we have been analyzing the behavior of the maximizers of a parameterized optimization problem. We often will also want to analyze how the value of a parameterized optimization problem changes with θ. The function

$$V(\theta) \equiv f(x^*(\theta); \theta)$$

is referred to as the **value function** or **indirect utility function** for the problem

$$\max_x \; f(x; \theta)$$

The sign of $\frac{\partial V(\theta)}{\partial \theta}$ tells us how an actor's utility changes for a small rise in θ given that she is maximizing. This is ostensibly a non-trivial exercise: θ has two effects on $V(\theta)$, a direct effect and an indirect effect that operates through $x^*(\theta)$.

Theorem 3 (The Envelope Theorem) *For some θ, suppose the unique maximizer $x^*(\theta)$*

is locally characterized by first order conditions

$$\frac{\partial f(x^*(\theta),\theta)}{\partial x} = 0$$

Suppose that the conditions of the implicit function theorem hold locally so that $x^*(\cdot)$ *is a differentiable function of* θ. *Then*

$$\frac{\partial V(\theta)}{\partial \theta} = \frac{\partial f(x^*(\theta),\theta)}{\partial \theta}$$

The proof of the envelope theorem is remarkably simple. By first order conditions, $\frac{\partial f(x;\theta)}{\partial x} = 0$ at the optimum. Then by the chain rule,

$$\frac{\partial V(\theta)}{\partial \theta} = \frac{\partial f(x^*(\theta);\theta)}{\partial x}\frac{\partial x^*(\theta)}{\partial \theta} + \frac{\partial f(x^*(\theta);\theta)}{\partial \theta}\frac{\partial \theta}{\partial \theta} = \frac{\partial f(x^*(\theta);\theta)}{\partial \theta}$$

This is a powerful result. As θ changes, so does $x^*(\theta)$. Therefore we need to take into account both the direct effect of θ on utility but also the change that comes indirectly from the change in $x^*(\theta)$. What the envelope theorem says is that the indirect effect does not matter! Only the direct effect of the exogenous parameter needs to be considered even though the exogenous parameter may enter the value function indirectly.

Example: Let's revisit the earlier fundraising example where $f(x;\theta) = \theta \ln(x) - x^2$ and $x^*(\theta) = \sqrt{\frac{\theta}{2}}$. This gives us $V(\theta) = \theta \ln(\sqrt{\frac{\theta}{2}}) - \frac{\theta}{2}$.

$$\frac{\partial V(\theta)}{\partial \theta} = \ln(\sqrt{\frac{\theta}{2}}) - \frac{1}{2} + \theta\frac{\partial}{\partial \theta}\ln(\sqrt{\frac{\theta}{2}})$$

$$\theta\frac{\partial}{\partial \theta}\ln(\sqrt{\frac{\theta}{2}}) = \theta\sqrt{\frac{2}{\theta}}\frac{\partial}{\partial \theta}\sqrt{\frac{\theta}{2}} = \theta\sqrt{\frac{2}{\theta}}\frac{1}{\sqrt{2}}\frac{\partial}{\partial \theta}\sqrt{\theta} = \frac{1}{2}\frac{\sqrt{\theta}}{\sqrt{\theta}} = \frac{1}{2}$$

Therefore

$$\frac{\partial V(\theta)}{\partial \theta} = \ln(\sqrt{\frac{\theta}{2}}) = \frac{\partial f(x,\theta)}{\partial \theta}\Big|_{x^*(\theta)}$$

For $\theta > 2$, this term is positive. ■

Example: For $f(x, y; \theta) = v(x) + \theta w(y) - c(x, y)$, $V(\theta) = v(x^*(\theta)) + \theta w(y^*(\theta)) - c(x^*(\theta), y^*(\theta))$. By the envelope theorem, $\frac{\partial V(\theta)}{\partial \theta} = w(y^*(\theta)) \geq 0$. ■

As with the implicit function theorem, this lecture ignores constraints. If a problem has constraints, the envelope theorem can still be applied using the Lagrangean.

Correspondences

We have so far been assuming that $x^*(\theta)$ is a singleton for all θ. We have similarly been assuming that $x \in \mathbb{R}$ and ignoring the possibility of constraints. We typically assume that $x \in G(\theta)$ where G is called a **feasible set**. If $G(\theta)$ depends on θ, then unless $G(\theta)$ is a singleton, then we need a more general concept than function to think about the properties of $G(\theta)$. (Why?)

Given two sets X and Y, **correspondence**, $G : X \rightrightarrows Y$ is a rule that assigns to every element $x \in X$ a subset $G(x)$ of Y. Correspondences are therefore sometimes referred to as "set valued functions."

Example: Let $X = \mathbb{R}_+$ and $Y = \mathbb{R}_+$. $G(x) = [0, x]$ is a correspondence. ■

Example: Let $X = \{1, 2, 3, 4, 5, 6\}$ and let $G(x) = \{z \in X : z \geq x\}$. This correspondence returns all the elements of x that are greater than or equal to the input e.g. $G(4) = \{4, 5, 6\}$. ■

We refer to $x^*(\theta)$ as a solution correspondence because there may be more than one solution to a problem given θ.

A correspondence if **compact valued** if $G(x)$ is a compact set. It is called **convex valued** if $G(x)$ is convex. It is called **closed valued** if $G(x)$ is closed. A **fixed point** of a correspondence is a point $x \in X$ such that $x \in G(x)$.

A correspondence $G : X \rightrightarrows Y$ is **upper hemicontinuous at** $x \in X$ if for every open subset V of Y with $G(x) \subseteq V$, there exists a $\delta > 0$ such that $G(N_\delta(x)) \subseteq V$. G is called

upper hemicontinuous if it is upper hemicontinuous on all of X.

Upper hemicontinuity is easier to visualize if G is closed valued and X is closed. The **graph** of $G : X \to Y$, is defined as

$$\Gamma(G) = \{(x, y) \in X \times Y : y \in G(x)\}$$

If G is upper hemicontinuous and closed valued and X is closed, then $\Gamma(G)$ is closed. If Y is compact, the converse is also true.

Uhc can also be characterized with sequences. If $x^*(\theta)$ is a solution correspondence and $\theta_n \to \theta_0$ is a sequence of parameters in the parameter space Θ, then θ_n induces a sequence of subsets of the feasible choice set, $x^*(\theta_n)$. Now consider a sequence of optimal choices that this sequence induces which converges to a choice x. The solution correspondence $x^(\theta)$ is uhc at $x^*(x_0)$ if for all such convergent sequences of choices and convergent sequences of parameters, $\lim x^*(x_n) = x \in x^*(x_0)$.

Maximum Theorem

As we have seen, it is rare that we will be able to find an explicit solution correspondence. Often, however, we can appeal to properties of the objective function and constraint set to infer properties of the constraint correspondence. We saw above that we can exploit the differentiability of an objective function to learn about the properties of locally smooth solution correspondences and value functions. Our problems, however, will not always be so well behaved. A common utility function in spatial models is the **absolute loss function**, $u(x; z) = -|x - z|$. We cannot use the implicit function for comparative statics here. Intuitively, however, we should be able to exploit the continuity of the objective function to learn about the solution correspondence. The **maximum theorem** does just that. A "light" version of it is presented here where it is assumed that the feasible set, G, is independent of θ.

Theorem 4 (Maximum) *If $f : X \times \Theta \to \mathbb{R}$ is continuous and the feasible set G is compact, then $V(\theta)$ is continuous and $x^*(\theta)$ is nonempty, compact-valued, and upper hemicontinuous.*

The result of the theorem is that if the elements of an optimization problem are sufficiently continuous, then some, but not all, of that continuity is preserved in the solutions. The Maximum theorem is also referred to as "The Theorem of the Maximum," "The Continuous Maximum Theorem," and "Berge's Theorem of the Maximum." We will revisit this and see the full version in the next lecture when we study the existence of solutions.

Monotone Comparative Statics

Occasionally, even continuity of an objective function may be too strong of an assumption. A powerful and underutilized tool in political research, **monotone comparative statics**, allows us to study the properties of $x^*(\theta)$ even when f is discontinuous.

We say that $f(x, \theta)$ satisfies the **single-crossing property** if for all $x > x'$ and $\theta > \theta'$,

$$f(x, \theta') - f(x', \theta') \geq 0 \implies f(x, \theta) - f(x', \theta) \geq 0$$

and

$$f(x, \theta') - f(x', \theta') > 0 \implies f(x, \theta) - f(x', \theta) > 0$$

Intuitively if the two objective functions for different θ cross only once, the function satisfies the single crossing property.

A sufficient condition for single crossing is **increasing differences**. A function displays increasing differences if

$$f(x, \cdot) - f(x', \cdot)$$

is weakly increasing in θ: if $f(x, \theta') - f(x', \theta') \geq 0$, then

$$f(x, \theta) - f(x', \theta) \geq f(x, \theta') - f(x', \theta') \geq 0$$

and f is single crossing.

The central result of monotone comparative statics comes from Milgrom and Shannon (1994).

Theorem 5 (Milgrom and Shannon) *If the single crossing condition holds, then $x^*(\theta)$ is weakly increasing in θ.*

This result generalizes to higher dimensions.

If $X \subseteq \mathbb{R}^n$ and $\Theta \subseteq \mathbb{R}^m$, the set $X \times \Theta$ is a **product set** if it can be represented by the Cartesian product of subsets of $\mathbb{R}$. For example, $[a, b] \times [a, b]$ but not $\{a, b : x \geq 0, y \geq 0, x + y \leq 1\}$. A product set is a special type of chain.

A function $f : X \times \Theta \to \mathbb{R}$ is **supermodular** if $X \times \Theta$ is a product set and f has increasing differences for all pairs of arguments of the function. Supermodularity generalizes the notion of "complementarity." If f is differentiable, then f is supermodular if all off-diagonal elements of H_f are positive i.e. if the cross partial derivative of all pairs of variables are positive.

Theorem 6 (Topkis) *Consider the problem*

$$\max_{x \in X} f(x, \theta)$$

where $\theta \in \Theta$ is a parameter. If f is supermodular, then every component of $x^*(\theta)$ is weakly increasing in θ

A rise in θ leads the decision maker to directly raise each x_i. Because the x_i are all complementary to one another, all indirect effects are also positive. Because the x_i are all complementary to one another, all indirect effects are also positive. Hence the "monotone" in monotone comparative statics. The power of this approach is its generality. Functions do not even need to be continuous for us to do comparative static analysis. All we need is supermodularity.

Exercises

1) Prove that if f is differentiable, then f has increasing differences if and only if $\frac{\partial^2 f(x,\theta)}{\partial x \partial \theta} \geq 0$.

2) Consider the parameterized optimization problem

$$\max_{x\in[0,1]} \quad (1-x)p(x) + q(1-p(x))$$

where $p(\cdot) > 0$ is strictly increasing and concave. Assume that x^* is on the interior of $[0,1]$.

i) Use the implicit function theorem to show how the optimal choice of x given q, $x^*(q)$ changes as q changes.

ii) How does the value function change as q changes?

3) Consider the parameterized optimization problem

$$\max_{x,z} \quad f(x,z;\theta)$$

where $x, y, \theta \in \mathbb{R}$. Assume f is twice continuously differentiable. Let f_{ij} denote $\frac{\partial^2 f}{\partial i \partial j}$ for $i, j \in \{x, z, \theta\}$ i.e. f_{xx} is the second derivative of f and f_{xz} is the cross partial derivative of f with respect to x and z. Let $(x^*(\theta), z^*(\theta))$ be a solution.

i) What conditions on f_{xx}, f_{zz}, and f_{xz} must hold for $(x^*(\theta), z^*(\theta))$ to be a local maximum? (Hint: what must be true of the Hessian matrix with respect to choice variables at a local maximum?)

ii) Use the implicit function theorem to characterize $\frac{\partial}{\partial\theta}x^*(\theta)$ and $\frac{\partial}{\partial\theta}z^*(\theta)$ in terms of f_{ij}.

iii) Let $f_{x\theta} = 0$, $fz\theta < 0$, $f_{xz} > 0$. Describe the comparative statics. Now let $f_{xz} < 0$ and describe the comparative statics. Interpret this result.

iv) Show that if f is supermodular, then $\frac{\partial}{\partial\theta}x^*(\theta) > 0$ and $\frac{\partial}{\partial\theta}z^*(\theta) > 0$.

Source Material

Ashworth, S. & Bueno de Mesquita. (2006). "Monotone Comparative Statics for Models of Politics." *American Journal of Political Science 50*(1), 214-231.

Carter, M. (2001). *Foundations of Mathematical Economics.* Cambridge, MA: The MIT Press.

McCarty, N., & Meirowitz, A. (2007) *Political Game Theory: An Introduction.* New York, NY: Cambridge University Press.

Ok, E.A. (2007). *Real Analysis with Economic Applications.* Princeton, NJ: Princeton University Press.