Comparing Optima: Do Simplifying Assumptions Affect Conclusions?

Paul Milgrom

Stanford University

Consider a family of maximization models in which the optimum trades off beneficial and costly effects. Then comparative statics derived under many kinds of simplifying assumptions about the benefits technology are also true for general (convex and nonconvex) technologies. For example, any comparative statics conclusion about investment by a risk-averse decision maker under uncertainty that holds when expected returns are described by a general linear function also holds for an arbitrary nonlinear expected return function.

I. Introduction

One of the continuing controversies that divide economic theorists is the question of how to regard results obtained from models incorporating particular functional forms. Nearly every subfield of economics has produced influential papers with models that incorporate such forms, especially for technology or demand. These forms are typically adopted for reasons of tractability or computability, not because they are empirically supported or are even a priori plausible. Critics observe that the conclusions derived from such models are sometimes misleading and worry that they might often be so. This paper addresses one aspect of that concern: When are the comparative statics of narrowly parameterized families of optimization models reliable guides to the general case?

My analysis focuses on the most commonly studied class of optimi-

I thank Joshua Gans and Armin Schmutzler for research assistance; Aaron Edlin, Avner Greif, Roy Navon, Randy Verbrugge, and a referee for comments; and the National Science Foundation for financial support.

[*Journal of Political Economy*, 1994, vol. 102, no. 3]

zation models for economic applications. This is the class in which the optimum is determined by trading off two or more effects of a particular choice variable. Formally, such problems can be expressed in terms of maximizing a function $V(x, f(x), \theta)$ subject to a constraint $x \in K \subset \mathbb{R}$, where x is the choice variable, $\theta \in \mathbb{R}$ is a parameter, and $f(x)$ is the benefits production function. I assume that V is continuously differentiable and that $V_2 > 0$. For many applications, I also assume that $V_1 \leq 0$. Then increasing x directly reduces utility through the first argument of V and may produce benefits (through f) that are valued through the second argument of V.

The main theorem of this paper establishes that for problems in this class, assumptions about that exact form of f are irrelevant for the comparative statics analysis. More precisely, if a comparative statics result is established for the linear functions $f(x) = ax + b$ or for the functions in any other family in which the level and slope can be varied independently, then the same conclusion is true for *any* function f. The theorem also establishes that the separate convexity assumptions governing the payoff function V and the constraint set K are always inessential for comparative statics analysis.

These conclusions do not mean that functional form assumptions are either useless or inconsequential for economic analysis. Functional form assumptions may be helpful for deriving explicit formulas for empirical estimation or simulations or simply to lend insight into the problem structure, and they certainly can help determine the magnitude of comparative statics effects. But with economic knowledge at its current state, functional form assumptions are never really convincing, and this lends importance to the question I ask and to its answer: One can indeed often draw valid general comparative statics inferences from special cases. I begin by illustrating that point with a particularly striking example.

II. An Example

Consider the problem of the risk-averse firm under price uncertainty first studied by Sandmo (1971). In the original version of the model, the firm is competitive and is run by an entrepreneur with exogenous income I, facing a price $P + \epsilon$ for its output, where ϵ is a mean zero random variable. The firm produces output x and earns an expected profit of $\pi(x) = Px - C(x)$. The entrepreneur chooses $x \geq 0$ to maximize $E[U_\theta(I + Px - C(x) + S_\theta(x) + x\epsilon_\theta)]$. Then θ may parameterize shifts in the cost or demand functions through S_θ, the risk attitude or income tax rate through U_θ, or the randomness of demand through ϵ_θ, or any combination of these. Assume that, for all θ, $U_\theta(\cdot)$ is increasing, smooth, and concave and that $S_\theta(\cdot)$ is nonincreasing and concave. Although Sandmo's original model focused on the competitive firm, I generalize it here to include models in which the firm may have some market power. Thus the nonrandom component of price is treated here as a function of output: $P = g(x)$.

The firm's problem can be represented in the general benefit-cost form with costly activity x and benefit y by taking $V(x, y, \theta) = E[U_\theta(y + S_\theta(x) + x\epsilon_\theta)]$, where $\theta \in T$ and T is an arbitrary subset of $\mathbb{R}$. Note that V is increasing in y and nonincreasing in x (because S is nonincreasing and U is increasing and concave). This formulation entails conceptualizing the problem as one in which increases in x lead to increases in risk that may be offset by increases in expected profits. The optimum is determined by balancing these effects, with the balance affected by θ. The following proposition is a corollary of my general theorem.

PROPOSITION 1. In the generalized Sandmo problem, the following three conditions (each of which is a joint restriction on U_θ, S_θ, and ϵ_θ) are equivalent.

1. In the special case of the model in which the firm is a price taker ($g(x) \equiv P$) and costs are zero ($C(x) \equiv 0$), $x^*(\theta|P, I)$ is monotone nondecreasing in θ for all prices $P > 0$ and all exogenous incomes I.
2. In the special case of the model in which demand is linear with slope -1 ($g(x) = A - x$) and costs are $C(x) = \frac{1}{2}x^2$, $x^*(\theta|A, I)$ is monotone nondecreasing in θ for all values of the parameter $A > 0$ and all exogenous incomes I.
3. For all expected inverse demand functions $g(\cdot)$, cost functions $C(\cdot)$, and exogenous incomes I, $x^*(\theta|g, C, I)$ is monotone nondecreasing in θ.

Thus, for each of the kinds of comparative statics exercises listed above, showing monotone comparative statics in a very special case such as case 1 or 2 is equivalent to establishing the same conclusion for the case of general demand and cost functions.

I have not yet defined monotone comparative statics for the case in which the optimum is not unique. A sensible definition must assert that for any two parameter values $\theta \geq \theta'$ for which the optimum $x^*(\cdot)$ is unique, $x^*(\theta) \geq x^*(\theta')$. The proposition does indeed imply that and more. The treatment of nonexistent and nonunique optima is explained in the next section.

The key to understanding the proposition is that the particular cases, though restrictive, are rich enough to incorporate what I call *semifull two-parameter families* of benefit production functions. The families are $f(x|I, P) = I + Px$ for case 1 and $f(x|I, A) = I + x(A - x) - \frac{1}{2}x^2$ for case 2. Both families have the crucial property that, for each x, one can select the parameters to fix both the value $f(x)$ and the slope $f'(x)$ at any desired levels.

Results such as proposition 1 can be used to expand the scope of earlier analyses. For example, one variation of the problem studied in case 1 of the proposition is identical to Arrow's (1971) problem of the demand for a risky asset, where P is reinterpreted to be the excess return on the risky asset. Arrow showed that, in that case, a sufficient condition for $x^*(I|P)$ to be monotone nondecreasing in I for all P is that U exhibit decreasing absolute risk aversion. If U and the distribution of ϵ are taken to be independent of θ and $S_\theta(x)$ is set equal to θ, proposition 1 then implies that the same condition is also sufficient to conclude that $x^*(I|g, C)$ is monotone nondecreasing in I for all demand functions g and all cost functions C in the generalized Sandmo model. Similarly, Rothschild and Stiglitz (1971) showed that if there is decreasing absolute risk aversion, then increasing the riskiness of investment in Arrow's model leads to reduced demand for the asset. It follows that under the same conditions in the Sandmo model, with general demand and cost functions, increasing riskiness of ϵ_θ causes the entrepreneur to produce less output.

III. The Main Theorem

In general, maximization problems may have multiple optima or no optimum at all, so I shall need to be careful about the meaning of the phrase "monotone comparative statics." Let $x^*(\theta|f, S)$ denote the possibly empty set of maximizers in a one-variable maximization problem with objective $V(x, f(x), \theta)$ subject to $x \in S$. I define monotonicity using Veinott's "strong set order" on the range of the function x^*. Formally, x^* is monotone nondecreasing in θ if, for every two parameter values $\theta' < \theta''$ for which an optimum exists, the following condition holds: If $x' \in x^*(\theta'|f, S)$ and $x'' \in x^*(\theta''|f, S)$, then $\max(x', x'') \in x^*(\theta''|f, S)$ and $\min(x', x'') \in x^*(\theta'|f, S)$. If one restricts attention to values of the parameter θ for which the optimum is unique, this condition is equivalent to the intuitive definition that the unique optimum is a monotone nondecreasing function of θ. Notice from the definition that if the set of optima is empty for some parameter value θ', then it is automatically true that $x^*(\theta) \geq x^*(\theta')$ and $x^*(\theta') \geq x^*(\theta)$. Hence, the claims about monotone comparative statics made below entail no conclusions about the existence of an optimum.

In the theorem below, we shall also have need of the concept of a *full two-parameter family* $\{g(\cdot|\alpha)\}$ of concave functions from $\mathbb{R}$ to $\mathbb{R}$. This is a family of concave functions with parameter α lying in some set A and with the property that, for all $x, a, b \in \mathbb{R}$, there is some parameter value $\alpha \in A$ such that $g(x|\alpha) = b$ and $\partial g(x|\alpha)/\partial x = a$. A *semifull two-parameter family* of concave functions is defined by the same restrictions, but the restrictions are imposed only when $a > 0$.

Finally, the (x, y)–level sets of V are the sets $S_{\theta v} = \{(x, y) | V(x, y, \theta) = v\}$. In the theorem below, I assume $V_2 > 0$ so that there is at most a single value $y(x|\theta, v)$ such that $(x, y(x)) \in S_{\theta v}$. The set $S_{\theta v}$ is a *curve* if $y(\cdot|\theta, v)$ is a continuous function. With the definitions in place, I can state the main theorem.

THEOREM 1. Let V: $\mathbb{R}^3 \to \mathbb{R}$ be continuously differentiable with $V_2 > 0$ and suppose that the (x, y)–level sets of V are curves. Let $S \subset \mathbb{R}$ be convex. Then the following two conditions are equivalent:

1. For all functions f: $\mathbb{R} \to \mathbb{R}$ and all sets $K \subset S$, the solution $x^*(\theta|f, K)$ of the following problem is monotone nondecreasing in θ:

$$\text{maximize } V(x, f(x), \theta) \quad \text{subject to } x \in K.$$

2. For all $x \in \text{int}(S)$ and all $y \in \mathbb{R}$, the function $V_1(x, y, \theta)/V_2(x, y, \theta)$ is monotone nondecreasing in θ.

Suppose, in addition, that V is quasi-concave. If either (i) $\{g(\cdot|\alpha)\}$ is a full two-parameter family of concave functions or (ii) $V_1 \leq 0$ everywhere and $\{g(\cdot|\alpha)\}$ is a semifull two-parameter family of concave functions, then conditions 1 and 2 are also equivalent to the following condition:

3. For all $\alpha \in A$, the solution $x^*(\theta|\alpha)$ of the following problem is monotone nondecreasing in θ:

$$\text{maximize } V(x, g(x|\alpha), \theta) \quad \text{subject to } x \in S. \tag{1}$$

Proof. The equivalence of conditions 1 and 2 is established in Milgrom and Shannon (1994), and it is obvious that condition 1 implies condition 3. So it is sufficient to verify that condition 3 implies condition 2 under the hypotheses of the theorem.

Choose any $\bar{x} \in \text{int}(S)$, $\bar{y} \in \mathbb{R}$, and $\theta \in \mathbb{R}$. Choose $\alpha \in A$ such that $g(\bar{x}|\alpha) = y$ and $\partial g(\bar{x}|\alpha)/\partial x = -V_1(\bar{x}, \bar{y}, \theta)/V_2(\bar{x}, \bar{y}, \theta)$. This choice of α means that the convex preferred set $\{(x, y) | V(x, y, \theta) \geq V(\bar{x}, \bar{y}, \theta)\}$ is tangent to the convex set $\{(x, y) | y \leq g(x|\alpha)\}$ at $(\bar{x}, \bar{y})$. Hence, $(\bar{x}, \bar{y})$ maximizes $V(x, y, \theta)$ subject to $x \in S$ and $y \leq g(x|\alpha)$. Therefore, since $V_2 > 0$, $\bar{x}$ maximizes $V(x, g(x|\alpha), \theta)$ subject to $x \in S$, that is, $\bar{x} \in x^*(\theta|\alpha)$. If, for some $\theta' > \theta$, condition 2 fails at $(\bar{x}, \bar{y})$, then the derivative with respect to x of $V(x, g(x|\alpha), \theta')$ is less than that of $V(x, g(x|\alpha), \theta)$ at $\bar{x}$ and, hence, negative. Therefore, by quasi concavity of the objective, any $x' \in x^*(\theta'|\alpha)$ satisfies $x' < \bar{x}$, in contradiction to condition 3. Q.E.D.

Several remarks may help clarify the scope and meaning of the theorem. First, the proof of the theorem works by deriving necessary conditions for the convex case from standard convex programming methods and then using a direct argument to establish that the same conditions are sufficient for the nonconvex case. This is a familiar idea from the analysis of demand relations, carried here to its logical conclusion, and allows us to unify convex and nonconvex comparative statics into one single theory. Although the theory is derived for comparative statics of global maxima in optimization problems, it can also be used to analyze local maxima, because the local maxima are just the global maxima for a family of problems with a small constraint set.

Second, although this example emphasizes that one can use convex methods on particular parameterized problems to obtain more general comparative statics, the theorem suggests an alternative approach: direct verification of condition 2. This approach, which Milgrom and Shannon (1994) have dubbed the "method of dissection," has several advantages for applications. It can be applied regardless of whether the problem is convex; it works whenever anything does (because condition 2 is *necessary* as well as sufficient, even within the class of convex problems); and it serves intuition by directing attention away from the irrelevant structure of the benefit production function and by emphasizing a marginal rate of substitution concept that is familiar to economists from demand theory.

The final remark concerns the possibility of strengthening conclusion 1 of the theorem in the case of multiple optima. As it stands, condition 1 still allows that when $x^*(\theta)$ is not single-valued, there may exist a selection $x(\theta)$ from the set of optimizers that is not monotone nondecreasing. A variation of the theorem that rules that out follows. In this variation, $x(\theta)$ is said to be a selection from $x^*(\theta)$ if the domain of $x(\cdot)$ is the set of parameter values such that $x^*(\theta)$ is nonempty and for all θ in the domain $x(\theta) \in x^*(\theta)$.

THEOREM 2. Let V: $\mathbb{R}^3 \rightarrow \mathbb{R}$ be continuously differentiable with $V_2 > 0$ and suppose that the (x, y)–level sets of V are curves. Let $S \subset \mathbb{R}$ be convex. Then condition 2 below implies condition 1:

1. For all functions f: $\mathbb{R} \rightarrow \mathbb{R}$ and all sets $K \subset S$, *every* selection $x(\theta)$ from $x^*(\theta | f, K)$ is monotone nondecreasing.
2. For all $x \in \text{int}(S)$ and all $y \in \mathbb{R}$, the function $V_1(x, y, \theta)/V_2(x, y, \theta)$ is monotone increasing in θ.

Suppose, in addition, that V is quasi-concave. If either (i) $\{g(\cdot | \alpha)\}$ is a full two-parameter family of concave functions or (ii) $V_1 \leq 0$ everywhere and $\{g(\cdot | \alpha)\}$ is a semifull two-parameter family of concave functions, then condition 3 below implies condition 2:

3. For all $\alpha \in A$, the solution $x^*(\theta|\alpha)$ of problem (1) is unique, is continuously differentiable in θ, and satisfies $\partial x^*/\partial\theta > 0$.

Thus, in the Sandmo example, if $x^*(\theta|\alpha)$ is single-valued and increasing in either case 1 or case 2, then every selection from $x^*(\theta|\alpha)$ is nondecreasing in the general case 3.

Many of the most general results of demand theory have been derived on the basis of simplifying assumptions, especially the assumption that the budget set is linear. That assumption is sometimes unnatural and, as our theory implies, often dispensable with no extra work. For example, the effect on labor supply of changing the tax rate on income is typically studied by analyzing the problem of maximizing $U((1 - \tau)(I + wx), \bar{x} - x)$ subject to $x \in [0, \bar{x}]$. Here, U is a quasi-concave utility function defined on income-leisure pairs, τ is the tax rate, I is nonlabor income, w is the wage, and x is the labor hours supplied out of the total labor endowment $\bar{x}$. Suppose that for some family of utility functions U it is concluded that, for all I and w, $x^*(\tau)$ is monotone nondecreasing (or nonincreasing) in τ. Then the results imply that the same conclusion holds when the linear wage function wx is replaced to be an arbitrary nonlinear function $f(x)$ and when the constraint $x \in [0, \bar{x}]$ is supplemented by an arbitrary constraint $x \in K$. Nonlinear functions can represent pay differentials for part-time and overtime work, which are prominent features of the real working environment, and the constraint set K can reflect realistic restrictions on the hours of work that are available in various jobs. To verify this claimed generalization of the comparative statics conclusion using theorem 1 or 2, one takes $V(x, y, \theta) = U(\theta y, \bar{x} - x)$ and $g(x|I, w) = I + wx$.

IV. Higher Dimensions

The theorems in the previous section are stated for problems in which the choice variable is a real number, but it is obvious that the same theorems imply that convexity assumptions can also be relaxed in problems in which the choice set is a subset of $\mathbb{R}^n$ for $n > 1$. The reason is that one can always perform a two-stage maximization to reduce a multivariate problem to one of choosing a single real variable. The following theorem and its proof illustrate this possibility.

THEOREM 3. Suppose that the function $f(x, \theta)$: $\mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ is continuously differentiable and that, for all θ, $f(\cdot, \theta)$ is concave. Let $S \subset \mathbb{R}^n$ be compact and convex. Suppose, in addition, that, for all $a \in \mathbb{R}$, the solution set $x_1^*(\theta|a)$ is monotone nondecreasing in θ, where $x_1^*(\theta|a)$ is the first component optimum of the problem

$$\text{maximize } f(x, \theta) + ax_1 \quad \text{subject to } x \in S.$$

Then for all functions g: $\mathbb{R} \to \mathbb{R}$ and all $K \subset \mathbb{R}$, the first component optimum $x_1^*(\theta|g)$ of the problem

$$\text{maximize } f(x, \theta) + g(x_1) \quad \text{subject to } x \in S \text{ and } x_1 \in K$$

is also monotone nondecreasing in θ.

Proof. Let f^*: $\mathbb{R}^2 \to \mathbb{R}$ be defined by $f^*(z, \theta) = \max\{f(x, \theta)|x \in S, x_1 = z\}$ and let S_1 be the projection of S onto its first component. Then $x_1^*(\theta|a, b) = x_1^*(\theta|a)$ is the solution set of the one-dimensional problem

$$\text{maximize } f^*(x_1, \theta) + ax_1 + b \quad \text{subject to } x_1 \in S_1.$$

Set $V(x, y, \theta) = y + f^*(x_1, \theta)$ and $g(x_1|a, b) = ax_1 + b$. Apply theorem 1. Q.E.D.

Adding $g(x_1)$ to the objective and introducing the additional constraint that $x_1 \in K$ generally destroy the convexity of the original problem, so convexity cannot be essential for conclusions about comparative statics. Still, convexity of the subproblem in the two-stage maximization is sometimes helpful for establishing the requisite properties of the second-stage (reduced) problem.

To illustrate, consider the problem of maximizing the total benefit $\Sigma_{i=1}^n B_i(x_i)$ obtained from a finite amount of total resource x, where $\Sigma x_i \leq \bar{x}$ and $x_i \geq 0$ for all i. It is a standard result of convex programming that if each B_i is concave, then $x_1^*(\bar{x})$ is monotone nondecreasing. To relax the concavity of the overall problem, let us conduct a two-stage optimization. Thus define $B_{-1}(z) = \max \Sigma_{i\neq 1} B_i(x_i)$ subject to $\Sigma_{i\neq 1} x_i \leq z$ and, for all i, $x_i \geq 0$. Then the original problem is reduced to maximize $B_1(x_1) + B_{-1}(\bar{x} - x_1)$. With $V(x, y, \theta) = y + B_{-1}(\theta - x)$, condition 2 in theorem 2 indicates that a sufficient condition for comparative statics is that $B_{-1}(\cdot)$ is concave, and the theorem indicates that this condition is also necessary if B_1 is to be arbitrary. A sufficient condition for the concavity of B_{-1} is that the functions $B_2, \ldots, B_n$ (but not B_1) are concave. When $n = 2$, the concavity of B_2 is a necessary condition as well. Thus, in this example, concavity in the subproblem, but not in the overall problem, is necessary as well as sufficient for the comparative statics analysis.

V. Conclusion

My theorems assert that comparative statics of concave maximization models derived in highly specific contexts are sometimes reliable guides to the comparative statics of much more general contexts. The critical sufficient condition in each context is the same: it is condition 2 of theorem 1. Milgrom and Roberts (1994) obtain similar results for equilibrium models. They show that the necessary conditions for monotone comparative statics in certain narrow, two-parameter families of fixed-point problems are sufficient conditions for a much wider family of fixed-point problems. Together, these results suggest that comparative statics conclusions obtained in models with special simplifying assumptions can often be significantly generalized. The theorems help to distinguish the critical assumptions of an analysis from the other assumptions that simplify calculations but do not alter the qualitative comparative statics conclusions. In that way, the theorems improve our ability to develop useful models of parts of the economy and to interpret those models accurately.

References

Arrow, Kenneth J. "The Theory of Risk Aversion." In *Essays in the Theory of Risk-Bearing*. Chicago: Markham, 1971.

Milgrom, Paul, and Roberts, John. "Comparing Equilibria." *A.E.R.* 84 (June 1994).

Milgrom, Paul, and Shannon, Chris. "Monotone Comparative Statics." *Econometrica* 62 (January 1994): 157–80.

Rothschild, Michael, and Stiglitz, Joseph E. "Increasing Risk II: Its Economic Consequences." *J. Econ. Theory* 3 (March 1971): 66–84.

Sandmo, Agnar. "On the Theory of the Competitive Firm under Price Uncertainty." *A.E.R.* 61 (March 1971): 65–73.