

Robust Comparative Statics

© Susan Athey, Paul Milgrom, and John Roberts

DRAFT ONLY - DO NOT CITE OR CIRCULATE

Comments Welcome

This Draft: October, 1998

OUTLINE

Chapter 0: Introduction.

Part I: Univariate Optimization Problems

Chapter 1: An Introduction to the Theory.

Chapter 2: Single Variable Problems with Separable Contexts.

Chapter 3: Single Variable Problems with non-Separable Contexts.

Part II: Multivariate Optimization Problems.

Chapter 4: Multivariate Comparative Statics on Product Sets.

Chapter 5: Elementary Lattice Theory and Comparative Statics on Lattices.

Part III: Stochastic Optimization Problems.

Chapter 6: Supermodularity and monotonicity.

Chapter 7: Log-Supermodularity.

Chapter 8: Single Crossing Properties.

Part IV: Equilibrium Models.

Chapter 9: One dimensional models.

Chapter 10: Multidimensional models.

Chapter 11: Conclusion.

0 Introduction

This monograph describes a new approach to comparative statics analysis that has developed rapidly in the past several years. *Comparative statics* – the study of how the solutions of an economic model change as the model parameters and specification are changed – is important because (1) most of the testable predictions of economic theory are comparative statics predictions and (2) many economic equilibrium analyses are built from comparative statics analyses of the model’s components. The results of comparative statics analyses thus form the basis for much of our understanding of the behavior of the economy.

Comparative statics predictions can be qualitative or quantitative, but the focus of the new methods is on the former. Typical examples of qualitative predictions include assertions that increasing employment taxes or a minimum wage will increase unemployment, or that relaxing trade restrictions will cause industrial pollution to fall in developed countries and rise in less developed ones. Other examples include statements about how interactions among decisions in an economic system will multiply or attenuate certain effects, such as the prediction that a plant closing will reduce income in a community by more than the lost wages of the laid-off workers, or that a reduction in oil supplies will increase oil prices by more in the short-run than in the long-run.

Notice that each of these predictions is about how a change in some condition will affect other variables in the system without regard to all the other conditions that define the economic environment. That is, the economist who makes these general predictions intends for them to be *robust* – independent of the specification of the fine details of the environment. The objective of this monograph is to develop and present the logic of robust comparative statics.

Alternative Methods of Comparative Statics Analysis

Traditionally, there have been three main mathematical approaches to comparative statics analyses of optimization problems. The first relies on building models that can be solved explicitly, and then deriving the functional relationship between the exogenous parameters and the endogenous variables. The disadvantage of this “*explicit solutions*” approach is that it usually entails making severe assumptions for tractability and leaves open the question of how crucially the qualitative conclusions depend on those restrictions. This approach gives us few hints about whether its conclusions are robust.

The second, “*implicit functions*” approach is based on calculus, although its main features have been extended in the area of operations research known as “convex programming.” In optimization problems it involves assuming that the first-order conditions characterize the solution to the problem and that these define the optimizing choices as a differentiable function of the varying parameter. A formula giving the derivative of this implicit function in terms of the first and second derivatives of the objective function can be calculated, and one then seeks to find conditions that determine the sign of the derivative. This approach has wider applicability than

the explicit solutions approach, and in fact the Implicit Function Theorem from calculus actually gives sufficient conditions for the solution function to exist locally. Still, in order to apply this approach, the objective function must satisfy a number of conditions that are economically restrictive and that in most cases are actually unrelated to the comparative statics conclusions. For example, the objective function must be suitably differentiable and the choice variables must be continuous, which bars applications to discrete choice problems and to some other problems as well. In addition, it is common to assume that the objective is strictly concave because this guarantees that the optimum of the problem is unique for each value of the parameter.

The implicit functions approach has a number of advantages. One is that when it applies to a particular problem, it gives necessary and sufficient conditions for the comparative statics conclusion to hold in that problem. A second advantage is that it can be applied before particular functional forms are assumed, giving some insight into the generality of the conclusions. Further, it is not limited to optimization models; it can be adapted for use on comparative statics for equilibrium analysis as well. Finally, these methods are systematic in the sense that the way to apply them across different problems is standardized and easily taught.

The third approach to comparative statics analysis is sometimes referred to as the “*revealed preference*” approach, where results are derived directly from the inequalities that define a global optimum. This approach has been used with great success in price theory to obtain very general results. However, because it is unsystematic and relies on the analyst’s ingenuity in crafting an argument that works in the particular problem at hand, this approach tends to be used only when the assumptions required for the application of the implicit function approach are not satisfied.

Besides these three general approaches, there is a fourth approach using dual functions (the “indirect utility” and “indirect profit” functions). This approach generates very elegant demonstrations of the comparative statics properties of the response of a competitive firm to changing market prices or of a consumer to changing prices and income. This approach is quite powerful when the relevant parameters are prices, and dual functions may help empirical researchers to draw inferences about production functions from data about prices and profits. However, dual functions are not even defined unless the relevant parameters are prices, so the scope of application of the duality approach is inherently limited. Because we are interested in more widely applicable methods, we will not study the duality approach in this monograph.

Robust Comparative Statics and Critical Sufficient Conditions

The defining characteristic of the new approach is that it systematically seeks the best possible conditions for robust comparative statics conclusions. Here, “*robust*” means that the qualitative conclusions should remain unchanged when the model is altered or generalized in particular ways that the modeler has in mind. “*Best*” means that the condition should be the weakest condition to yield the desired kinds of robustness. We call these best conditions for robust conclusions “*critical sufficient conditions*.³”

Critical sufficient conditions are defined by splitting the economic model into a part that is isolated to be the focus of analysis and a second part about which we profess to make only very

“weak” assumptions, or no assumptions at all. For example, suppose that the government considers new regulations which alter the cost of training and benefits for low skill workers. Our analysis of the problem might focus on the portion of costs affected by the new regulations, while imposing only weak restrictions on other factor prices and the production possibility set. These latter restrictions combined with the way the two parts of the model fit together, in this case as components of the firm’s profits, define what we call the *context* of the analysis. A *critical sufficient condition* is the weakest condition on the isolated part of the model that imply the desired conclusion for all fully specified models in the context. For example, the critical sufficient condition for the new regulation to reduce employment among such workers could be that the regulation increases the incremental cost of employing them.

We emphasize robustness because it is our view that even our best models are always incomplete and approximate. We must therefore be concerned that the conclusions we draw from our models are valid predictions about a more complex reality. A conclusion that does not hold robustly may be a highly misleading indication about how the real economy would actually respond to the sort of changes being investigated. This leads us to ask: how robust is the particular comparative statics prediction? That is, over what range of models does the prediction hold? Also, once we have specified a range of models we believe is most relevant, what is the weakest assumption that supports a robust conclusion? The new methods provide a general framework and a collection of theorems to answer such questions.

One surprising finding of our analysis that contributes greatly to the appeal of this approach is that the extra price to be paid to obtain robust predictions is often quite small. Price theory provides familiar examples to illustrate how this can be so. Consider, for example, a two input model of the competitive firm. In such a model, the condition that the mixed partial derivative $f_{kl}(k,l)$ of the production function be non-negative for all pairs (k,l) , which is necessary and sufficient for the inputs to be complements at all input price levels in case the production function is quadratic, is still necessary and sufficient for that conclusion in the case of general smooth concave production functions, and also in the general case of smooth nonconcave functions.¹ A necessary and sufficient condition derived in a narrowly prescribed model is applicable without modification in much more general models. We shall find that this generalizability of suitably stated critical sufficient conditions from specific models to general ones is a widespread phenomenon.

Expanding the Domain of Analysis

Price-theoretic intuition and reasoning, which is based ultimately on comparative statics analysis, is a hallmark of the professional economist. One of the main things we hope to accomplish in our study of the comparative statics of optimization problems is to investigate just

¹ This assumes also that the set of feasible input levels for each input may be constrained, so that the “nonconvex parts” of the production function are potentially relevant. This point will be considered further in Part I of the monograph.

how far the intuitions of price theory can be extended. Before embarking on the formal analysis, it is instructive to think about the issue in intuitive terms. Which elements of the standard price theory model are “critical” for the models conclusions?

At an intuitive level it is not surprising that the main ideas of price theory do not in fact depend on all the special structural elements built into the standard mathematical models. For example the idea that taxing an input will lead a firm to use less of it does not depend on the linear pricing structure of the traditional, competitive model. It doesn’t matter whether the purchases are subject to quantity discounts, or whether the input is infinitely divisible, or whether the firm is a price-taker in the input or output markets, provided that none of these things also changes when the tax is introduced.

Not only is the linear structure of the model typically irrelevant for the comparative statics conclusions of price theory, but even the fact that the changing parameter is a price plays almost no role. For example, suppose input A is a (short- or long-run) substitute for input B for a particular firm with a convex technology. Then, the firm responds qualitatively in just the same way to an increase in the price of input A, as it does to a government regulation that rations or bans use of the input, or to the CEO’s decision to limit purchases of the input in retaliation for the input manufacturer’s introducing a competing product in another market. In each case, the firm responds by using more of input B (in the short- or long-run). *The ideas of substitutes and complements, though they originate in price theory, are not exclusively — or even principally — about prices.*

An important characteristic of prices as parameters in the theory of the firm is that each price interacts directly with just one choice variable. Prices are not the only parameters to have this property. Rationing or restricting the use of an input or expanding the market for a single output have the same characteristic. So do non-linear commodity tax rates. The corresponding parameter changes are just as easy to analyze as price changes. In contrast, a tax reduction or a technological change that not only reduces the marginal cost of input A but also reduces that of input B generally has ambiguous effects — generally, only a quantitative assessment of the effects allows a definite prediction to be made about the change in use of input A. Some older approaches, particularly the revealed preference approach and the duality approach, treat price changes differently from other parameter changes on the basis of the fact that they enter the objective function linearly. However, if the goal is to obtain robust comparative statics conclusions, any linear structure that may be present is just a distraction.

Since the *linearity* of profit functions play no special role in price theory, it follows that input and output *quantities* have no special characteristics as endogenous variables. Indeed, the main thing that distinguishes quantities from more general classes of choice variables is that they enter the firm’s objective linearly, being multiplied by prices.

Once one dispenses with the irrelevant parts of the structure of price theory, one can begin to apply its logic rigorously to a much wider range of variables, including ones that are not neatly quantifiable. In economic history, one can speak rigorously of developing *substitutes* for trade routes that have been closed by political or military developments. One can apply the same ideas

to institutions, such as international banking institutions that are *complementary* to long-distance trade or accounting systems that *complement* the emergence of large firms by allowing the owner to monitor parts of the organization that s/he can't physically visit. In industrial organization, one can speak of the *substitution* between inventory systems that buffer uncertain demand and communication systems that reduce the uncertainty about demand. One can also speak rigorously of the *complementarity* between the flexibility of a firm's manufacturing technology and the amount of variety in the firm's product line.

Of course, many economists already talk and write this way, sometimes carefully and correctly, and sometimes not. Our objective is to provide you with methods that let you use these terms and ideas in this broad fashion and yet in ways that are completely precise and that match the formal mathematics exactly. One of the keys will be to learn to use the terms "substitutes" and "complements" in the fashion of the examples just given, but in which the terms denote precisely the same mathematical relationships as they do when the objects of study are inputs to a production function.

Thinking Clearly about Comparative Statics

Besides expanding the domain of comparative statics analysis, the elimination of extraneous assumptions that these new methods allows has another advantage: By identifying the minimal assumptions needed for a robust comparative statics result, these methods provide a much more reliable guide to intuition.

The methods based on the implicit function theorem, which are by far the most commonly used of the traditional approaches, are burdened by restrictive assumptions about divisibility, convexity, differentiability and the like. Direct, revealed preference arguments long ago showed that such extra restrictions are unnecessary for the comparative statics of demand theory, and we will show below that they play no role in any of much wider class of problems. Yet the use of these extra conditions in traditional formal analyses has sometimes convinced researchers to incorporate them in their intuitive explanations of comparative statics results. The following is typical of the kind of mistake most frequently made.²

Consider the benefit-cost problem of maximizing $B(x; \theta) - C(x)$ by choice of $x \in R$, with parameter θ . If (1) x is divisible, (2) the benefit function is strictly concave in x and differentiable and (3) the cost function is convex and differentiable, then the solution is characterized by the first-order condition $B_x(x; \theta) = C'(x)$, where the subscript denotes a partial derivative. A formal convex programming analysis shows that if the cross-partial $B_{x\theta} > 0$, then the optimum $x^*(\theta)$ is an increasing function of θ . A misguided intuitive interpretation might go like this: "If $B - C$ is strictly concave and x is divisible, then the optimal x equates the marginal benefit $B_x(x; \theta)$ to the marginal cost $C'(x)$. If increasing the parameter θ increases the marginal benefit B_x for each x ,

² Examples of such misguided arguments are not difficult to find in the books and articles of even leading economists.

then the new solution must involve a higher value of the marginal cost C' . Since C is convex, so that C' is an increasing function of x , that requires that x increase. Thus, if $B_{x\theta} > 0$, then an increase in θ leads to an increase in x .”

This purported explanation misguides the intuition by emphasizing the assumption that C is convex. For if $B_{x\theta} > 0$, then, as either direct arguments of the revealed-preference sort or robust methods arguments make clear, $x^*(\theta)$ is monotone nondecreasing regardless of whether C is convex, regardless of whether $B-C$ is concave or even quasiconcave, and regardless of whether x is divisible. Notice that if x is limited to a discrete set or if the optimum lies on a boundary of the constraint set, then the first-order condition $B_x = C'$ need not hold. That this fact is of no consequence at all for the comparative statics conclusion casts doubt on the usefulness of first-order conditions for intuitive economics. The purported explanation, as plausible as it may sound, is an example of misdirection: it highlights extraneous assumptions that are quite unrelated to the monotonicity conclusion.

A second insight concerns the basic nature of critical sufficient conditions for robust comparative statics. To illustrate, consider the modeler's choice about whether to measure the degree of uncertainty in some model by the variance or the standard deviation of some random variable. This formulation choice cannot affect any conclusion about monotonicity of the degree of uncertainty with respect to changes in the model parameters, because the variance and standard deviation are increasing functions of one another, so monotonicity of an endogenous variable with respect to one measure of uncertainty is equivalent to monotonicity with respect to the other. Any condition that prevents the rescaling of a variable in this way can always be relaxed, and therefore cannot be a critical condition for a wide family of models. For example, a condition that requires that the objective function $f(x;\sigma^2)$ be quasiconcave in its two arguments cannot be a critical sufficient condition for a wide family of models because it is not equivalent to the condition that the function g , defined as $g(x;\sigma) \equiv f(x;\sigma^2)$, be quasiconcave, despite the equivalence of the two formulations from the point of view of comparative statics.

In general, no multivariate convexity restriction can ever be (part of) a critical sufficient condition for a sufficiently broad family of models. Such conditions are not preserved when variables are subjected to (non-linear) increasing transformations, but the monotonicity conclusions of comparative statics are preserved under such transformations. For the same reason, contraction mapping conditions cannot be critical sufficient conditions for comparative statics in fixed point problems, because the defining condition that $|f(x)-f(y)| \geq \alpha|x-y|$ for some $\alpha < 1$ is not invariant to (non-linear) monotone transformations of the domain of f . The only conditions that are candidates to be critical conditions are ones whose validity is unaffected when each variable and parameter is replaced by some arbitrary increasing function of itself. That rules out a great many of the conditions that have formerly been emphasized in comparative statics analyses.

The Philosophy of Economic Modeling

Besides leading to confused analyses, the custom of including extraneous assumptions in our basic methods has been damaging in another important way: It lulls modelers into a willingness

to accept truly outrageous assumptions with hardly a raised eyebrow. Many of the assumptions that are built into economic models are made because they are asserted to be necessary “for tractability,” despite their being sharply in conflict with observations about the world. Modelers routinely assume that production sets are convex, despite Adam Smith’s observations more than two centuries ago that there are ubiquitous increasing returns to scale. Modelers assume that consumers maximize, that governments represent the median voter, that managers serve the interests of shareholders, that expectations about future prices and behavior are fulfilled. All of these assumptions are typically described in seminar presentations as “standard,” as though that somehow justified using them. Moreover, we routinely omit effects from our models that are plausibly as large as, or even larger than, the effects we have included. How is a serious person to regard models whose analyses are based on assumptions such as these?

Questions about the assumptions of economics have drawn the attention of several Nobel laureates. Ronald Coase (1937) endorsed Joan Robinson’s view that the assumptions of an economic analysis need to balance descriptive accuracy against tractability. Others, however, have taken more extreme views. Milton Friedman (1953) has even suggested that unrealism in assumptions is a virtue of a theory: “Truly important and significant hypotheses will be found to have ‘assumptions’ that are wildly inaccurate descriptive representations of reality, and, in general, the more significant the theory, the more unrealistic the assumptions.” Herbert Simon (1963) attacked Friedman’s view, arguing that “if the conditions of the real world approximate sufficiently well the assumptions of an ideal type, the derivations from these assumptions will be approximately correct.... Unreality of premises is not a virtue in scientific theory; it is a necessary evil — a concession to the finite computing capacity of the scientist that is made tolerable by the principle of continuity of approximation.”

Simon’s “principle of continuity of approximation” is not often helpful, however, in deciding which assumptions are realistic enough. How is one to measure, for example, whether people deviate “very far” from rational behavior, or even whether one irrational behavior is “more nearly rational” than another. Besides, modelers are typically highly uncertain about some of the facts of a situation and would like to identify conclusions that they can be confident are true, or nearly true, even if they have no hope of specifying a detailed model that is “nearly correct,” whatever that might mean. It is far more important, we think, to be able to identify the assumptions on which the various conclusions of the model critically depend.

A view about the role of assumptions that is close to our own has been eloquently expressed by Robert Solow (1956): “All theory depends on assumptions which are not quite true. That is what makes it theory. The art of successful theorizing is to make the inevitable simplifying assumptions in such a way that the final results are not very sensitive. A ‘crucial’ assumption is one on which the conclusions do depend sensitively, and it is important that crucial assumptions be reasonably realistic. When the results of a theory seem to flow specifically from a special crucial assumption, then if the assumption is dubious, the results are suspect.”

Of course, Solow’s distinction is not helpful if one cannot distinguish what he calls the “crucial” assumptions from those that are merely simplifying. The new theory’s emphasis on *critical* sufficient conditions – the weakest sufficient conditions for robust conclusions – is

particularly important to economists, like ourselves, who share this philosophy of economic modeling.

The Mathematics of Robust Comparative Statics

The methods we develop here involve some mathematics that has not been a standard part of the economist's tool kit. This mathematics is specifically focused on monotonicity, which is the central property of concern to us in doing qualitative comparative statics analyses. Monotonicity is, of course, a matter of *order*: Does a *higher* level of an exogenous variable lead to a *higher* level of the endogenous variables? Does using *more* of this input favor using *more* of some other? Is the long-run response *larger* than the short-run adjustment?

We shall find that giving robust answers to such questions entails using ideas from a branch of mathematics known as “*lattice theory*,” a theory which takes as its primitives a set and a partial order which can be used to compare pairs of elements of that set. Although ideas from this theory are ubiquitous in our analysis, we postpone introducing the notation of lattice theory until Chapter 5 by limiting the analysis to problems in which the choice set is a subset of \mathbb{R}^N . The general theory developed in Chapter 5 is useful for studying and understanding constrained optimization problems in \mathbb{R}^N as well as problems in which the choice variables are not finite vectors of real numbers. Examples of such choice variables include investment plans in an infinite horizon investment model, compensation contracts in a principal-agent model with infinitely many potential outcomes, and sets of coalition partners in a coalition formation problem. Besides expanding the scope of the theory, the lattice theoretic treatment deepens our understanding of the results obtained in Chapters 2-4.

The Road Ahead

The early chapters of our monograph — the first two Parts, consisting of Chapters 1 through 5 — develop the theory for comparative statics on optimization problems. This is the most completely developed part of the new theory, and the results allow one to do comparative statics in all sorts of economic models as easily and intuitively as for the standard models of price theory, for which refined methods and intuitions are already available. Later chapters take up the search for critical sufficient conditions in stochastic optimization problems, in games, and in other equilibrium models. There are both positive and negative conclusions here, and both are useful. The negative conclusions highlight the non-robustness of comparative statics derived from special kinds of equilibrium models, casting doubt on the generalizability of the conclusions of those models. The positive results point the way toward generalizing some standard results and incorporating them into more complex models.

Part I: Univariate Optimization Problems

Our analysis of univariate optimization problems is divided into three chapters. The first chapter provides an introduction to the theory in the context of three familiar economic problems. These examples illustrate the ideas of robustness and “critical sufficient conditions” in problems to which the implicit function methods can sometimes be applied, facilitating comparison of the old and new methods. The second chapter then provides a more complete and formal analysis of problems which have an additively or multiplicatively separable structure, developing ways to analyze comparative statics in problems with multiple optima and without differentiability assumptions. The last chapter provides a similar analysis for problems which are not additively or multiplicatively separable.

1 An Introduction to the Theory

We open our introduction to univariate optimization problems with three simple maximization problems of the kind that are the staple of formal economic analysis.

$$(P1) \quad \underset{x}{\text{Maximize}} \ U(x, \theta)F(x)$$

$$(P2) \quad \begin{aligned} & \underset{x \in [0, \bar{x}]}{\text{Maximize}} \ p\theta - rx - wl \\ & \text{s.t. } F(k, x) = \theta \end{aligned}$$

$$(P3) \quad \underset{x \in [0, \theta]}{\text{Maximize}} \ U(\theta - x, g(x))$$

The first problem is drawn from bidding theory. The decision maker chooses a bid x which has a probability $F(x)$ of being the highest bid. If its bid is highest, the bidder enjoys utility of $U(x, \theta)$, where θ is the bidder's type. Utility is normalized to be zero when the bidder loses. The analysis of whether the optimal bidding function $x = b^*(\theta)$ is monotonic is a first step in determining the equilibrium of the bidding model, which in turn determines the distribution F . Thus, F is endogenous to the equilibrium problem and so its properties cannot be specified in advance (when the individual optimization problems are being solved). The objective is to identify the weakest condition on U that implies the comparative statics conclusion for all distribution functions F , or for as wide range of distribution functions F as possible.

Problem (P2) is from the neoclassical theory of the competitive firm. A firm with production function F chooses capital (x) and labor (l) at prices r and w to meet an output target θ at minimum cost. The firm's revenues, given by $p\theta$, are exogenous for any given output target. A goal of comparative statics analysis in this problem is to determine when the cost-minimizing quantity of capital (x) that will be used in producing output is a nondecreasing function of the output target, that is, to determine whether capital is a “normal input.” An unstated tradition of price theory is to seek results that are not sensitive to the particular specification of prices. In that tradition, one seeks the weakest possible restriction on the production function F that implies the normality of the capital input regardless of the input prices r and w .

Problem (P3) is the classic two-good utility maximization problem in which the consumer

can transform some of the units of her endowment θ of good 1 into units of good 2 according to the transformation function g . To represent an ordinary budget constraint, one may take g to be any increasing linear function $\gamma + \rho x$, where x is the consumer's sales of good 1, $\rho \geq 0$ is the price of good 1 in units of good 2, and γ is the agent's initial endowment of good 2. In a two-period consumption-savings problem with goods 1 and 2 representing consumption in periods 1 and 2, x would be the savings and g would be the production or investment returns function linking current savings to future consumption opportunities. The problem also includes the a multiperiod consumption-savings problem. In that case, g could represent the “maximum utility of future consumption after the initial period,” which depends on future consumption and production opportunities. In each interpretation, one comparative statics question is: Under what conditions do increases in the endowment θ of good 1 lead to increases in the “savings” x ?

Notice that the particular interpretation of problem (P3) affects the kind of robustness that one might want. For example, in the budget problem interpretation, one might want a conclusion that does not depend on the particular specification of the goods prices. Then, the conclusion should hold for all prices ρ , or at least for a large range of prices. If the problem has linear budget constraints but multiple periods and $g(x)$ represents the maximum utility of future consumption when the initial level of savings is x , then one would want the conclusion to hold for any function g that could emerge from that kind of calculation. If one is studying a Robinson Crusoe economy with production that is subject to constant or decreasing returns to scale, one might want to know when the conclusion holds for all increasing, concave functions g . If general nonconvexities in production are also a concern, one might want a conclusion that holds for all functions g , or all increasing functions g . We take it as our starting point that the modeler has at least a rough idea about the class of functions g across which the comparative statics conclusion should hold.

In all three examples, what is wanted is a comparative statics result that holds not for a single problem specification but for a range of specifications. The task we have set for ourselves is to find the weakest possible restrictions on some elements of the problem that implies a conclusion with the desired degree of robustness. In problem (P1), we seek the weakest restriction on U that implies the desired monotonicity for all functions F , or all F in some large class. In (P2), we seek the weakest restrictions on F that imply the conclusion for all input price pairs (r, w) . In problem (P3), for each possible restriction on g , we seek to identify the weakest condition on U that implies the desired conclusion for all functions g satisfying the restriction.

Thus, in each of (P1)-(P3), we are interested in a *family of problems*. We will refer to such a family as the *context* in which we would like a particular comparative statics result to hold. In general, we will use $g()$ to denote the part of the problem which varies across members of the context. For example, in (P1) $g(x) = F(x)$.

We begin by analyzing problem (P1) using the traditional implicit function approach as we expect it would be applied by a skilled analyst. Such an analyst might start by limiting attention to the domain where the objective is non-negative and working with the logarithm of the objective. Because the logarithm function is monotonically increasing, this transformation does not change the optimizing bid for any θ and F and so does not alter the comparative statics analysis. We assume that there is an interior optimum, that the functions are differentiable as

many times as needed and that $U_x < 0$. The first-order optimality condition then characterizes the optimum for a class of problems including all those for which $\log(U) + \log(F)$ is strictly concave in x . The first-order optimality condition is:

$$\frac{U_x(x, \theta)}{U(x, \theta)} + \frac{F'(x)}{F(x)} = 0.$$

When the implicit function theorem applies, the maximizing bid (a function of θ and F) satisfies

$$\frac{\partial b}{\partial \theta} = \frac{-\partial^2 \ln U / \partial x \partial \theta}{\partial^2 \ln U / \partial x^2 + \partial^2 \ln F / \partial x^2}$$

and, according to the second-order optimality condition, the denominator of the fraction is negative. Hence, $\partial b^*(\theta; F) / \partial \theta \geq 0$ if and only if the condition $\partial^2 \log(U(x, \theta)) / \partial x \partial \theta \geq 0$ holds at $x = b^*(\theta; F)$.

In common practice, the formal analysis of monotonicity of b^* would cease at this point. Sometimes, however, the analyst may study robust comparative statics, observing that if the condition on $\log(U)$ holds globally, that is for all pairs (x, θ) , then b^* is nondecreasing in x for *all* F for which the implicit function approach applies (or, more precisely, for all F such that there is a uniquely defined, differentiable function satisfying the first-order conditions, which were assumed to characterize the solution). In other words, $\partial^2 \log(U(x, \theta)) / \partial x \partial \theta \geq 0$ is a sufficient condition for the comparative statics result to hold across the *context* defined by varying F as just described.

The new approach asks some additional questions about robust comparative statics. Are there alternative useful sufficient conditions, or is the global condition identified above the weakest one (at least in the case where the log-utility is concave in x)? Could a weaker condition be found if one sought more limited robustness by restricting the functional form of F ? Is the global condition identified above still sufficient if we allow a more general class of functions F for which the implicit function approach does not apply?

For the first question, one can get an answer by examining the first order condition. Assume that $\log(U)$ is a concave function of x for each θ . Then, for any utility function and any pair (x, θ) , there is some log-concave probability distribution F such that the first-order condition is satisfied uniquely at (x, θ) . (Since U is decreasing in x and F is increasing in x , one can simply set $F'(x)/F(x)$ to be the positive number necessary to make the condition hold at (x, θ) and set F elsewhere to make it a log-concave probability distribution.) Consequently, if the condition that $\partial^2 \log(U(x, \theta)) / \partial x \partial \theta \geq 0$ ever fails, it is not true that b^* is always nondecreasing in θ .

Could limiting the form of F lead to a weaker sufficient condition? We will treat the general question later, but the traditional approach is to study particular examples. Thus, suppose the class of probability distributions is large enough to include the exponential distributions given by $F(x) = 1 - \exp(-rx)$ for some $r > 0$. Referring back to the first-order optimality condition, one can see that for any values of x and θ and any function U , there is some exponential parameter r such

that (x, θ) satisfies the first-order condition. Thus, for a comparative statics conclusion in this model to apply for all F in the relevant class of distributions, it is necessary that $\partial^2 \log(U(x, \theta)) / \partial x \partial \theta \geq 0$ hold for *all* pairs (x, θ) such that $U(x, \theta) > 0$.

This emphasis on the exponential family is, of course, entirely *ad hoc*. Insisting that the exponential family be included is just one way to ensure that the family of possible distributions be “large enough.” Still, the conclusion derived this way is a representative one: requiring that the class of distributions be large in the sense of including some suitable one-parameter family of distributions will prove sufficient to establish the necessity of the mixed partial derivative condition. We will treat this issue more completely in Chapter 2.

The final issue about whether the assumptions needed for the implicit function approach can be relaxed without affecting the monotonicity of b^* cannot, of course, be examined using the implicit function approach.

The new methods work by developing a set of theorems that classify problems according to the kind of robustness that is desired and then specifying necessary and sufficient conditions for robust comparative statics on those problems. What we shall do next is to describe one such theorem that is quite useful for the problem just examined. The theorem applies to general problems of the following form:

$$\underset{x \in S}{\text{Maximize}} f(x, \theta) + g(x) \quad (1.1)$$

Notice that the objective consists of the sum two terms (or two groups of terms), one of which involves the choice variable and the parameter and the second of which involves only the choice variable. This form includes the bidding problem, in which the first term is $\log(U(x, \theta))$ and the second term is $\log(F(x))$. As we shall later see, objectives of this form are quite common in economic modeling. Our approach will be to seek to identify conditions on the function f for robust comparative statics of the choice x in the parameter θ : this is natural, because all the interaction between the two variables occurs in f . Meanwhile, we vary g to generate the context.

Let $X^*(\theta; g)$ denote the set of optimizers of (1.1) and let $x^H(\theta; g)$ and $x^L(\theta; g)$ the largest and smallest elements of $X^*(\theta; g)$, each defined on the parameter domain where such an element exists.¹

¹ Thus, $x^H(\cdot; g)$ is a function whose domain, for each g , is the subset of the parameter space Θ for which a largest maximizer exists and, for each g , $x^L(\cdot; g)$ is another function defined on the possibly different subset of Θ on which a smallest maximizer exists. The assertion in Theorem 1.1 that these functions are nondecreasing means, of course, that they are nondecreasing on the domains on which they are defined.

Theorem 1.1. Assume the following about the problem(1.1):

(S1.1) $S \subseteq \mathfrak{R}$.

(F1.1) The function f is twice continuously differentiable.

(G1.1) No restrictions are imposed on g .

Then, $x^H(\theta; g)$ is nondecreasing in θ for all such functions g if and only if $\frac{\partial^2}{\partial x \partial \theta} f(x, \theta) \geq 0$ for all pairs (x, θ) . Similarly, $x^L(\theta; g)$ is nondecreasing in θ for all such functions g if and only if $\frac{\partial^2}{\partial x \partial \theta} f(x, \theta) \geq 0$ for all pairs (x, θ) .

Theorem 1.1 (which, like the other theorems in this chapter, will be proved as part of a more general theorem in the next two chapters) is representative of the kind of comparative statics theorem offered in this monograph. Each theorem opens with a set of conditions like (S1.1), (F1.1) and (G1.1) to indicate the domain on which it applies. In each case, separate assumptions are made about various elements of the formulation; these assumptions together describe the context, or set of problems, for which we would like the conclusion to hold. The conclusion of each theorem identifies what we call a “critical sufficient condition,” which is the weakest sufficient condition for the desired conclusion. In this case, the theorem identifies the weakest sufficient condition to imply that the extremal maximizers of (1.1) are nondecreasing in θ regardless of the specification of g . That is, the theorem gives critical sufficient conditions for the conclusion to hold in the context defined by fixing f and S but varying g across the family of *all* functions on the real line.

There are some technical details to be accounted for in understanding the theorem. First is the issue of multiple optima. We will give a more complete treatment of that issue later. For now, we just select the largest and smallest maximizers as the focus of our attention. Second is the issue of existence. In this problem, a sufficient condition for existence of a largest optimum is that S be finite or that S be compact and g be upper semi-continuous. For analytic clarity, however, we adopt a style here and throughout this book in which we separate the assumptions useful to ensure existence of a solution – in this case an optimum – from assumptions needed for comparative statics. We do this by simply stating our results so that they apply whenever an optimum exists.

Theorem 1.1 already takes some of the steps we indicated were desirable in our discussion of problem (P1). The primary advance is that it eliminates the restrictions on the function $f + g$ that were necessary for application of the implicit function approach, allowing discrete choice variables (through the specification of S), multiple optima, and discontinuous changes in the optimum. Thus, we already have answered one of the open questions, namely, we have established that the sufficient condition $\partial^2 \log(U(x, \theta)) / \partial x \partial \theta \geq 0$ is still sufficient even when assumptions needed for the implicit function approach do not hold. However, we shall want to take our study of robustness even further.

On one hand, we will want to consider the effect of imposing restrictions on the function g (that is, considering a more limited context) in order to identify when such restrictions can lead to

weaker sufficient conditions. A remarkable conclusion of the theory is that it takes very restrictive assumptions about the form of g to change the critical sufficient condition of Theorem 1.1.

On the other hand, we will seek additional generality of several kinds. We shall eliminate the assumption that f is differentiable, incorporate a parameter into the constraint sets, treat the question of strict monotonicity, and give a more satisfactory treatment of non-unique optima. And, of course, we seek to treat problems like (P3) in which there is no additive structure as well as problems with a multiplicity of choice variables.

In order to highlight the first of these issues – the robustness of the conclusions when the set of possible functions g is restricted – let us turn to the example of problem (P2). To apply our theorems, we first rewrite this problem in a form similar to (1.1). To facilitate that, we will maintain the assumption that the function F is increasing, that is, that inputs are productive. Let $L(x, \theta)$ be the isoquant function denoting the minimum amount of labor required to produce output θ when the capital input is fixed at x , that is, $L(x, \theta) = \min\{l | F(x, l) = \theta\}$. Then the objective to be maximized is this:

$$\underset{x}{\text{Max}} \ p\theta - wL(x, \theta) - rx$$

This problem has the form specified in (1.1), with one group of terms $f(x, \theta) = p\theta - L(x, \theta)$ that depends on the choice variable and the parameter, and a separable term $g(x) = -rx$ that depends on the choice variable but not the parameter.

First, consider the question of sufficient conditions for comparative statics predictions. Theorem 1.1 tells us that, if the isoquant function is appropriately differentiable, then capital is a normal input if the isoquant function satisfies $L_{x\theta} < 0$. That is, if increasing the level of output required makes the isoquant steeper at a fixed level of capital, then the demand for capital will increase with output. What does this condition mean in terms of conditions on the production function? First, observe that an isoquant function satisfies $F(x, L(x, \theta)) = \theta$. Thus, if F is continuously differentiable and the isoquant is well-defined, we conclude that $-L_x(x, \theta) = F_x(x, l)/F_l(x, l)|_{l=L(x, \theta)}$, which will be increasing in θ if $F_x(x, l)/F_l(x, l)$ is increasing in labor for a given level of capital. This is the standard condition on the marginal rate of technical substitution from price theory.

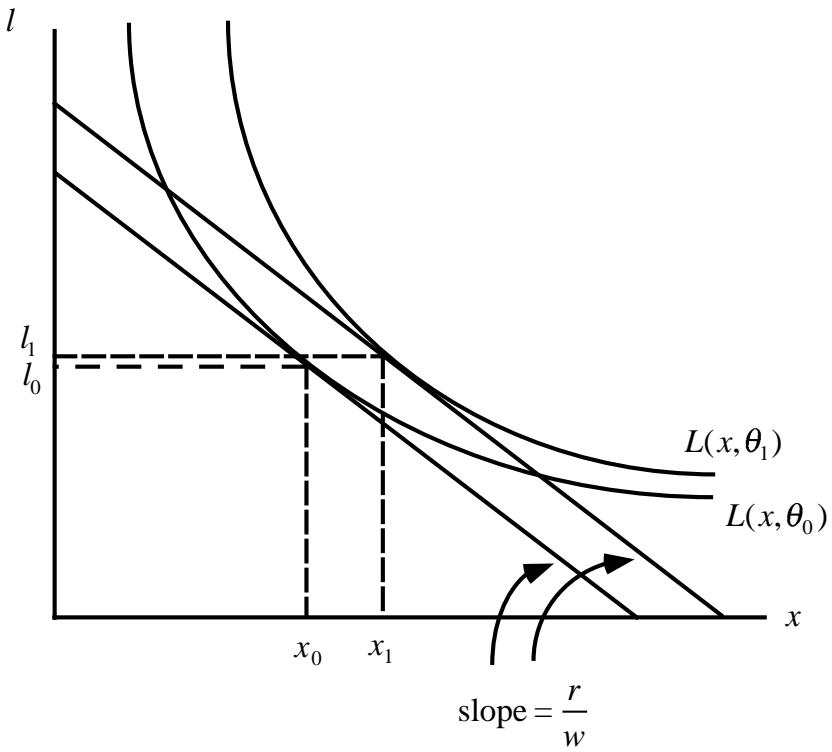


Figure 1.1 When the isoquant $L(x, \theta)$ becomes steeper when output (θ) increases, the choice of capital (x) increases with output.

Is there a weaker condition which will yield this conclusion? We cannot apply Theorem 1.1 directly to answer these questions, for several reasons. To start, the function g is not unrestricted in problem (P2): it is decreasing and linear. To add further structure to the problem, it is common to assume that the inputs are divisible and that the isoquant function L is convex in x . In principle, this extra condition might allow the analyst to obtain a weaker sufficient condition on L . But does it? This question is posed in Theorem 1.2 using the abstract formulation of problem (1.1), and the answer is found to be negative.

Theorem 1.2. Assume the following about the problem (1.1):

(S1.2) $S \subseteq \mathcal{R}$ is a convex set.

(F1.2) The function f is twice continuously differentiable and increasing and concave in x .

(G1.2) The function g is decreasing and linear.

Then, $x^H(\theta; g)$ is nondecreasing in θ for all such functions g if and only if $\frac{\partial^2}{\partial x \partial \theta} f(x, \theta) \geq 0$ for all pairs (x, θ) . Similarly, $x^L(\theta; g)$ is nondecreasing in θ for all such functions g if and only if $\frac{\partial^2}{\partial x \partial \theta} f(x, \theta) \geq 0$ for all pairs (x, θ) .

Applying Theorem 1.2 to problem (P2) leads to the conclusion that if the isoquant is convex, then capital is a normal input if and only if the isoquant function satisfies $L_{x\theta} < 0$, that is, if and

only if $F_x(x,l)/F_l(x,l)$ is increasing in l for any given level x of capital. A typical intuitive argument based on the standard theory might go as follows: If θ_0 is the initial output level, the optimal capital-labor combination (x_0, l_0) must satisfy $r/w = F_x(x_0, l_0)/F_l(x_0, l_0)$. Further, convexity implies that $F_x(x,l)/F_l(x,l)$ is diminishing in x along an isoquant. Since moving to a higher isoquant ($\theta=\theta_1$) leads to an *increase* in marginal rate of substitution at the old choice of capital, then to get to a new optimum, capital must be increased to $x_1 > x_0$ in order to diminish the marginal rate of substitution so that the new optimal point, (x_1, l_1) , satisfies $r/w = F_x(x_1, l_1)/F_l(x_1, l_1)$.

However, this cannot be a correct intuition for the robust comparative statics conclusion, since Theorem 1.1 tells us that $L_{x\theta} < 0$ is sufficient for comparative statics *even in the absence of convexity assumptions*. Using Theorems 1.1 and 1.2 together, the new methods lead to an improvement over the best previously known result. The assumptions that the isoquants are convex and capital is divisible, which were needed in older treatments of this problem, are here found not to be necessary for the comparative statics conclusion. Furthermore, imposing those extra assumptions does *not* lead to any weakening of the sufficient condition.

Next, we consider problem (P3). We do this in two steps. First, we will impose some additional structure on the problem, namely additive separability between utility for consumption of good 1 and good 2, and we will apply Theorem 1.1 to that special case. We will then treat the more general, non-separable case, introducing several new theorems in the process.

We begin by assuming that the agent's utility can be written as the sum of his first period utility of consumption, and the utility of future consumption. That is, $U(\theta - x, g(x)) = U^1(\theta - x) + g(x)$, where good 1 is current consumption and g is the agent's (endogenously determined) value function for consumption in all future periods, given that the agent saves x in period 1. Thus, in order to determine the properties of g necessary for the application of the implicit function theorem, we would need to perform further analysis of an appropriately defined dynamic programming problem, analysis which might lead to further restrictions on the problem. For the moment, however, let us simply assume that U^1 and g are concave and differentiable everywhere. A more careful analysis could sharpen these conditions slightly, but some assumptions of this sort are necessary to use the implicit function approach.

If we apply the implicit function theorem to this problem, the concavity assumptions lead us to conclude that $x^*(\cdot) \geq 0$ everywhere. Under the implicit functions approach, this conclusion is a joint implication of the various concavity and differentiability assumptions. One cannot easily distinguish the separate roles that the various concavity assumptions play.

Our approach to this problem, using Theorem 1.1, does better. According to the theorem, the conclusion that savings increases with initial wealth depends critically on the concavity assumption on first period utility, since this assumption is equivalent to $\frac{\partial^2}{\partial x \partial \theta} U^1(x, \theta) \geq 0$. On the other hand, the concavity assumption on the value function g is completely irrelevant: the conclusion holds for any function g . We will show in Chapter 2 that the differentiability assumptions are irrelevant as well. So the application of Theorem 1.1 gives a more precise answer about what conditions are required for savings to increase with initial wealth. Concavity

of first period utility for consumption is the critical sufficient condition, and nothing about g is important at all. In contrast, the question cannot even be attempted using the implicit function theorem without some restrictions on g , restrictions which we have here proved are irrelevant for the comparative statics question.

We now turn to the more general form of problem (P3), dropping the assumption of additive separability, so that the utility function is given by $U(\theta - x, g(x))$. This represents the case where the agent's marginal utility for current consumption might depend on the future wealth available for consumption and investment. We then repeat the question, what are the weakest assumptions on the utility function so that the choice of savings is increasing in the initial endowment?

As before, we can get a certain distance towards answering this question through the use of the implicit function approach, where it applies. Let us consider an even more general formulation of the problem, as follows:

$$\max_{x \in S} V(x, g(x), \theta) \quad (1.2)$$

If g and V are each suitably differentiable and further are such that the objective function has a unique, interior local maximum satisfying the first and second order conditions, then the optimal choice x^* can be studied by studying those conditions. The first order condition is:

$$V_1(x^*, g(x^*), \theta) + V_2(x^*, g(x^*), \theta) \cdot g'(x^*) = 0$$

and the second order condition is:

$$\begin{aligned} & V_{11}(x^*, g(x^*), \theta) + V_{12}(x^*, g(x^*), \theta)g'(x^*) \\ & + V_{21}(x^*, g(x^*), \theta)g'(x^*) + V_{22}(x^*, g(x^*), \theta)(g'(x^*))^2 + V_2(x^*, g(x^*), \theta)g''(x^*) < 0 \end{aligned}$$

Applying the implicit function theorem to the first order condition and solving for $x^{*''}$ (suppressing function arguments), we find that

$$x^{*''}(\theta) = -[V_{13}(x^*, g(x^*), \theta) + V_{23}(x^*, g(x^*), \theta) \cdot g'(x^*)] / \text{Negative Term}.$$

where “*Negative Term*” is the term that was determined to be negative in the second order condition. Provided that $V_2 \neq 0$, we may solve the first order condition to find $g'(x^*) = -V_1/V_2$ and substitute that into the preceding expression. That leads to:

$$\begin{aligned} \text{sign}[x^{*''}(\theta)] &= \text{sign} \left[V_{13}(x^*, g(x^*), \theta) - V_{23}(x^*, g(x^*), \theta) \cdot \frac{V_1(x^*, g(x^*), \theta)}{V_2(x^*, g(x^*), \theta)} \right] \\ &= \text{sign} \left[\frac{\partial}{\partial \theta} \frac{V_1}{|V_2|}(x^*, g(x^*), \theta) \right] \end{aligned}$$

Thus, a sufficient condition for $x^{*''}(\theta)$ to be non-negative for all g is that $V_1(x, y, \theta)/|V_2(x, y, \theta)|$ is nondecreasing in θ for all (x, y) . Using the implicit functions approach, this result can be understood as a consequence of the first and second order conditions for maximization.

However, as we argued above, there are many economic problems where we do not have enough a priori information about the function g to apply the implicit function theorem. When g is endogenously determined, we may not be able to rule out non-concavities of the objective function or nondifferentiabilities of the function g . In such cases, the implicit function theorem cannot instruct us about how the optimal choice of x changes with θ . However, it turns out that, just as in the case of the additively separable problem above, the additional concavity and differentiability assumptions required for the implicit function theorem are irrelevant for the comparative statics conclusions. We will, however, require that $V(x,y,\theta)$ has well-behaved x - y indifference curves, that is, $\{(x,y) | V(x,y,\theta)=k\}$ is a closed curve for every θ and k . In other words, the projection of each curve onto the x -axis is a convex set; this rules out the possibility that the indifference curve “becomes infinite” at some interior x .

Theorem 1.3. Assume the following about the problem (2):

(S1.3) No restrictions on S .

(V1.3) The function $V(x,y,\theta)$ is continuously differentiable in its first two arguments, $V_2 \neq 0$, and the indifference set $\{(x,y) | V(x,y,\theta)=k\}$ is a closed curve for all (θ,k) .

(G1.3) No restrictions imposed on g .

Then, $x^H(\theta, g)$ is nondecreasing in θ for all functions g if and only if V_1/V_2 is everywhere nondecreasing in θ . Similarly, $x^L(\theta, g)$ is nondecreasing in θ for all functions g if and only if V_1/V_2 is everywhere nondecreasing in θ .

In Chapter 3, we will provide the proof of this result, and explain the intuition graphically. For now, let us show how this result applies to our savings decision problem. Letting $V(x,g(x),\theta) = U(\theta-x, g(x))$, and assuming that the utility function is continuously differentiable and strictly increasing in both arguments ($U_1 > 0$ and $U_2 > 0$) and further satisfies the closed indifference curve condition, we can apply Theorem 1.3. The critical sufficient condition is that $-U_1/U_2$ (the marginal rate of substitution of current for future wealth) is nondecreasing in initial wealth. As long as this condition holds, savings will increase in initial wealth no matter what form g takes.

Now let us consider a different interpretation of the problem, an interpretation which involves substantially more structure. Suppose that goods 1 and 2 are both consumption goods, and the agent is endowed with θ units of good 1 and γ units of good 2. Further, the ratio of the price of good 1 to the price of good 2 is given by ρ . The agent can buy or sell each of these goods at those prices on a competitive market; we let x indicate the amount of good 1 which is sold. Thus, the agent’s maximization problem is

$$\max_{x \in [0, \theta]} U(\theta - x, \gamma + \rho x)$$

The comparative statics question is, how does the amount of good 1 sold change as the initial endowment of good one increases? This problem takes a very similar form to the consumption-

savings problem, except that we have made a functional form assumption about transformation function g . Instead of allowing g to be any function, we have now restricted ourselves to a very special class of functions which are linear in x . We might wonder whether this extra structure will lead us to less restrictive sufficient conditions for our comparative statics conclusion. A surprising result, which is analogous to our result from the additively separable case, is that the additional structure does not relax our requirements on the utility function at all. We will state our result in terms of the general problem, (2), posed above.

Theorem 1.4. Assume the following about the problem (2):

(S1.4) *No restrictions on S .*

(V1.4) *The function V is continuously differentiable in its first two arguments and $V_2 > 0$, and $V_1 < 0$. Moreover, for all (θ, k) , the indifference set $\{(x, y) | V(x, y, \theta) = k\}$ is a closed convex curve*

(G1.4) *The function $g(x)$ can be written $g(x) = ax + b$, where a is nonnegative.*

Then, $x^H(\theta; g)$ is nondecreasing in θ for all functions g if and only if $V_1 / |V_2|$ is everywhere nondecreasing in θ . Similarly, $x^L(\theta; g)$ is nondecreasing in θ for all functions g if and only if $V_1 / |V_2|$ is everywhere nondecreasing in θ .

Theorem 1.4 has a structure parallel to Theorem 1.3. Notice first that we have restricted our class of functions to be the class of linear functions with nonnegative slope. The restriction (V1.4) has been modified to place sign restrictions on the derivatives of V and to ensure that the objective is quasiconcave. The sign restrictions are matched with the sign restrictions on the slope of g . What is surprising about this theorem is that, even though (G1.4) is much more restrictive than (G1.3), so that the context over which robustness is sought is much narrower, the critical sufficient condition is unchanged.²

Thus, the pairs of theorems (1.1 and 1.2, and 1.3 and 1.4) illustrate several of the themes which will arise throughout our analysis of optimization problems. By formulating the problems in terms of “contexts,” we are able to compare problems which have a similar structure but require different amounts of robustness (i.e., smaller or larger sets of functions g). In both pairs of theorems, we saw that the same condition was critical for very general contexts and for relatively narrow contexts – ones which require robustness only across families of *linear* functions. Thus, we were able to distinguish critical assumptions from other, “simplifying” assumptions in the problems: the restriction that g is linear is a simplifying assumption, while the critical sufficient conditions are stated in the corresponding theorems. In the next two chapters, we develop the general theory associated with the theorems we have introduced in this chapter.

² Applying the necessity part of Theorem 1.4 to the two good consumption problem does entail an assumption that the second period endowment γ can be either positive or negative. If one restricts γ to be positive, then some consumption patterns for the two periods may be inconsistent with any set of prices and the local changes in the slopes of the indifference curves in those regions cannot be so sharply restricted. Such problems can still be addressed using the results of the Chapter 3, in which domain restrictions on V may be invoked.

2 Single Variable Problems with Separable Contexts

In the preceding chapter, we made several compromises to keep the analysis familiar and relatively simple. First, we assumed differentiability of various functions so that implicit function methods could be applied. Second, we have avoided dealing squarely with the case of multiple optima, choosing to focus instead on the largest and smallest maximizer when multiple maximizers exist. This, too, is a compromise dictated by a desire to stay close to the implicit function approach, which is limited to studying particular optima that vary smoothly with changing parameters.

In this chapter, we will gradually eliminate these restrictions, trying at the same time to build new intuitions that corresponds closely to the new methods. At the end of this chapter, we show how the theorems developed in this chapter can be used to generalize and clarify several classic results in producer theory. In particular, we examine the LeChatlier Principle and comparative statics questions about how a firm's input demands, profits, and supply change with input prices. The formal analysis in both cases is simple and general, dispensing with extraneous assumptions such as convexity of the production technology that are required only by the implicit functions approach; further, the analysis mirrors the economic reasoning.

Our analysis of robust comparative statics for optimization distinguishes several kinds of "contexts" for optimization problems. The context treated in this chapter is of a type that arises in the classical theory of the firm, in which the objective is a sum of several terms, with some terms parameterized and others of which are not.

While we will not make much use of the formalism associated with contexts, it may be helpful to be precise about what a context is. Formally, a context for maximization problems is a triple $(f, G = \{g\}, \lambda; f \times G \rightarrow \Pi)$ where $\Pi = \{\pi | \pi: X \times \Theta \rightarrow \mathbb{R}\}$ is a set of potential objective functions for the problem and f and g are elements from which the objective is constructed, using a rule of construction λ . Thus, $\pi(x, \theta) = \lambda(f, g)(x, \theta)$. *Additively separable* contexts are ones for which $\lambda(f, g) = f + g$. These are the focus of this chapter. Note, however, that by using log transformations, results about additively separable contexts can also be applied to *multiplicatively separable* contexts, that is, ones in $\lambda(f, g) = f \cdot g$. Problem (P1) is multiplicatively separable, while (P2) is additively separable. In each context, we seek conditions on f such that for all $g \in G$, $x^*(\theta, g) = \text{Argmax}_{x \in X} \lambda(f, g)(x, \theta)$ is nondecreasing in θ .

We will also study other kinds of contexts in the rest of the monograph; for example, Chapter 3 studies problems where $f: X \times \mathbb{R} \times \Theta \rightarrow \mathbb{R}$, $g: X \rightarrow \mathbb{R}$, and $\pi(x, \theta) = f(x, g(x), \theta)$. In Part III of this monograph, we will study contexts which correspond to stochastic problems, for example, where $f: X \times \mathbb{R} \rightarrow \mathbb{R}$, $g: \Theta \times \mathbb{R} \rightarrow \mathbb{R}$, and $\pi(x, \theta) = \int f(x, s)g(\theta, s)ds$.

2.1 General Sufficient Conditions

This section treats the following maximization problem:

$$\max_{x \in S} f(x, \theta) + g(x). \quad (2.1)$$

As in Chapter 1, we let $x^H(\theta, g)$ denote the highest solution to (2.1), while $x^L(\theta, g)$ denotes the lowest solution, where these functions are defined on the domain where such solutions exist. Theorems 1.1 and 1.2 stated that, when f is smooth, $\frac{\partial^2}{\partial x \partial \theta} f(x, \theta) \geq 0$ is a critical sufficient condition for comparative statics in several contexts, for example, the context corresponding to the family of functions $G = \{g | g: \mathbb{R} \rightarrow \mathbb{R}\}$. If one thinks of f as a “benefit” function and g as a “cost” function, then Theorem 1.1 implies that increasing the marginal benefit of the choice variable leads to higher optimal values of the choice variable. If one reverses the roles, thinking of f as a cost function and g as a benefit function, then Theorem 1.1 implies that reducing marginal costs leads to an increase in the choice variable. According to Theorems 1.1 and 1.2, this is a critical sufficient condition. That is, there are classes of problems in which one cannot hope to find a weaker sufficient condition. We have based our choice of classes of problems on the minimal requirements of price theory, but these classes will prove to have much wider application than that.

Remark: Throughout the monograph, we will state our results in terms of critical sufficient conditions for solutions to be increasing in a parameter. Note, however, that if a solution x is decreasing in a parameter θ , then $-x$ is increasing in θ (and, equivalently, x is increasing in $-\theta$). Thus, it is easy to handle problems where the natural comparative statics are in terms of increases in the parameter leading to decreases in the solution, simply by reformulating the problem to replace the natural choice variable by its negative or the parameter by its negative. We also focus our results on maximization problems, but a similar simple reformulation takes care of minimization problems. The key is to note that $\max f(x, \theta) \equiv -[\min(-f(x, \theta))]$ and that the solutions of the problems “ $\max f(x, \theta)$ ” and “ $\min[-f(x, \theta)]$ ” are identical. Thus, results on maximization problems can be used to obtain parallel results for minimization problems.

The first generalization of Theorems 1.1 and 1.2 proceeds by dropping the hypothesis of differentiability. To do that, the idea of a “marginal” benefit or cost is replaced by the idea of an “incremental” benefit or cost. If f is a benefit function and $x'' > x'$, then the incremental benefit of increasing from x' to x'' is $f(x'', \theta) - f(x', \theta)$.

The principle that an increase in the incremental benefit or a reduction in the incremental cost of an activity makes the activity more desirable is quite an intuitive one. If increasing (“incrementing”) a particular activity is profitable when the incremental benefit is low, then it must be even more profitable when the incremental benefit is high, so increasing the incremental benefits cannot favor reducing the level of an activity.

The mathematical development tracks this intuition exactly and, importantly, identifies it as the *only* principle of robust comparative statics in additively separable contexts. The formalities begin with a definition.

Definition: The function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ has **increasing differences** if for all $x'' > x'$, the difference $f(x'', \theta) - f(x', \theta)$ is nondecreasing in θ .

Using this definition, we can develop a generalization of Theorem 1.1 that drops the

assumption (F1) that the function f is differentiable.

Theorem 2.1. Assume the following about the problem (2.1):

(S2.1) $S \subseteq \mathfrak{R}$.

(F2.1) No restrictions are imposed on f .

(G2.1) No restrictions are imposed on g .

Then, $x^H(\theta; g)$ is nondecreasing in θ for all functions g if and only if f has increasing differences. Similarly, $x^L(\theta; g)$ is nondecreasing in θ for all functions g if and only if f has increasing differences.

Notice the close formal comparison between Theorem 1.1 and Theorem 2.1. Conditions (S1.1) and (G1.1) are identical to (S2.1) and (G2.1). The restriction (F1.1) on f is replaced by the null restriction (F2.1). Thus, Theorem 2.1 applies over a larger domain than of Theorem 1.1. To specialize Theorem 2.1 to imply Theorem 1.1 directly, one needs to characterize when a twice continuously differentiable function f has increasing differences. For completeness, the following theorem does more: It also treats the case where f may be just once differentiable or differentiable in just one variable, as occurs, for example, in applications where the choice variable or parameter is discrete.

Theorem 2.2. Let $f(x, \theta): \mathfrak{R}^2 \rightarrow \mathfrak{R}$.

(1) If f is twice continuously differentiable, then f has increasing differences if and only if for all (x, θ) , $\frac{\partial^2}{\partial x \partial \theta} f(x, \theta) \geq 0$.

(2) If, for all θ , $f(x, \theta)$ is continuously differentiable¹ in x , then f has increasing differences if and only if for all x , $\frac{\partial}{\partial x} f(x, \theta)$ is nondecreasing in θ .

(3) If, for all x , $f(x, \theta)$ is continuously differentiable in θ , then f has increasing differences if and only if for all θ , $\frac{\partial}{\partial \theta} f(x, \theta)$ is nondecreasing in x .

Proof: We will prove (2) only; the other proofs are similar. First, suppose that $\frac{\partial}{\partial x} f(x, \theta)$ is nondecreasing in θ . Pick any $x'' > x'$. Then, $f(x'', \theta) - f(x', \theta) = \int_{x'}^{x''} \frac{\partial}{\partial x} f(x, \theta) dx$ by the fundamental theorem of calculus. But, the assumption that $\partial f / \partial x$ is nondecreasing in θ implies that $f(x'', \theta) - f(x', \theta)$ must be nondecreasing in θ . Applying the result for all $x'' > x'$ yields the desired result, that f must satisfy increasing differences.

Next, suppose that there exists an x and a $\theta'' > \theta'$ such that $\frac{\partial}{\partial x} f(x, \theta'') < \frac{\partial}{\partial x} f(x, \theta')$. Since f is continuously differentiable in x , there must be two values of x such that $x'' > x > x'$,

¹ The proof given below actually applies as well when f is differentiable only *almost everywhere*, provided that f is also absolutely continuous. This strengthened version of Theorem 2.2 is useful in dynamic programming contexts and in applications that require the Envelope Theorem, such as the producer theory application considered later in this chapter.

such that $\frac{\partial}{\partial x} f(x, \theta') < \frac{\partial}{\partial x} f(x, \theta)$ for all $x \in (x', x'')$. This in turn implies that $\int_{x'}^{x''} \frac{\partial}{\partial x} f(x, \theta'') dx < \int_{x'}^{x''} \frac{\partial}{\partial x} f(x, \theta') dx$. Applying the fundamental theorem of calculus one more time yields that $f(x'', \theta'') - f(x', \theta') < f(x'', \theta) - f(x', \theta)$, so f does not have increasing differences.

Examples:

1. The function $f(x, y) = axy$ has increasing differences on \mathbb{R}^2 when $a \geq 0$ and nonincreasing differences when $a \leq 0$, because $f_{xy} \equiv a$.
2. The function $f(x, y) = ax^\alpha y^\beta$ has increasing differences on the positive quadrant $\{(x, y) \mid x > 0, y > 0\}$ if $a\alpha\beta \geq 0$, because the mixed partial derivative is $f_{xy}(x, y) = a\alpha\beta x^{\alpha-1} y^{\beta-1}$. The function has nonincreasing differences on the same set when $a\alpha\beta \leq 0$.
3. For any convex function g , the function $f(x, y) = g(x+y)$ has increasing differences. Similarly, if g is concave, then $f(x, y) = g(x-y)$ has increasing differences. These results are easy to check using Theorem 2.2 when g is twice differentiable but are also true when g is not smooth.
4. For any increasing function f with increasing differences, and for any increasing, convex function g , the composite function $g(f(x, y))$ has increasing differences. Once again, when f and g are twice continuously differentiable, this result is easy to verify by applying Theorem 2.2.

2.2 Multiple Optima and Global Comparative Statics

The next step in the development of the theory is to deal more completely with the possibility that there may be multiple local or global optima. Situations with multiple optima are particularly problematic for the implicit function approach, which traces the change in a continuous selection of a local optimum. The difficulty is that it is possible that a parameter increase causes every local optimum to increase and yet leads to a decrease in the global optimum. Figure 2.1 illustrates how that can happen.

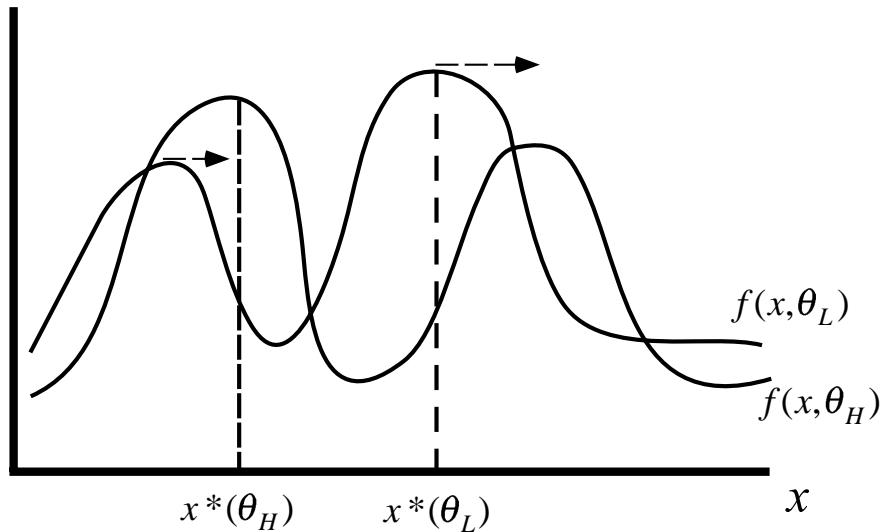


Figure 2 .1: Each local optimum is increasing in θ , yet the global optimum decreases in θ .

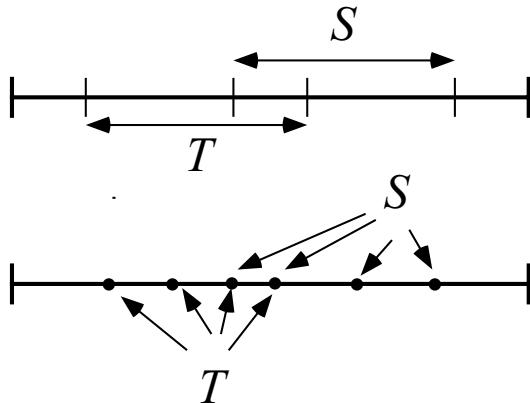
Henceforth, we shall let $X^*(\theta, g)$ denote the set of maximizers in problem (2.1), sometimes dropping the explicit reference to the function g when no confusion is likely to result.. While it is obvious what it means for a unique optimum or for the highest or lowest maximizer – $x^H(\theta, g)$ or $x^L(\theta, g)$ – to shift upward, it is less obvious what it must mean for the set $X^*(\theta, g)$ to move up as θ increases. Yet this is the question we must answer to make a comparative statics assertion for sets of maximizers. The notions we shall use to resolve this problem are given by the following definitions, which utilize what is sometimes called the “strong set order” (SSO).

Definition: A set $S \subseteq \mathbb{R}$ is **as high as** another set $T \subseteq \mathbb{R}$, written $S \geq_S T$, if for every $x \in S$ and $y \in T$, $y \geq x$ implies both $x \in S \cap T$ and $y \in S \cap T$. S is **higher than** T (written $S >_S T$) if S is as high as T but T is not as high as S .

Definition: A set-valued function $H: \mathbb{R} \rightarrow 2^{\mathbb{R}}$ is **nondecreasing** if for $x > y$, $H(x) \geq_S H(y)$.

An alternative, equivalent definition of the “as high as” relation goes as follows. The set S is as high as the set T (in the Strong Set Order) if and only if (i) each $x \in S \setminus T$ is greater than each $y \in T$ and (ii) each $x' \in T \setminus S$ is less than each $y' \in S$.

Sets satisfy $S \geq_S T$



Sets fail $S \geq_S T$

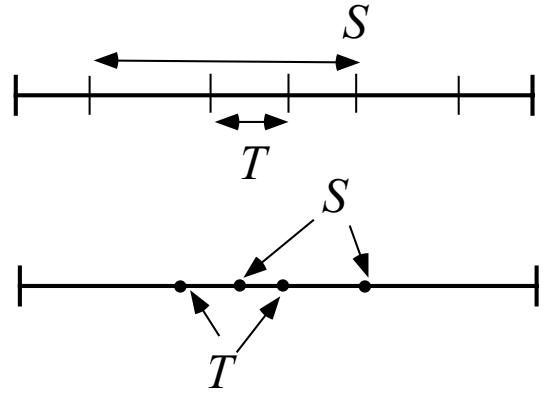


Figure 2 .2: Illustrations of the Strong Set Order

Figure 2 .2 gives some examples of some pairs of sets that satisfy the condition of the “higher than” definition and other pairs that fail the condition. The condition is quite demanding.

Notice that when the sets A and B are singletons, say $A = \{x\}$ and $B = \{y\}$, then $A \geq_S B$ corresponds to $x \geq y$. So, the strong set order can be regarded as an extension of the usual order from points to sets. However, this particular extension is relatively demanding; it is much stronger than merely asserting that the points in A tend “on average” to be higher than the points in B . Indeed, if A and B are closed sets of real numbers, then $A \geq_S B$ if and only if there exist numbers a and b in $\mathbb{R} \cup \{-\infty, +\infty\}$ such that $A = (A \cup B) \cap \{x \mid x \geq a\}$ and $B = (A \cup B) \cap$

$\{x \mid x \leq b\}$. That is, A is higher than B if A consists of all the points in the “top part” of $A \cup B$ and B consists of all the points in the “bottom part.” Either by using this rule or working directly from the definition, one can verify that $\{2\} >_S \{1,2\} >_S \{1\}$ and $[2,4] >_S [1,3]$, but $\{2,4\}$ is not higher than $\{1,3\}$. Yet, increasing differences is sufficient to imply a comparative statics theorem using even the demanding notion of monotonicity associated with this set order.

The proof of the foregoing characterization of the strong set order is left as an exercise below. A similar characterization holds when A or B is not closed, but then the inequality $x \geq a$ may need to be replaced by $x > a$ and $x \leq b$ may need to be replaced by $x < b$. It is also worth noting explicitly that $B \subset A$ does not imply either $A \geq_S B$ or $B \geq_S A$. For example, although $[2,3] \subset [1,4]$, the two sets are unordered in the strong set order: $[1,4]$ is not greater than $[2,3]$, and $[2,3]$ is not lower than $[1,4]$.

When restricted to non-empty sets *on the real line* (we will see in Part II that things get more complicated in \Re^n), the “higher than” relation is transitive ($A \geq_S B$ and $B \geq_S C$ implies $A \geq_S C$), reflexive ($A \geq_S A$), and antisymmetric ($A \geq_S B$ and $B \geq_S A$ implies $A = B$), but not complete. As we have just seen, $[1,4]$ and $[2,3]$ are unordered; neither is higher than the other. Similarly, the sets of odd integers and even integers are unordered according to the strong set order.

There is one peculiarity of the definition that requires special mention. It is that, according to the definition, for all sets $S \subset \Re$, $\emptyset \geq_S S \geq_S \emptyset$: the empty set is both higher and lower than any other set. Thus, if the solution $X^*(\theta)$ of the model exists (that is, $X^*(\theta) \neq \emptyset$) for some value θ' of the parameter and fails to exist ($X^*(\theta'') = \emptyset$) for some other value $\theta'' > \theta'$, then the solution is both increasing and decreasing in θ . We take advantage of this property in the statements of our theorems, omitting language that would otherwise be required restricting comparisons to situations in which the set of optimizers is non-empty.

Having defined what it means for a set to be an increasing function of a parameter, we can now extend our theory to include comparative statics on sets of optimizers.

Theorem 2.3. Assume the following about the problem (2.1):

(S2.3) $S \subseteq \Re$.

(F2.3) No restrictions are imposed on f .

(G2.3) No restrictions are imposed on g .

Then, $X^*(\theta; g)$ is nondecreasing in θ for all functions g if and only if f has increasing differences.

Proof: First, we show that if f has increasing differences, then $X^*(\theta; g)$ is nondecreasing in θ for all functions g . Pick a g , and a $\theta' > \theta'$. Pick $x' \in X^*(\theta'; g)$ and $x'' \in X^*(\theta''; g)$. Suppose further that $x' > x''$. Now, $x' \in X^*(\theta'; g)$ implies that $f(x', \theta') + g(x') \geq f(x'', \theta') + g(x'')$. Recall that increasing differences implies that

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

and thus

$$f(x', \theta'') + g(x') - f(x'', \theta'') - g(x'') \geq f(x', \theta') + g(x') - f(x'', \theta'') - g(x'').$$

These two facts together imply that $f(x', \theta'') + g(x') \geq f(x'', \theta'') + g(x'')$. Now, since $x'' \in X^*(\theta''; g)$, then the reverse inequality must hold as well, which implies that $f(x', \theta') + g(x') = f(x'', \theta'') + g(x'')$ and hence that $x' \in X^*(\theta''; g)$. A symmetric argument shows that $x'' \in X^*(\theta'; g)$. This completes the proof that $X^*(\theta''; g) \geq_S X^*(\theta'; g)$.

Now, suppose that f fails to satisfy increasing differences. That is, there exists a $x'' > x'$ and a $\theta' > \theta'$ such that $f(x'', \theta') - f(x', \theta') > f(x'', \theta'') - f(x', \theta'')$. Then, define a function $g(x)$ by specifying its values at x', x'' , and elsewhere, as follows:

$$\begin{aligned} g(x') &= f(x'', \theta') - f(x', \theta') \\ g(x'') &= 0 \\ g(x) &= \min\{f(x'', \theta') - f(x, \theta'), f(x'', \theta'') - f(x, \theta'')\} - 1, \text{ for } x \in S \setminus \{x', x''\}. \end{aligned}$$

With this specification, $f(x'', \theta') + g(x'') = f(x', \theta') + g(x')$, and further $f(x'', \theta') + g(x'') = f(x'', \theta') > f(x, \theta') + g(x)$ for all $x \in S \setminus \{x', x''\}$. This implies that $X^*(\theta'; g) = \{x', x''\}$.

Now, consider θ' . Substituting in for g and applying our hypothesis about f , we have that

$$\begin{aligned} f(x', \theta') + g(x') &= f(x', \theta') + f(x'', \theta') - f(x', \theta') \\ &> f(x'', \theta'') \\ &= f(x'', \theta'') + g(x'') \end{aligned}$$

Further, for $x \in S \setminus \{x', x''\}$, $f(x, \theta') + g(x) < f(x'', \theta'')$. Thus, $X^*(\theta'; g) = \{x'\}$, which is lower than $X^*(\theta'; g) = \{x', x''\}$ (because, by hypothesis, $x'' > x'$). This provides the required contradiction.

It is useful to compare Theorems 2.1 and 2.3. The conditions (S2.3), (F2.3), and (G2.3) are identical to the corresponding restrictions of Theorem 2.1. If we are interested in exploring the consequences of the assumption that f has increasing differences, then the monotonicity conclusion of Theorem 2.3 is stronger. While Theorem 2.1 concludes only that the highest and lowest maximizers move up, Theorem 2.3 concludes that the whole set moves up in a more restrictive sense.

However, Theorem 2.3 does not imply Theorem 2.1. Because both theorems provide necessary and sufficient conditions for a monotonicity conclusion, the stronger monotonicity conclusion of Theorem 2.3 makes that theorem sharper in one direction (sufficiency) and weaker in the other (necessity). Theorem 2.1 establishes that increasing differences are necessary to imply either one of the weaker monotonicity conclusions that $x^H(\theta, g)$ or $x^L(\theta, g)$ is nondecreasing in θ for all g . The proof of necessity in Theorem 2.1, however, employs the same counterexample that was used in the corresponding proof for Theorem 2.3, and thus restricting the

conclusion of Theorem 2.3 will not change anything.

We will pay more attention later to the necessary conditions for various comparative statics conclusions. Focusing on increasing differences as a sufficient condition, Theorem 2.3 is our most powerful result. We will now argue that the theorem's conclusion is stronger than anything that can be derived from the implicit function approach, even when that approach applies. We make this comparison first for global maximizers and then for local maximizers.

For global maximizers, Theorem 2.3 tells us that the assumption of increasing differences rules out examples like that illustrated in Figure 2 .1. The implicit function theorem applies only when each maximizer is locally unique, so let us assume that is the case. Notice that in Figure 2 .1, increasing the parameter adds a new, lower global maximum. But, according to the theorem, this cannot happen if f has increasing differences.

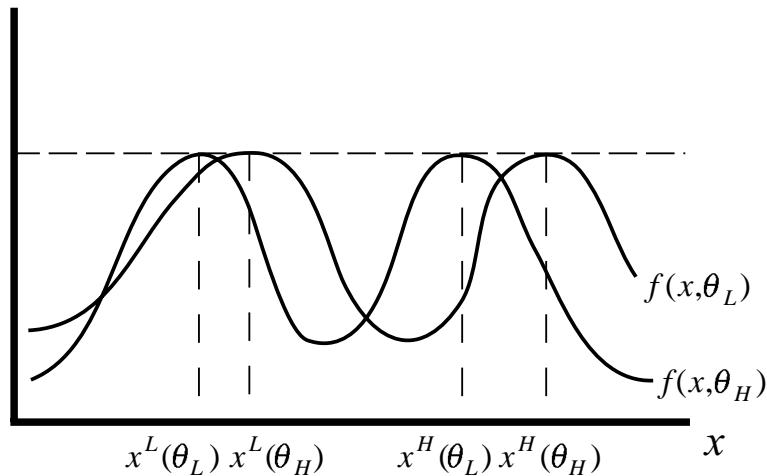


Figure 2 .3: Two optima which both increase in θ . We will show that this f cannot satisfy increasing differences.

Another possibility which is ruled out by the theorem is illustrated in Figure 2 .3. In this illustration, there are two global maxima, and both shift up when the parameter value increases. However, this is a violation of an increase in the strong set order, as discussed above. To see why this cannot happen, consider the case where f is differentiable in x . Figure 2 .4 graphs f_x as a function of x . By the definition of increasing differences, moving from θ_L to θ_H increases this curve. If we are indifferent between $x^L(\theta_L)$ and $x^H(\theta_L)$, then the area marked A must be equal to the area marked B (this follows since $f(x^L(\theta_L), \theta_L) = f(x^H(\theta_L), \theta_L)$ if and only if $\int_{x^L(\theta_L)}^{x^H(\theta_L)} f_x(x, \theta_L) dx = 0$, by the fundamental theorem of calculus). But, increasing θ to θ_H must decrease A and increase B, so that $x^H(\theta_H)$ is strictly preferred to $x^L(\theta_H)$.

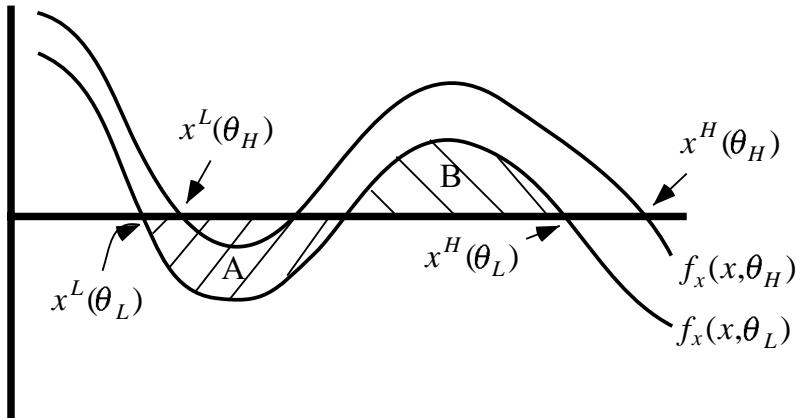


Figure 2.4: Increasing differences rules out the case of two optima which both increase in θ .

In general, monotonicity of the set of maximizers implies that if a small increase in the parameter changes some local maximizer other than $x^H(\theta)$, then that local maximizer and all smaller ones cannot be global maximizers for the higher parameter value. Thus, it is not possible that $x^L(\theta)$ increases continuously with θ over any range of the parameter for which the number of maximizers is finite and at least two, for then the set of maximizers would not increase in θ according to our definition of the strong set order. Similar conclusions about elements entering and leaving the set of global maximizers are not derivable using implicit function methods.

We now compare the two methods for the problem of deriving comparative statics on local optima. We know that when implicit function methods apply, they give the best possible results about how local optimizers change with exogenous parameters. However, one does not need to rely on implicit function methods for that. The same conclusions can also be derived using Theorem 2.3 and under weaker assumptions than those required by the implicit function approach. Intuitively, our argument is that *local maximizers are simply global maximizers on a restricted domain*, and the critical sufficient conditions for monotonicity of global maximizers, when applied correctly to a narrow domain, are the critical sufficient conditions for monotonicity of local optimizers.

More formally, suppose that for a particular parameter value θ' , there is a local maximizer at $x(\theta')$. For the implicit function approach to apply, several things must be true: There must be some interval S' around $x(\theta')$ in which there are no other local maximizers. The same conclusion must be true using the same interval S' , for some interval Θ of parameter values around θ' . And, the local optimum function $x(\theta)$ and the objective function $f(x, \theta) + g(x)$ must be twice continuously differentiable in a neighborhood of $(x(\theta'), \theta')$. When these conditions are satisfied, the implicit function approach provides a way to calculate the derivative $\frac{\partial}{\partial \theta} x(\theta')$.

Given these assumptions, however, the “local maximizer” is precisely the unique solution to an optimization problem with choice set S' . The preceding theorems apply to this problem, giving sufficient conditions for the local maximizer to be nondecreasing in the parameter throughout a

neighborhood of θ' . In the case of smooth functions f and g , the condition is that $f_{x\theta}$ be non-negative in the neighborhood of θ' , which is the same as the conclusion from the implicit function approach. However, according to the theorems, the monotonicity of the local maximizer can still be correctly inferred even when, for example, g is not differentiable or the derivative $\frac{\partial}{\partial \theta} x(\theta')$ is undefined at some points.

We close this section with a theorem to show how the notion that one set is “higher than” another can be useful as a sufficient condition in a theorem, and not merely as part of the comparative static conclusion. Examples of economic problems where constraint sets vary include pricing problems with price floors or price ceilings, or quantity problems with capacity constraints.

Theorem 2.4. Let $X^*(\theta, S)$ be the set of maximizers of (1) and let us regard the constraint set S as varying. If $S \geq_S S'$, then for all θ , $X^*(\theta, S) \geq_S X^*(\theta, S')$.

Proof: Choose $x \in X^*(\theta, S)$ and $x' \in X^*(\theta, S')$. Suppose that $x > x'$. Since $S' > S$, by the definition of the strong set order, $x' \in S$ and $x \in S'$. Since both are available in both choice problems, the choices must yield equal returns, and thus $x, x' \in X^*(\theta, S) \cap X^*(\theta, S')$.

Thus, increasing the constraint set in the strong set order must (weakly) increase the set of optimizers, since increasing the constraint set takes away low options and makes new, higher options available. To see an example, suppose that an agent is choosing x to maximize $\sin(x)$, and let $S = [0, 4\pi]$ and $S' = [\pi, 5\pi]$. In this case, the set of optimizers $X^*(S) = \{\frac{\pi}{2}, \frac{5\pi}{2}\}$, and $X^*(S') = \{\frac{5\pi}{2}, \frac{9\pi}{2}\}$. Clearly, the second set is higher in the strong set order.

2.3 Critical Sufficient Conditions: The Choice of Context

Our main emphasis in Sections 2.1 and 2.2 was on finding sufficient conditions for monotone comparative statics in problem (1) which are robust across certain families of problems. Our analysis culminated in Theorem 2.3, which establishes that increasing differences is a sufficient condition for a strong form of monotonicity regardless of the specification of g . Moreover, the theorem establishes that increasing differences is the weakest such condition in the specified context – what we have called a *critical sufficient condition*. However, the degree of robustness we sought was quite extensive. Consequently, it is natural to ask whether a weaker condition than increasing differences might be sufficient if we were somehow less demanding. Is increasing differences still a *critical* sufficient condition in plausibly narrow contexts? The main task of this section is to identify precisely the contexts for which increasing differences is a necessary condition for robust comparative statics.

In our analysis of the bidding problem (P1), we discussed why increasing differences might be critical for monotone comparative statics in that problem. There, we used the implicit function approach to argue that if the set of possible distribution functions for competing bids included the exponential family, then the $f(x, \theta) = \log(U(x, \theta))$ must satisfy the increasing differences condition that, for all (x, θ) , $f_{x\theta}(x, \theta) \geq 0$ everywhere. The idea was that the exponential family is large enough that for any value of the parameter θ , the optimal bid x might be anything in the domain of

$\log(U(x, \theta))$. Then, the implicit function approach requires that the mixed partial derivative must be non-negative everywhere, which by Theorem 2.2 is equivalent to the increasing differences condition.

In our study of the bidding example, we acknowledged that the particular choice of the exponential family was arbitrary and suggested that some – and perhaps many – other families might have worked equally well, provided the set of functions g that are included in the family is sufficiently large. How large must the family be? The next theorem provides the answer for smooth, concave maximization problems. First, however, we shall need a definition which applies for families of functions whose domain is a convex set.

Definition: The one-parameter family of functions $\{g_t(x)\}$ mapping a convex set $S \subseteq \mathbb{R}$ into \mathbb{R} is a full one parameter family if for all $x \in S$ and all $a \in \mathbb{R}$ there exists a $t \in \mathbb{R}$ such that $g_t'(x) = a$. It is a positive (negative) one parameter family if for all $x \in S$ and all $a > 0$ ($a < 0$) there exists $t \in \mathbb{R}$ such that $g_t'(x) = a$.

Thus, a family of functions is a full one parameter family if, for every value of x and every constant a , there is some member of the family whose first derivative at x is a . An example of such a family is $\{g_t(x) = t \cdot x, t \in \mathbb{R}\}$. The family is a positive one parameter family if its derivative at each argument can be any positive number. In the log-transformed bidding example, the family of functions we used was the logarithm of the exponential distribution functions. Although we did not use this language, our argument in the example depended only on the idea that this family is a positive one parameter family. In the normal inputs problem (P2), the relevant family of functions was the set of cost functions of the form $-rx$ which, if we restrict attention to positive input prices, forms a negative one parameter family.

The next theorem generalizes the arguments made in our examples and reveals how narrow the context must be before increasing differences can fail and yet the desired comparative statics conclusion can hold “robustly.”

Theorem 2.5. Assume the following about the problem (2.1):

(S2.5) $S \subseteq \mathbb{R}$ is convex and compact.

(F2.5) The function $f(x, \theta)$ is twice continuously differentiable in (x, θ) and strictly concave in x .

(G2.5) The function $g_t(x)$ is continuous in t and differentiable and concave in x .

(a) If $\{g_t(x)\}$ is a full one parameter family, then the following conclusion holds:

$x^H(\theta; f, g_t)$ is nondecreasing in θ for all t , **if and only iff** has increasing differences.

(b) If $\{g_t(x)\}$ is **not** a full one parameter family, then there exists an f satisfying (F2.5) such that the following holds:

$x^H(\theta; f, g_t)$ is nondecreasing in θ for all t , but f does not have increasing differences.

On one hand, part (a) of Theorem 2.5 implies that increasing differences is critical for robust comparative statics in the context where the objective is smooth and concave, and when the family of functions $\{g(x)\}$ is sufficiently rich. “Sufficiently rich” means that the family $\{g(x)\}$ is a full one parameter family. This means that the conditions required for robust comparative statics in these classes of concave problems are no different than those for general problems. To those who were taught that comparative statics analysis is different and perhaps easier for smooth, concave maximization problems because the second order conditions are satisfied globally, this conclusion of Theorem 2.5 may come as a surprise.

However, part (b) of Theorem 2.5 states the converse to part (a): it asserts that if the set of relevant g functions in problem (1) do not include a one parameter family, then the condition of increasing differences is *not* critical. Weaker sufficient conditions are possible.

What is the nature of these weaker sufficient conditions? A simple example to illustrate them involves specifying the family $\{g(x)\}$ to consist just of the zero function. In that case, the implicit function approach implies that a necessary and sufficient condition for monotone comparative statics is that $f_{x\theta}(x^H(\theta), \theta) \geq 0$ for all θ . Increasing differences needs to hold only along the curve in (x, θ) -space where the maximum may lie. In general, if the possible (x, θ) pairs that can arise at an optimum excludes a neighborhood of some point (x', θ') , then one cannot generally conclude that $f_{x\theta}(x', \theta') \geq 0$. One point of the theorem is that if the family $\{g(x)\}$ includes some g for which $f_x(x', \theta') + g'(x') = 0$, then such a case cannot arise.

Since we are interested in qualitative comparative statics, for most of what follows we shall assume that every point that the modeler includes in the constraint set S can actually be an optimum for some specification of g . Then, in problems of the form (1), increasing differences is indeed a critical sufficient condition.

Another consequence of Theorem 2.5 is that checking that a comparative statics conclusion holds for all linear cost functions is equivalent to checking that the result holds for arbitrary cost functions, because the linear functions are a full one parameter family. Thus, existing comparative statics results which make restrictive assumptions, such as perfectly competitive input markets, automatically generalize to the case of more general cost functions, so long as the cost functions satisfy increasing differences (this is further developed in Section 2.4).

We can further extend Theorem 2.5 to problems where some monotonicity is assumed, for example, if f is a benefit function and g is a cost function, or vice versa.

Theorem 2.5'. Suppose that

(S2.5) $S \subseteq \mathfrak{R}$ is convex and compact.

(F2.5') The function $f(x, \theta)$ is twice continuously differentiable in (x, θ) and strictly concave in x , and further f is decreasing (respectively, increasing).

(G2.5) The function $g_i(x)$ is continuous in t and differentiable and concave in x .

(a') If $\{g_i(x)\}$ is a positive (respectively, negative) one parameter family, then the following conclusion holds:

$x^H(\theta; f, g_t)$ is nondecreasing in θ for all t , if and only if f has increasing differences.

(b') If, instead, $\{g_t(x)\}$ is not a positive (respectively, negative) one parameter family, then there exists an f satisfying (F2.5') such that the following holds:

$x^H(\theta; f, g_t)$ is nondecreasing in θ for all t , but f does not have increasing differences.

Theorem 2.5' shows that if the slope of f is restricted to be positive, the relevant one-parameter family $\{g_t(x)\}$ need only include all functions with negative slope. We do not need to ask for our result to be robust to linear cost functions with negative prices, if we know that f is a benefit function.

2.4 Strict Comparative Statics

In some problems, it can be useful to know when a selection from the set of optimizers is strictly increasing in an exogenous parameter. This problem has been analyzed by Edlin and Shannon (1996a, 1996b). A first intuition might suggest that a sufficient condition for a selection to be strictly increasing would be strictly increasing differences, defined as follows:

Definition: The function f has strictly increasing differences if for all $x'' > x'$, $f(x'', \theta) - f(x', \theta)$ is strictly increasing in θ .

Strictly increasing differences allows us to strengthen our comparative statics conclusions: if $f(x, \theta)$ satisfies strictly increasing differences, then for $\theta_H > \theta_L$, if $x' \in x^*(\theta_L)$ and $x'' \in x^*(\theta_H)$, then $x'' \geq x'$ (the proof of this is left as an exercise). This conclusion can be useful for proving that an agent's optimal choice will never decrease in θ . However, Figure 2.6 illustrates why strictly increasing differences is not sufficient for strict comparative statics, which would entail $x'' > x'$ in the latter example.

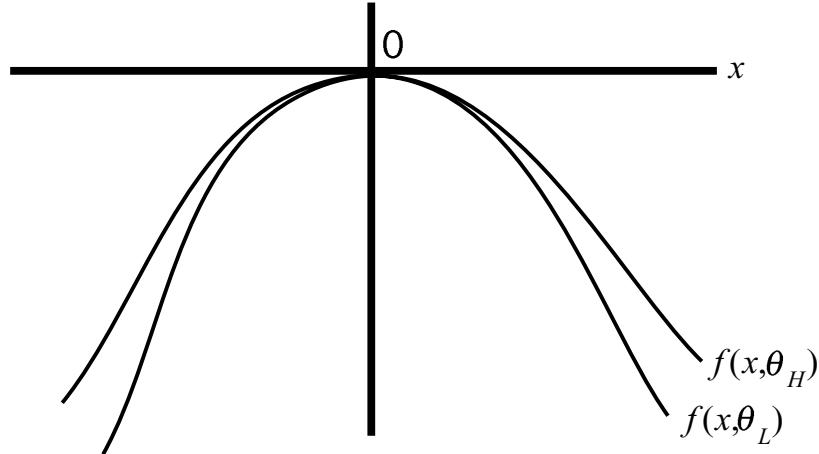


Figure 2.6: The function f satisfies strictly increasing differences, yet the optimal choice of x is constant in θ .

Figure 2.6 plots the function $f(x, \theta) = -6x^2 + \theta x^3$ for $x \in [-1, 1]$ and for two distinct values of the parameter θ satisfying $0 < \theta \leq 1$. Notice that the maximum remains fixed at $x^* = 0$, even though

$f_{x\theta}=3\theta x^2 \geq 0$ with equality only at $x=0$. This function has strictly increasing differences, but the maximum $x^*(\theta)$ does not increase with θ . To conclude that the maximizers are strictly increasing, stronger assumptions are required.

It might seem surprising that there is no strict analog of Theorem 2.2, which shows that increasing differences of f is equivalent to the conclusion that $\frac{\partial}{\partial x} f(x, \theta)$ is nondecreasing. However, it turns out that even if a differentiable function $f(x, \theta)$ has strictly increasing differences, then $\frac{\partial}{\partial x} f(x, \theta)$ may still be constant in θ throughout most of its domain.² This is not a problem if we are only interested in a weak conclusion, but it does create difficulties for strict results. In the above example, $\frac{\partial}{\partial x} f(0, \theta)$ is a constant, independent of θ at the optimum, where $x=0$.

For the special case where f is differentiable in x , we have the following result:

Theorem 2.6. Assume the following about the problem (2.1):

(S2.6) $S \subseteq \mathbb{R}$.

(F2.6) $f(x, \theta)$ is continuously differentiable in x .

(G2.6) g is continuously differentiable in x .

(I2.6) The highest solution $x^H(\cdot; g)$ is interior.

Then, $x^H(\theta; g)$ is strictly increasing in θ for all functions g if and only if $\frac{\partial}{\partial x} f(x, \theta)$ is strictly increasing in θ .

A final observation about strict monotonicity involves comparative statics on sets of optimizers. Recall that, if f has increasing differences, then the set of optimizers increases in θ in the strong set order. Fix g to be continuously differentiable. Suppose that $\frac{\partial}{\partial x} f(x, \theta)$ is strictly increasing in θ and that there exist interior maximizers for all values of the parameters. Let $\theta > \theta'$. The sets $X^*(\theta)$ and $X^*(\theta')$ can have no elements in common in the interior of S (because no x cannot satisfy the first order condition for two distinct values of θ). Also, it follows from Theorem 2.3 that $X^*(\theta) \geq_S X^*(\theta')$. It follows from these facts and the assumption that both sets have interior elements that every element of $X^*(\theta)$ is greater than every element of $X^*(\theta')$.

2.5 A Generalized Envelope Theorem

Our emphasis in this monograph is on theorems about comparative statics, and further our comparative statics theorems apply to large as well as small changes in exogenous parameters. However, a number of classic results in economics are derived using the “envelope theorem,” which studies the question of how the maximized value of an objective function (i.e. a firm’s profits evaluated at its profit-maximizing input choices) changes with an exogenous parameter

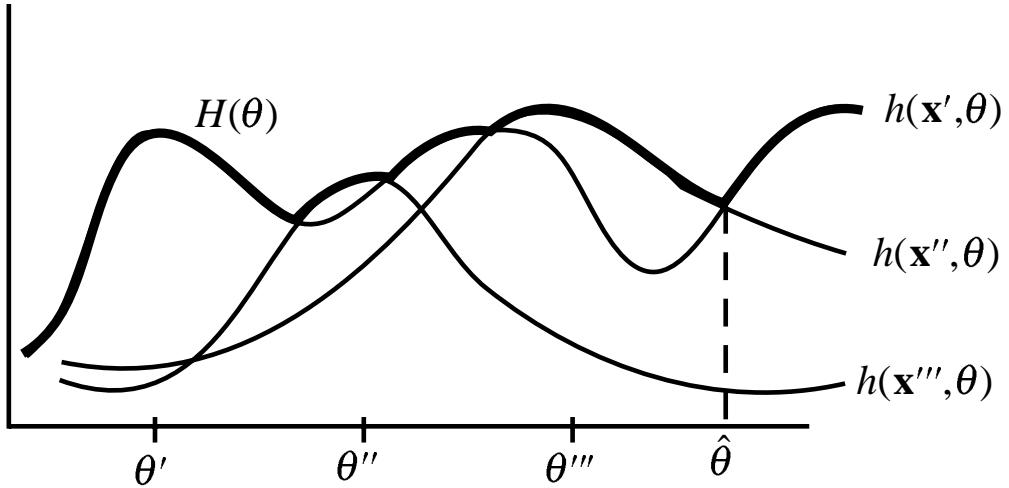
² See Edlin and Shannon (1996b) for more discussion of the relationship between the discrete and differentiable conditions.

(such as an input price). Since we are interested in problems where many of the standard assumptions have been relaxed, it will be useful to be able to make reference to a version of the envelope theorem which applies to the kinds of problems we allow for in this monograph (problems with possibly non-divisible choices, non-differentiable objective functions, and potentially multiple optima).

There are two standard explanations of the envelope theorem, one algebraic and one graphical.³ The algebraic analysis is based on differentiability of the objective function in the choice variable, while the graphical analysis is clearly more general. In both cases, we will state our result for the case where the choice variable \mathbf{x} is a real vector.

The algebraic approach proceeds as follows. Assume that the objective function $h(\mathbf{x}, \theta)$ is strictly quasi-concave and differentiable in all arguments. Then the optimum is determined by the solution to the system of first order conditions, $\nabla_{\mathbf{x}} h(\mathbf{x}^*, \theta) = \mathbf{0}$. The value function is then $H(\theta) \equiv h(\mathbf{x}^*(\theta), \theta)$. It follows that $H'(\theta) = \frac{\partial}{\partial \theta} h(\mathbf{x}^*(\theta), \theta) = \nabla_{\mathbf{x}} h(\mathbf{x}, \theta)|_{\mathbf{x}=\mathbf{x}^*(\theta)} \cdot \frac{\partial}{\partial \theta} \mathbf{x}^*(\theta) + \frac{\partial}{\partial \theta} h(\mathbf{x}, \theta)|_{\mathbf{x}=\mathbf{x}^*(\theta)}$. Since $\nabla_{\mathbf{x}} h(\mathbf{x}, \theta)|_{\mathbf{x}=\mathbf{x}^*(\theta)} = \mathbf{0}$, we conclude $H'(\theta) = \frac{\partial}{\partial \theta} h(\mathbf{x}^*(\theta), \theta) = \frac{\partial}{\partial \theta} h(\mathbf{x}, \theta)|_{\mathbf{x}=\mathbf{x}^*(\theta)}$.

While the algebraic approach emphasizes the role of differentiability with respect to \mathbf{x} , the graphical analysis makes it more clear that only differentiability of h in θ is important. Figure 2.8 plots $h(\mathbf{x}, \theta)$ as a function of θ for different values of \mathbf{x} . For the moment, let us maintain the assumption that the optimal choice of \mathbf{x} at each θ is unique. Choose three values of θ , θ' , θ'' , and θ''' , and denote the optimal choices of \mathbf{x} at these parameter values by \mathbf{x}' , \mathbf{x}'' , and \mathbf{x}''' , respectively. Figure 2.8 illustrates the function h for each of these choices of \mathbf{x} ; notice that the highest of the three functions at $\theta=\theta'$ is $h(\mathbf{x}', \theta)$, and likewise for the other choices. The figure further illustrates the value function $H(\theta) \equiv h(\mathbf{x}^*(\theta), \theta)$. Notice that $H(\theta') = h(\mathbf{x}', \theta')$. Thus, $H(\theta)$ is the “upper envelope” of all functions $h(\mathbf{x}, \theta)$, since $H(\theta)$ is by definition the best we can do at any θ .



³ For a general textbook treatment, see Mas-Colell, Whinston, and Green, 1995, pp. 964-965.

Figure 2 .8: The value function $H(\theta)$ is the upper envelope of the functions $h(\mathbf{x}, \theta)$.

Now, we know that $H(\theta) \geq h(\mathbf{x}', \theta)$ for all θ (as pictured), and further $H(\theta) = h(\mathbf{x}', \theta)$. Together, these facts imply that if H is differentiable at θ' , then there must be a neighborhood of θ' (call it $N(\theta')$) such that $\frac{\partial}{\partial \theta} H(\theta) \leq \frac{\partial}{\partial \theta} h(\mathbf{x}', \theta)$ for all $\theta < \theta'$ such that $\theta \in N(\theta')$, and $\frac{\partial}{\partial \theta} H(\theta) \geq \frac{\partial}{\partial \theta} h(\mathbf{x}', \theta)$ for all $\theta > \theta'$ such that $\theta \in N(\theta')$. Otherwise, $h(\mathbf{x}', \theta)$ would cross $H(\theta)$, which is not permitted. But, we can then conclude that $\frac{\partial}{\partial \theta} H(\theta) = \frac{\partial}{\partial \theta} h(\mathbf{x}', \theta)$: the value function must be tangent to $h(\mathbf{x}', \theta)$ at θ' .

However, we have continued to make use of the assumption that optima is unique and that derivative exists. Notice that there is a close relationship between these assumptions. Consider a problem where $x \in \mathbb{R}$ and $h(x, \theta) = g(x) + \theta \cdot x$. Suppose that there are two optimal choices of x at $\theta = \hat{\theta}$, so that $x^*(\hat{\theta}) = \{2, 4\}$. If the envelope theorem applies here and the derivative of the value function exists, we would have $H'(\hat{\theta}) = \frac{\partial}{\partial \theta} h(x^*, \hat{\theta}) = x^*$. But, we have two possible values for x^* , and we will get a different answer depending on which one we choose! Notice that increasing θ slightly will eliminate the lower optimum ($x=2$), while decreasing θ slightly will eliminate the higher optimum ($x=4$). A scenario with multiple optima is depicted in Figure 2.6. We see that the slope of the value function changes discretely at $\hat{\theta}$, so the right-hand derivative is different from the left-hand derivative. Thus, the derivative of the value function does not exist at $\hat{\theta}$. More generally, if there are multiple optimal at $\theta = \hat{\theta}$, then $H'(\hat{\theta})$ exists only if $\frac{\partial}{\partial \theta} h(\mathbf{x}, \hat{\theta})$ does not vary with \mathbf{x} for $\mathbf{x} \in \mathbf{x}^*(\hat{\theta})$.

Thus, a general version of the envelope theorem must be qualified to allow for points of non-differentiability of the value function. The following theorem does more: it states that when the objective function is continuous in \mathbf{x} , the value function is differentiable almost everywhere, and further, the standard envelope result holds whenever the derivative exists.

Theorem 2.7 (Envelope Theorem). Suppose that $S \subseteq \mathbb{R}^N$ is a compact set, that $h(\mathbf{x}, \theta)$ is differentiable in θ and continuous in \mathbf{x} , and that the derivative h_θ is bounded. Let $H(\theta) = \max_{\mathbf{x} \in S} h(\mathbf{x}, \theta)$. Then H is absolutely continuous and differentiable almost everywhere, and for all θ such that $H'(\theta)$ exists, $H'(\theta) = h_\theta(\mathbf{x}, \theta)$ for $\mathbf{x} \in \mathbf{x}^*(\theta)$.

This powerful result can be used to calculate the derivative of the maximum profit function with respect to parameters for general concave and non-concave programming problems.

2.6 Applications in Producer Theory

This section shows how the analysis of this chapter can be applied to classic questions in producer theory. Part II of this monograph will provide additional analysis of the multi-input decision problem for a firm. However, the special case of two inputs can be studied using the methods of this chapter, and our analysis illustrates several ways in which two-variable problems can be reformulated when we are interested in asking questions about how only one of the inputs changes in response to changing prices.

2.6.1 Changes in Input Prices

This section studies three classic comparative statics questions in producer theory: When input prices change, what happens to (1) the firm’s maximum profits, (2) the firm’s input demand, and (3) the firm’s output? We show that all three of these questions can be answered quite generally using the methods from this chapter. Notice that the firm’s maximum profits and output may not be choice variables in standard formulations of the problem; further, we are interested in two choice variables instead of one. We will show how to rephrase the problems so that the techniques of this chapter can be used to analyze the questions. However, for a more general treatment of multivariate problems, see Chapter 4.

Consider the classical model of the profit-maximizing price-taking firm with capital and labor as inputs. Letting r be the rental rate of capital and w the wage rate of labor, the firm chooses the input levels k for capital and l for labor to solve the following problem (which is essentially (P2), except that we have changed the notation slightly and we allow the firm to choose both capital and labor):

$$\max_{k,l} \pi(k, l; r, w) \equiv pF(k, l) - rk - wl \quad (\text{P4})$$

First consider question (1), which asks how profits are affected by changing prices. To approach this problem, we will make use of the general “envelope theorem” analyzed in Section 2.5. Fix the price r , and define the following value function:

$$V(w) \equiv \max_{k,l} \pi(k, l; r, w).$$

Since π is differentiable in the input prices, Theorem 2.7 implies that $V'(w)$ exists almost everywhere, and where it exists $V'(w) = -l^*(r, w)$. That is, when there is a small change in the wage, the firm’s profits decrease by exactly the amount of labor which is consumed at the optimum. If the optimum is not unique, the function V still generally has left-hand and right-hand derivatives at each point, but these are generally unequal. For example, in problem (P4) with r fixed, if two different values of labor, $l^L < l^H$ are optimal, then one can show that an increase in the wage reduces profits at rate l^L while a decrease in wages raises profits at rate l^H .

Now consider question (2), which asks how input demands change with prices. To apply the theorems of this chapter, we must reduce the two variable problem to a one variable problem, which we do by dynamic programming. Notice that in a two-variable maximization problem, the order in which maximization occurs does not affect the agent’s choices. So, without loss of generality, we will consider a “two-stage” maximization approach, with labor being chosen for each given level of capital and then capital being chosen. (This corresponds roughly to the usual short-run versus long-run story.) Then, it is immediate that

$$\max_{k,l} \pi(k, l; r, w) = \max_k \left\{ \max_l \pi(k, l; r, w) \right\}$$

While the special structure of this problem implies that if the production function is continuous, the optimal choices of capital and labor are unique for almost all prices in this

market,⁴ we still need to specify how to handle the cases of multiple optima. While we could perform comparative statics on the set of optimizers (and will do so explicitly in the treatment of Chapter 4), for simplicity of exposition in this section we will study the highest optimizers. Thus, we define:

$$\hat{l}(k; r, w) = \sup_l \left\{ \arg \max_l \pi(k, l; r, w) \right\} \quad (2.2)$$

$$\hat{k}(r, w) = \sup_k \left\{ \arg \max_k \pi(k, \hat{l}(k; r, w); r, w) \right\} \quad (2.3)$$

With this notation in place, we can begin to use the Theorems of this chapter to answer comparative statics questions about input demands. First, we ask how the profit-maximizing choice of capital changes with changes in the rental price, r .

Our first observation concerns the function $\hat{l}(\cdot)$. Notice that

$$\arg \max_l \pi(k, l; r, w) = \arg \max_l pF(k, l) - rk - wl = \arg \max_l pF(k, l) - wl$$

so that the choice of labor in this sub-problem is independent of the rental price of capital (since capital is fixed, its price does not affect the optimization). Thus, we write $\hat{l}(k; r, w) = \hat{l}(k; w)$. Further, let $v(k; w) \equiv pF(k, \hat{l}(k; w)) - wl\hat{l}(k; w)$. Then, the profit-maximizing level of capital and the optimal value of the objective in the original problem are the same as in the following one variable problem:

$$\max_k v(k, w) - rk \quad (\text{P4}')$$

It should be clear that, no matter what the properties of F , the objective function in (P4') satisfies increasing differences in $(k, -r)$, and thus the firm's optimal choice of k will be nonincreasing in the rental price. Further, this result will hold for any input cost function $h(r, k)$ which satisfies increasing differences, not just the function $h(r, k) = r \cdot k$. Applying Theorem 2.5, we see that if we admit a rich enough class of production functions, increasing differences of a generalized input cost function $h(r, k)$ will be a critical sufficient condition for comparative statics.

We next ask how the optimal choice of labor changes with the rental price of capital (r). A standard price-theoretic definition holds that capital and labor are substitutes or complements depending on whether an increase in the price of capital increases or reduces the optimal quantity of labor. We shall use a different definition in this monograph – one that is not so closely tied to price theory. The new definition is mathematically equivalent to this older definition in the two input case if we limit attention to firms with strictly concave production functions.

⁴ So long as the production function is continuous, then the Envelope theorem (2.7) states that the value function H is differentiable almost everywhere, with $H'(\theta) = h_\theta(\mathbf{x}, \theta)$ for all $\mathbf{x} \in \mathbf{x}^*(\theta)$. But in this problem, $\pi_w = -l$ and $\pi_r = -k$, which of course vary with the choices. In order for the value function to have partial derivatives almost everywhere, it must follow that almost everywhere, there is a unique optimizer.

Specifically, we define two inputs to be *complements* if an exogenous increase in one input increases the returns to using more of the other input; two inputs are *substitutes* if such an exogenous increase in one input *decreases* the returns to using more of the other. For the problem at hand, then, capital and labor are complements in this sense if $F(k,l)$ satisfies increasing differences in (k,l) , while capital and labor are substitutes if $F(k,l)$ satisfies increasing differences in $(-k,l)$. (In this latter case, we will say that F satisfies *decreasing differences* in (k,l)). By Theorem 2.1, complementarity between the inputs implies that an exogenous increase in capital will lead to an increase in the optimal choice of labor; when the inputs are substitutes, such a change will lead to a decrease in the choice of labor.

The intuitive correspondence of these ideas and their price theoretic counterparts is straightforward. For example, consider the effect of an increase in the rental price of capital (r). There is no direct effect of such a change on the returns to capital in the profit function π defined in (P4); the entire effect comes indirectly through the effect of r on the choice of capital. Raising r reduces the use of capital, and that will increase or decrease the use of labor according to whether the reduction in capital increases or decreases the returns to labor, that is, whether the production function has decreasing or increasing differences. So, if capital and labor are substitutes in the sense that the production function has decreasing differences, then they are also substitutes in the sense that an increase in the rental price of capital increases the use of labor. Similarly, if they are complements in the sense that the production function has increasing differences, then they are complements in the price theoretic sense that an increase in the rental price of capital reduces the use of labor. Moreover, as we shall show, these conditions are critical sufficient conditions in case the production function is concave and the input prices are arbitrary positive numbers.

One attractive feature of the new theory is that intuitive arguments such as the one just given often have exact formal analogues. We now give such a formal argument below. Recall from above that we may write $\hat{l}(k;w)$ for $\hat{l}(k;r,w)$: the level of r does not directly affect the choice of l , given that the level of k is fixed. For the moment, let us fix w and suppress it in our notation. Then, a pair that solves problem (P4) can be written as $(\hat{k}(r), \hat{l}(\hat{k}(r)))$. By Theorem 2.3, a sufficient condition for $\hat{l}(k)$ to be nondecreasing is that F has increasing differences, which in the smooth case means $F_{kl} \geq 0$. In that event, one may conclude that $\hat{l}(\hat{k}(r))$ is nonincreasing, which matches the condition of the intuitive argument.

These conclusions are special cases of more general results to be proved in Chapter 4; Chapter 4 will also show that the conditions derived above are critical sufficient conditions for the conclusion that the inputs are price-theoretic substitutes or complements. For now, we formalize our statements of the sufficient conditions for the two input case as follows:

Proposition 2.8 Consider problem (P4). Let $k^(r,w) \equiv \hat{k}(r,w)$, and let $l^*(r,w) \equiv \hat{l}(\hat{k}(r,w);w)$.*

- (i) *Then (with no restrictions on $F(k,l)$), $k^*(r,w)$ is nonincreasing in r , and $l^*(r,w)$ is nonincreasing in w .*
- (ii) *If $F(k,l)$ satsfies increasing differences in (k,l) , $k^*(r,w)$ and $l^*(r,w)$ are nonincreasing in both*

arguments; in contrast, if $F(k,l)$ satisfies decreasing differences in (k,l) , $k^*(r,w)$ is nondecreasing in w and $l^*(r,w)$ is nondecreasing in r .

(iii) Results (i) and (ii) continue to hold if the objective function in (P4) is replaced by $\hat{\pi}(k,l;r,w) \equiv p F(k,l) - h^k(k;r) - h^l(l;w)$, so long as the functions $h^k(k;r)$ and $h^l(l;w)$ satisfy increasing differences.

Finally, consider question (3), which concerns how the total output $F(k,l)$ varies with the wage rate. Economists wishing to apply the implicit function approach to this problem would probably first characterize the optimal choices of k and l , and then substitute those choices into the production function to see what might be learned about the optimal level of output. Such an approach can be difficult to execute even when the profit objective is concave, and it becomes doubly so when the profit function is not concave.

So that we may apply our comparative statics theorems, it will be convenient to reformulate the problem as a standard cost-minimization problem and then examine the choice of output. Thus, we will study a two step maximization problem, with the firm first choosing x and then choosing k and l to maximize profits subject to $F(k,l)=x$. In more familiar terms, we define the following cost function, which represents the minimum cost to the firm of producing output x given wages w :

$$C(x,w) = \min_{k,l} rk + wl \quad .$$

subject to $F(k,l) = x$

Further, let

$$l^{**}(x,w) = \left\{ l \left| (\exists k) (k,l) \in \arg \min_{k,l \text{ s.t. } F(k,l)=x} rk + wl \right. \right\}.$$

The firm's problem can be restated as follows (where we now allow the price of output to vary):

$$\max_x px - C(x,w) \quad (\text{P4}'')$$

Notice that the family of functions px generated by varying the one parameter p is a positive one parameter family. So, according to Theorem 2.5', if $-C$ is smooth, decreasing and concave (C is smooth, increasing and convex), then the critical sufficient condition for output to decrease with the wage rate w in this problem is that C has increasing differences. Moreover, according to Theorem 2.3, this condition is still sufficient even if C is not convex.

Then, we are left with the final question, when does C satisfy increasing differences? According to the Envelope Theorem, $C_w(x,w) = l^{**}(x,w)$ when this set is a singleton. Moreover, by Theorem 2.2, C has increasing differences exactly when $C_w(x,w)$ is nondecreasing in x , that is, when $l^{**}(x,w)$ is nondecreasing in x . By definition, this means that the critical sufficient condition for total output to decrease with the wage rate is that labor be a "normal input," that is, an input whose use grows with the level of output. But, we showed in our analysis of (P1) in Chapter 1 that a critical sufficient condition for labor to be a normal input is that the marginal rate of technical substitution of labor for capital is diminishing in the amount of output required.

Notice how the formal analysis corresponds to the economic intuition of the problem: Suppose that a firm is currently optimizing, taking wages as given. A rise in the wage reduces output robustly only if it raises the incremental cost of producing output. And, it does that precisely when increased output leads to an increased use of labor, that is, when labor is a normal input. The simple principles embedded in the theorems are the only relevant ones for qualitative comparative statics in separable models like the classical model of the firm

Summarizing our result, we have the following proposition.

Proposition 2.9 Consider problem (P4). Assume that the production function F is continuous. Then if labor is a normal input, the profit-maximizing level of output is a nonincreasing function of the wage rate w .

The sufficient conditions we have obtained in this section do not exploit any hypotheses about convexity. Furthermore, our analysis of critical sufficient conditions implies that adding convexity assumptions does not permit us to weaken our sufficient conditions. In other words, for the models studied in this section, concavity of the objective is a uselessly restrictive condition that is entirely dispensable condition for the study of comparative statics.

2.6.2 The LeChatelier Principle: The Two Input Case

One success of the new methods of analysis for the theory of the firm is its extension of the LeChatelier principle. In traditional price theory, the principle is used to compare long- and short-run demand for an input, but it also has other important applications in game theory and the economic theory of organizations.

In its price theoretic version, the principle partially formalizes the intuitive idea that the demand for an input, say oil, is more elastic in the long-run than in the short-run because the firm has more ways to economize on oil in the long-run. For example, if the price of oil rises, the firm can respond by substituting other energy forms, closing down operations that are oil-intensive, investing in fuel efficient machinery and insulation, and so on. Over longer horizons, there are more ways to economize on the use of oil, and so there can be a greater response. That is the intuitive formulation, but the usual textbook treatments lead to a very much narrower formal conclusion.

Suppose (as in (P4)) there are two inputs to production, capital k and labor l , and that labor is variable but capital is fixed in the short run. Holding the output price p and the rental rate r fixed, we will vary the wage w , and ask the following question: how does the response of labor to a change in the wage differ between the short run, when capital is fixed, and the long run?

As in equations (2.2) and (2.3), let $\hat{k}(r, w)$ denote the (long-run) optimal quantity of capital to employ at each r and w , and let $\hat{l}(k; w)$ denote the short run optimal choice of labor. Since we are holding r fixed, we will suppress it in our notation. For a given wage w , the long-run demand for labor will be denoted $l_{LR}(w) \equiv \hat{l}(\hat{k}(w); w)$. According to the usual treatment of these issues in textbook analyses, if the relevant functions are differentiable at w , then

$l_{LR}'(w) \leq \frac{\partial}{\partial w} \hat{l}(k, w) \Big|_{k=\hat{k}(w)}$.⁵ This is supposed to formalize the idea that demand varies more in the long-run than in the short-run. Strikingly, *it requires no assumptions beyond differentiability of the factor demand functions.*

Paul Samuelson regarded this principle as one of the great contributions of modern mathematical economics and even emphasized its significance in his Nobel prize lecture (Samuelson, 1971). Nevertheless, there are several important reasons to be disappointed in both the principle and its proof. First, the mathematical conclusion in the traditional version of the principle falls far short of the intuitive conclusion. It is a local result that holds only for infinitesimal price changes.⁶ Moreover, the traditional principle is limited to the case where the production function is strictly convex over the relevant neighborhood, so that demand for the input will vary continuously. Finally, the traditional mathematical analysis involves ideas about convexity that are not present in intuitive argument. The traditional analyses do not answer the questions: Why does the general result hold only locally? Are there plausible, intuitive conditions under which it holds globally? What role does convexity play in supporting the conclusion?

The explanations of the principle that are typically offered in graduate textbooks do not answer these questions and serve only to deepen the mystery. In contrast, the robust comparative statics approach mimics and amplifies the intuitive argument, clarifying its assumptions, adding rigor, showing when the conclusion holds globally, and explaining why the global conclusion can sometimes fail.

That analysis proceeds as follows. Suppose first that capital and labor are complements, that is, that F has increasing differences. By Theorem 2.1, this implies that the demand for labor given capital, $\hat{l}(k; w)$, is nondecreasing in k and nonincreasing in w . (Intuitively, this is because capital increases the marginal product of labor, while higher w makes labor more expensive.) Also, as we showed in Proposition 2.7, $\hat{k}(w)$ is nonincreasing, since labor will fall in response to higher wages and lower labor decreases the returns to increasing capital. Now, consider an initial wage w_0 . Then, for any increase in the wage to $w > w_0$, $\hat{k}(w_0) > \hat{k}(w)$, and we have $\hat{l}(\hat{k}(w_0); w) \geq \hat{l}(\hat{k}(w); w) = l_{LR}(w)$: the quantity of labor demanded falls more in response to a price increase in the long run than in the short run. The intuition is simple: the higher short-run level of labor will be chosen when the marginal product of labor is higher (due to the higher fixed capital stock).

⁵A typical treatment is in Varian (1992), page 47. The formal analysis follows from the fact that the difference between long run and short run profits are minimized at the wage level where fixed level of capital happens to be optimal. Since a differentiable function must be locally convex at a minimum, the marginal effect of changing the wage must be increasing, and thus the long-run labor demand must be changing faster than short-run labor demand.

⁶This qualification is often not emphasized; for example, Varian (1992) gives the formal analysis, then claims that “this analysis confirms the intuition” that “the firm will respond more to a price change in the long run, since, by definition, it has more factors to adjust in the long run than in the short run.” Nowhere does he mention the qualification that the principle is not valid for large changes in prices.

On the other hand, when F has decreasing differences so that capital and labor are substitutes, then by Theorem 2.1, $\hat{l}(k; w)$ is nonincreasing in both arguments and, as we have seen, $\hat{k}(w)$ is nondecreasing. So, again, for all $w > w_0$, $\hat{l}(\hat{k}(w_0); w) \geq \hat{l}(\hat{k}(w); w) = l_{LR}(w)$: the quantity of labor demanded falls more in the long run than in the short run.⁷

Note well that this is a global argument using only the assumptions of increasing or decreasing differences. It applies to both concave and non-concave production functions F and works for price changes of any size. Its logic is essentially the same as the verbal logic. Suppose that oil and other inputs are substitutes, and the price of oil rises. The initial fall in the quantity of oil demanded leads to a higher marginal product of the substitute input, resulting in increased use of the substitute. That change, in turn, further reduces the marginal product of oil and encourages additional reductions in oil consumption. Complementarity between oil and other inputs leads to a similar amplification of the effect of the initial price change, as the initial reduction in the use of oil leads to long-run reductions in the use of complementary inputs (whose marginal products fall as oil usage falls), which in turn encourages further reductions in oil usage. This intuitive idea is reflected exactly in the mathematics.

The new treatment of the LeChatelier principle involves an extra assumption compared to the treatment using traditional methods, namely, the assumption that F has either increasing or decreasing differences. Is this assumption dispensable? It is not. Here is a counterexample.

		Output	
		0	1
New Capital	0	0	56-2p
	1	-24	32-p

Net revenues when the price of oil is p .

Suppose a firm can produce either zero or one units of output using oil and capital as inputs. Capital is fixed in the short run, but in the long run can be increased by a unit, which costs \$24. Oil usage is variable. With the original level of capital, producing a unit of output requires two units of oil. If a unit of new capital is purchased, then producing a unit of output requires only one unit of oil. The price of the output is fixed at \$56.

Let the price of oil initially be $p = \$20$. Then the long-run optimal plan is to produce a unit of output using only the original capital and two units of oil, earning a profit of $\$56 - 2p = \16 . Purchasing a unit of new capital, the maximum profit would be only \$12. If the price of oil rises to \$30, the optimal short-run response is to cease production, netting zero, rather than to continue

⁷A similar treatment can be found in Pollak (1969). Like all the early treatments, this one still incorporates irrelevant assumptions, including the smoothness of demand functions and therefore, implicitly, the strict convexity of the production technology.

producing with the original capital and lose \$4. However, in the long run over which capital is free to adjust, the optimal plan is to increase capital and use one unit of oil to produce a unit of output, netting $\$32 - p = \2 . Thus, in response to the oil price increase, oil consumption falls more in the short run (from 2 units to zero) than the in long run (when usage falls to one unit).

Intuitively, fuel efficient equipment can be a substitute for oil at low oil prices, because its use reduces oil consumption at the current level of output, but also be a complement for oil at high oil prices, when oil might not be used at all but for the availability of the efficient equipment. There is nothing strange or unconventional or non-robust about this possibility.

Although the new version of the LeChatelier principle incorporates an extra assumption (that the production function has either increasing or decreasing differences) compared to the traditional version, it nevertheless encompasses the traditional version as a special case. The reason is that the local version of the LeChatelier principle implies a conclusion about long- and short-run demand only over the range of inputs prices where demand is well approximated by a linear function, and hence over which the production function F is well approximated by a quadratic function. It is easy to see, however, that the new analysis applies to any quadratic function even for non-infinitesimal price changes. For any quadratic function F , $F_{kl}(k,l)$ is a constant, independent of k and l . If $F_{kl} \geq 0$, then F has increasing differences. If $F_{kl} \leq 0$, then F has decreasing differences. Over any domain where F_{kl} does not change sign, the two inputs are either substitutes or complements, and so the new version of the LeChatelier principle applies. To put the same point another way, the traditional assumption that input demands are differentiable imply that, locally, the production function is approximately one with either increasing or decreasing differences. Then the traditional local version of the principle follows from the new one.

The LeChatelier principle is not restricted to the case where the changing parameter is a price or the changing decision is a quantity demanded. The same logic (and proof) has a wide range of applications. For example, suppose a change in the technology of production calls for a long-run increase in the number of product varieties the firm will offer and complementary changes in worker training or organization, but that resistance by affected workers or a shortages of funds or the requirements of planning lead to delays in changing the training or organization. If the product variety decision is adjusted optimally in view of these delays, the result will be a slower increase in product variety.

2.7 Problems

1. Prove that if A and B are compact subsets of \mathbb{R} , then $A \geq B$ if and only if there exist numbers a and b such that $A = (A \cup B) \cap \{x \mid x \geq a\}$ and $B = (A \cup B) \cap \{x \mid x \leq b\}$.
2. Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be an increasing function. Show that if $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ has increasing differences, then so does the composite function defined by $(x,y) \equiv f(g(x),y))$.
3. Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be continuously differentiable. Prove that the function $f(x,\theta) = g(x + \theta)$ has increasing differences if and only if g is convex. [Extra: Prove the same result under the weaker assumption that g is continuous but not necessarily differentiable. Can one prove the same conclusion without the assumption that g is continuous?]
4. Here is a variation on one of the implications of Theorem 2.1. Suppose that $f(x,\theta)$ has strictly increasing differences. Show that if $t' > t$, then for all $x' \in x^*(t')$ and all $x \in x^*(t)$, $x' \geq x$.
5. Consider a firm that chooses its product price p to maximize profits, which are given by $(p - c)Q(p,\theta)$, where Q is a demand function with parameter θ which is decreasing in p . Let $\partial \ln(Q(p,\theta)) / \partial \ln(p) = \varepsilon(p) + \theta$ be the elasticity of demand at various prices, so that an increase in θ simply makes demand less elastic at each price. Show that $p^*(\theta, c)$ is nondecreasing in each argument.
6. A firm uses capital (K) and labor (L) to produce its output. In doing so, it has access to a range of production functions of the form $f(K,L) = \alpha \log K + (1 - \alpha) \log L$, for $\alpha \in [0,1]$. Let r be the cost of capital and w that of labor.

- a) Show that, as r increases, the firm will prefer to use a more labor-intensive technology.
- b) What is the problem with using the implicit function theorem to obtain this conclusion?
7. A firm is completing the development of a new product and is evaluating how long it should wait before launching it. A longer development time allows the firm to improve its production technology, which results in cost savings and better product quality. On the other hand, the firm knows that a direct competitor is working on a similar product, and it realizes that whoever introduces its product first will capture a significant share of the market.

Specifically, suppose that if the competitor enters the market first, the firm will be left with profits equal to π , while if it introduces its product at time t and the competitor hasn't entered yet, it enjoys a profit of $\pi(t) > \pi$ ($\pi' > 0$). Finally, the firm believes that its competitor's time of entry is exponentially distributed with parameter θ :

$$Pr(\text{competitor enters at time } t) = \theta e^{-\theta t}$$

How will the firm react to an increase in (its estimate of) θ ?

8. The arrival of customers at a bank follows a Poisson process with parameter θ :

$$\text{Prob}(k \text{ arrivals}) = \frac{\theta^k e^{-\theta}}{k!}$$

Assume that if no teller is available for service, arriving customers leave rather than wait in line. Assume also that if x tellers are open, the bank incurs a cost of $C(x)$, and that the bank levies a service charge of δ dollars for each transaction.

Show that x is nondecreasing in θ .

9. Consider an economy in which two goods, tillip and quillip, are available. Tillip, which is exclusively imported (in fixed-size lots) by Firm 1, can be consumed or used as an input to make quillip.

Hence, Firm 1 sells a share x of the total amount of tillip imported every year, T , to a firm called Firm 2, which has a production function of the form $Q = \tau^\beta$, $\beta \in (0,1)$, where τ is the amount of tillip that Firm 2 decides to purchase. The market for quillip is competitive, and the prevailing price is P_Q . Assume that Firm 1 can quote any price it deems reasonable, and Firm 2 does *not* attempt to bargain with 1: it just takes that price as if exogenously given. Finally, quillip is subject to a $t\%$ value-added tax; that is, for any unit of quillip sold, Firm 2's net revenue is $(1-t)P_Q$.

Firm 1 sells the remaining portion of tillip to directly to consumers, acting as a monopolist facing demand given by $P = a - b\tau$, where τ is the amount of tillip supplied for consumption. Note that the price P prevailing in the consumption market for tillip need not be the same as the price that Firm 1 quotes to Firm 2. Show that x is non-increasing in t , and give intuition for the result. Would things be different if Firm 1 employed a non-linear tariff when selling to consumers?

10. Consider two competitive firms, Firms 1 and 2 operating in distinct markets. Firm 1 produces a product called O_1 using factor X_1 as its only input; it is characterized by a production function $f(x_1)$ and costs $c(x_1)$. Firm 2 produces O_2 using factor X_2 as its only input, but its output is adversely influenced by the amount of input X_1 utilized by Firm 2: its production function is thus $g(x_2, x_1)$ and satisfies $g(x_2, x'_1) < g(x_2, x_1)$ for $x'_1 > x_1$. For example, Firms 1 and 2 might be located next to each other, X_1 might be “clean air” and X_2 labor (the idea being that workers become less productive if the pollution level increases). Firm 2's costs are given by $c_2(x_2)$.

In light of this interpretation, suppose that we wish to increase the amount of labor used by Firm 2.

- a) Consider introducing a tax t on Firm 1's revenues, so that Firm 1's net profit after taxes is given by $f(x_1)(1 - t) - c_1(x_1)$. (i) Under what conditions will an increase in the tax rate result in less consumption of X_1 , independent of the cost structure $c_1(x_1)$? (ii) If g has the property given two paragraphs above, and is increasing in x_2 , does it necessarily follow that an increase in x_1 leads to a lower optimal value of x_2 ? Provide a counterexample.

- b) Indicate what assumptions are “critically sufficient” to ensure that the introduction of a revenue tax on Firm 1’s output actually induces Firm 2 to hire more labor, X_2 .

3 Single Variable Problems with Non-Separable Contexts

Note: this chapter is very preliminary and probably has typos

This Chapter studies problems which can be written in the following form:

$$\max_{x \in S} V(x, g(x), \theta) \quad (3.1)$$

Alternatively, we can write the objective function as $V(x, y, \theta)$, with the constraint that $y = g(x)$. Problem (P3) fits into this form, and we analyzed this class of problems in Theorems 1.3 and 1.4.

Notice that the separable contexts we studied in Chapter 2 can be incorporated in the framework described by (3.1): two special cases of (3.1) are $V(x, g(x), \theta) = f(x, \theta) + g(x)$, and $V(x, g(x), \theta) = f(x, \theta) \cdot g(x)$. We treated separable contexts as separate cases in our analysis because they arise frequently in economic applications. The methods described in this section are most useful for all contexts which are not additively or multiplicatively separable, but where the family of functions which varies across the context can be written $g(x)$, a function does not depend on θ . In our formal notation for a context as a triple $\{f, G=\{g\}, \pi: f \times G \rightarrow H\}$, problem (3.1) has $f: X \times \Re \times T \rightarrow \Re$, $g: X \rightarrow \Re$, and $\pi(f, g) = f(x, g(x), t)$. Throughout this section, we will use the notation V instead of f , to indicate the difference in structure from the separable contexts.

To motivate further the results of this section, let us consider some additional examples of problems which take the form of (3.1). First, consider a classic “screening,” or price-discrimination, problem. A monopolist faces heterogeneous consumers, so that she would like to sell different price-quantity bundles to different consumers. However, the monopolist cannot distinguish the different consumers, and so instead offers a menu of price-quantity bundles from which the consumers may choose. For any quantity q chosen, let $p(q)$ be the corresponding price on the menu. The consumers’ utilities are given by $u(q, p(q), \theta)$, where θ represents the agent’s type, and utility is increasing in the first argument and decreasing in the second argument. For each type θ , the monopolist would like to specify a price-quantity bundle which is incentive-compatible for that type, so that type will choose their assigned bundle over all other bundles. In solving this problem, the comparative statics question of interest is, how must type affect utility so that higher types always choose higher quantities? This monotonicity is the first step for establishing the existence of a menu which is incentive-compatible.

A related problem is the Spence education signaling model, where different type workers face different costs of becoming educated, but education is not directly productive. However, type is correlated with productive ability. In this model, equilibrium wages are a function of educational decisions, $w(e)$. An agent of type θ has utility given by $u(e, w(e), \theta)$, where utility is decreasing in the first argument and increasing in the second argument. The question then becomes, when do higher ability workers choose higher levels of education, irrespective of the wage function?

A final example is the savings-growth model studied by Frank Ramsey, Tjalling Koopmans, and David Cass. Koopmans (1965) studied a general formulation of recursive utility which allows a consumer’s rate of time preference to be sensitive to her level of periodic consumption. Thus, the utility of the infinite consumption stream $c = (c_0, c_1, \dots)$ is given by $U(c_0, u_1)$, where $u_1 =$

$U(c_1, c_2, \dots)$. The agents maximize utility subject to the constraints that (i) wealth w_t is always nonnegative, (ii) periodic consumption c_t plus savings s_t is equal to current wealth w_t , and (iii) tomorrow's wealth is determined as a function of today's savings as follows: $w_t = F(s_t)$. In this problem, the question of interest is, When does savings increase with current wealth?

In the next few sections, we will analyze critical sufficient conditions for comparative statics in these and other problems which take the form of (3.1).

3.1 Sufficient conditions: Single Crossing of Indifference Curves

Theorems 1.3 and 1.4 provide two results which can be used to analyze problem (3.1), so long as the hypotheses of the theorems are satisfied. In this section, we will generalize Theorems 1.3 and 1.4 to apply to problems which are not everywhere differentiable. Further, we will analyze the question of what kinds of contexts, as determined by the family of functions g we wish to consider, have the same critical sufficient condition for comparative statics conclusions.

We saw in Theorems 1.3 and 1.4 that, for a suitably differentiable function, the critical sufficient condition for comparative statics in several contexts was V_1/V_2 nondecreasing in θ . In this section, we generalize this condition to what we call the “single crossing condition,” a condition which requires that the x - y indifference curves of $V(x, y, \theta)$ cross once as a function of θ .

However, in many cases it is also helpful to consider a more “algebraic” intuition as well. We can think of V as incorporating two “effects,” a “direct effect” and an “indirect effect.” The direct effect of x comes through the first argument of V , while the indirect effect results from the effect of x on y , through $g(x)$. In many of our applications, one of the following two scenarios is true: either the direct effect of x is positive and higher values of y decrease the objective, or else the direct effect of x is negative, and higher values of y increase the objective. In either case, the condition that V_1/V_2 is nondecreasing in θ , derived in Chapter 2, can be interpreted as saying that the ratio of the marginal benefits of x to the marginal cost (through g) must be nondecreasing in θ . However, it is important to keep in mind that no monotonicity restrictions have been placed on g itself; g is not a part of the restriction on V_1/V_2 .

More generally, the requirement that V_1/V_2 is nondecreasing in θ is equivalent to single crossing of the x - y indifference curves. In order to discuss indifference curves and upper contour sets without ambiguity, we require that the indifference sets are indeed closed curves and that the preferred sets are comprised of the points above or below the indifference curves:

Either $V(x, y, \theta)$ is strictly increasing in y everywhere or else V is strictly decreasing in y everywhere, and $\{(x, y) | V(x, y, \theta) = k\}$ is a closed curve for each θ, k . (CC)

Under (CC), we know that for every x , θ , and k , there is a unique y , denoted $\tilde{y}(x; \theta, k)$, which satisfies $V(x, y, \theta) = k$. Using that notation (which will be used throughout the remainder of this chapter), we can formally define single crossing of indifference curves.

Definition: Suppose V satisfies (CC). For each $\theta_H > \theta_L$ and $(\hat{x}, \hat{y}) \in \mathbb{R}^2$, denote the indifference curves through (\hat{x}, \hat{y}) as follows: $\tilde{y}_L(x) \equiv \tilde{y}(x; \theta_L, V(\hat{x}, \hat{y}, \theta_L))$ and $\tilde{y}_H(x) \equiv \tilde{y}(x; \theta_H, V(\hat{x}, \hat{y}, \theta_H))$. In the case where $V_2 > 0$, $V(x, y, \theta)$ satisfies the **single crossing condition (SCC)** if, for all $x'' > x'$, $\tilde{y}_H(x') \leq (<) \tilde{y}_L(x'')$ implies that $\tilde{y}_H(x'') \leq (<) \tilde{y}_L(x'')$. If $V_2 < 0$, then the inequalities are reversed.

The definition can be interpreted as follows. We consider two indifference curves, corresponding to θ_H and θ_L , and two arbitrary constant levels of utility. Consider the case where $V_2 > 0$, so that y is a good. We then require that the curve associated with θ_H crosses the curve associated with θ_L at most once, from above. The definition allows that the indifference curves might be equal to one another throughout a region of x , but that once the θ_H -curve is strictly below the θ_L -curve, it will remain strictly below for all higher values of x . Figure 3.1 illustrates a scenario where $V_2 > 0$, and the indifference curves satisfy SCC.

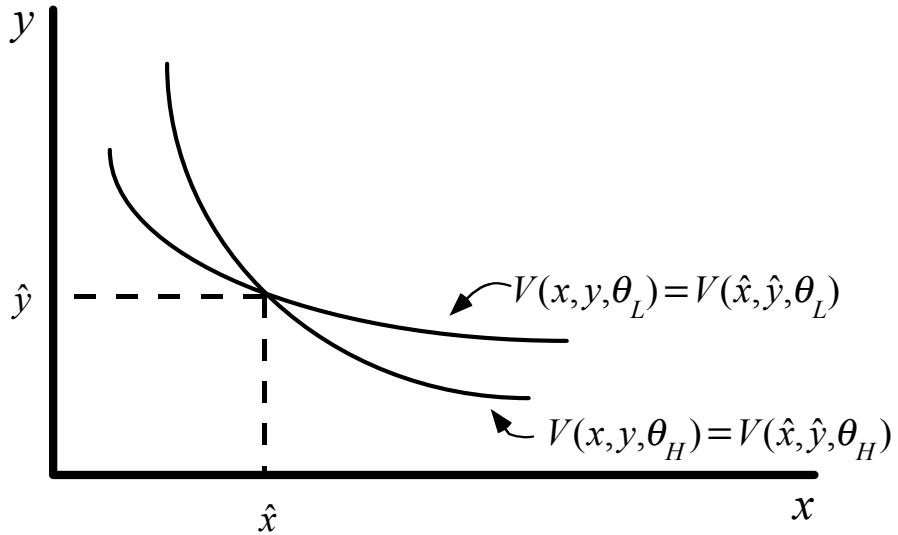


Figure 3.1: The indifference curves satisfy the single crossing condition.

Just as increasing differences can be shown to be equivalent to a condition on a cross-partial derivative when the relevant function is appropriately differentiable, the single crossing condition is equivalent to monotonicity of the function $V_1 / |V_2|$. This is formalized in the following theorem:

Theorem 3.1: Suppose that $V(x, y, \theta)$ is continuously differentiable in x and y , and further satisfies (CC). Then V satisfies the SCC if and only if $V_1 / |V_2|$ is nondecreasing in θ .

Proof: It suffices to show that, for $x'' > x'$, if $V(x', y', \theta_L) = V(x'', y'', \theta_L) \equiv k$ then $V(x', y', \theta_H) \leq V(x'', y'', \theta_H)$. In the following argument, \tilde{y} always refers to $\tilde{y}(x, \theta_L, k)$ and its derivative \tilde{y}' refers to the partial derivative with respect to x . By the implicit function theorem, $\tilde{y}' \equiv \frac{\partial}{\partial x} \tilde{y} = -(V_1 / V_2)(x, \tilde{y}, \theta_L)$. So,

$$\begin{aligned}
& V(x'', y'', \theta_H) - V(x', y', \theta_H) \\
&= \int_{x'}^{x''} \frac{d}{dx} V(x, \tilde{y}(x, \theta_L, k), \theta_H) dx \\
&= \int_{x'}^{x''} (V_1 + V_2 \tilde{y}') dx = \int_{x'}^{x''} \left(\frac{V_1}{V_2} (x, \tilde{y}, \theta_H) + \tilde{y}' \right) V_2 (x, \tilde{y}, \theta_H) dx \\
&\geq \int_{x'}^{x''} \left(\frac{V_1}{V_2} (x, \tilde{y}, \theta_L) + \tilde{y}' (x, \theta_L, k) \right) V_2 (x, \tilde{y}, \theta_H) dx \\
&= \int_{x'}^{x''} V_2 (x, \tilde{y}, \theta_H) dx = 0
\end{aligned}$$

The inequality step follows by comparing the integrands point by point using the hypothesis that V_1/V_2 is higher for θ_H than for θ_L . If $V_2 > 0$, then this hypothesis implies that V_1/V_2 is nondecreasing in θ , while if $V_2 < 0$ the hypothesis implies that V_1/V_2 is nonincreasing. In either case, the left-hand side of the inequality will be higher than the right-hand side point by point.

$\frac{V_1}{V_2}(\hat{x}, \hat{y}, \theta_H) < \frac{V_1}{V_2}(\hat{x}, \hat{y}, \theta_L)$. Since V is continuously differentiable in x and y , this inequality

must hold for an open set $D \subset \mathbb{R}^2$. Then, pick (x', y') and (x'', y'') on a θ_L -indifference curve which passes through D , and use the above arguments to show that $V(x'', y'', \theta_H) - V(x', y', \theta_H) < 0$. This violates the single crossing condition.

Theorem 3.2. Assume the following about the problem (3.1):

(S3.2) No restrictions on S .

(V3.2) For all (x, θ) the function $V(x, y, \theta)$ is strictly increasing or strictly decreasing in y . For all (θ, k) , the indifference set $\{(x, y) | V(x, y, \theta) = k\}$ is a closed curve.

(G3.2) No restrictions imposed on g .

Then, $x^H(\theta; g)$ is nondecreasing in θ for all functions g if and only if V satisfies the single crossing condition. Similarly, $x^L(\theta; g)$ is nondecreasing in θ for all functions g if and only if V satisfies the single crossing condition (SCC).

Proof: Let us consider the case where V is strictly increasing in y ; the case where V is decreasing is analogous. First, we show that the monotone comparative statics conclusion implies SCC. Choose $\theta_H > \theta_L$, and (\hat{x}, \hat{y}) . Let $\tilde{y}_L(x) = \tilde{y}(x; \theta_L, V(\hat{x}, \hat{y}, \theta_L))$, and let $\tilde{y}_H(x) = \tilde{y}(x; \theta_H, V(\hat{x}, \hat{y}, \theta_H))$. Then, define $g(x) = \tilde{y}_L(x)$ when $x \leq \hat{x}$, while $g(x) = \min(\tilde{y}_L(x), \tilde{y}_H(x)) - \varepsilon$ for $x > \hat{x}$. In this case, the highest optimizer of $V(x, g(x), \theta_L)$ is \hat{x} .

First suppose that there exists an $x' < \hat{x}$ such that $\tilde{y}_H(x') < \tilde{y}_L(x')$. Then, we have $V(x', g(x'), \theta_H) = V(x', \tilde{y}_L(x'), \theta_H) > V(x', \tilde{y}_H(x'), \theta_H) \equiv k_H$, by our definition of g and since $V_2 > 0$. Likewise, for all $x'' > \hat{x}$, $k_H = V(x'', \tilde{y}_H(x''), \theta_H) > V(x'', g(x''), \theta_H)$. Thus, the agent

prefers x' to all $x \geq \hat{x}$, and the monotone comparative statics conclusion fails.

Now suppose that V satisfies SCC. Pick any g , and choose $\theta_H > \theta_L$. Let $\hat{x} = \sup\{\arg \max_x V(x, g(x), \theta_L)\}$. Let $\tilde{y}_L(x) = \tilde{y}(x; \theta_L, V(\hat{x}, g(\hat{x}), \theta_L))$, and let $\tilde{y}_H(x) = \tilde{y}(x; \theta_H, V(\hat{x}, g(\hat{x}), \theta_H))$.

First, since V is strictly increasing in y and \hat{x} is optimal at θ_L , $\tilde{y}_L(x) \leq g(x)$. Consider $x' < \hat{x}$. Since V satisfies SCC, $\tilde{y}_H(x') \geq g(x')$, where the inequality is strict unless $\tilde{y}_H(x') = \tilde{y}_L(x')$ on $[x', \hat{x}]$. Thus, $V(x', g(x'), \theta_H) \leq V(\hat{x}, \tilde{y}_H, \theta_H)$, and either (i) \hat{x} is strictly preferred to x' , or (ii) the agent's preferences over $[x', \hat{x}]$ do not change between the two values of θ . Second, if new optimizers appear which are greater than \hat{x} , this can only further increase in x^H . Thus, $x^H(\theta_H, g(\cdot)) > x^H(\theta_L, g(\cdot))$.

The intuition behind the algebraic proof can be illustrated quite easily with a graph. Let us refer to the set of (x, y) points where $y = g(x)$ as the “feasible set.” Figure 3.2 illustrates two indifference curves through the point (\hat{x}, \hat{y}) , corresponding to the two different values of the parameter θ , and where the indifference curve corresponding to θ_H crosses the curve corresponding to θ_L from above. If we suppose that agent θ_L prefers (\hat{x}, \hat{y}) to any feasible (x, y) pair such that $x < \hat{x}$, then all feasible (x, y) pairs must lie below θ_L 's indifference curve through (\hat{x}, \hat{y}) . But, since θ_H 's indifference curve through (\hat{x}, \hat{y}) is *above* θ_L 's indifference curve throughout the region $x < \hat{x}$, agent θ_H clearly must find (\hat{x}, \hat{y}) preferable to all feasible (x, y) pairs in that region as well!

On the other hand, points in the shaded area satisfy $x > \hat{x}$ and lie between the two agents' indifference curves. Throughout this region are points which agent θ_L finds inferior to (\hat{x}, \hat{y}) , but which agent θ_H prefers to (\hat{x}, \hat{y}) . Thus, if any points in this region are feasible, agent θ_H will choose them over (\hat{x}, \hat{y}) . So we have proved our claim: if increasing θ leads to single crossing of indifference curves, then increasing θ will lead to a (weak) increase in the highest choice of x , *no matter what the feasible set, defined by the function g , looks like!* For example, g could be discontinuous or nondifferentiable, and the same result obtains.

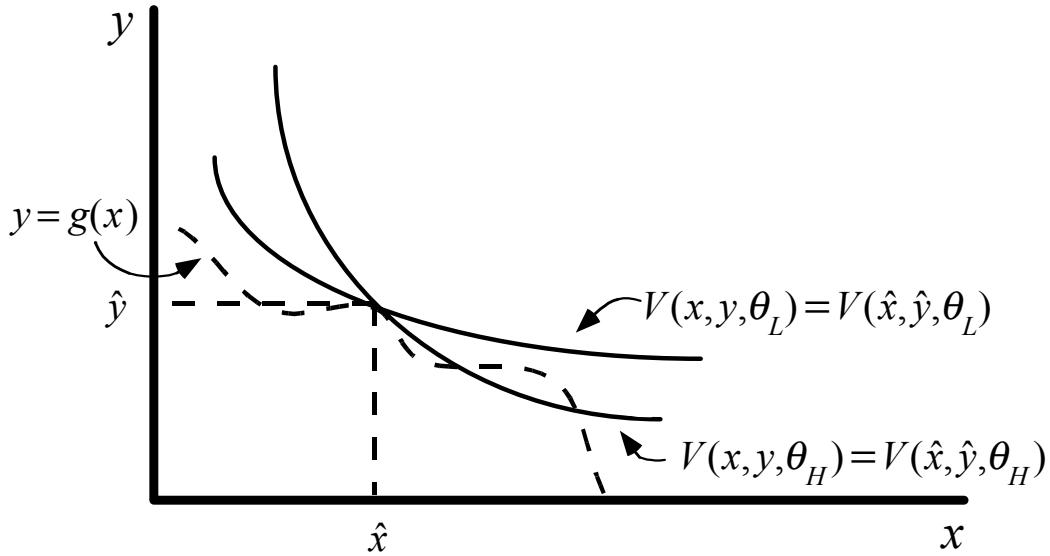


Figure 3.2: The shaded region is the set of (x,y) points which θ_L finds inferior to (\hat{x}, \hat{y}) , but which agent θ_H prefers to (\hat{x}, \hat{y}) . Some points in the feasible set may lie in this region, and θ_H will choose these points over (\hat{x}, \hat{y}) .

To see how Theorem 3.2 can be applied, consider the example of education signaling described at the beginning of this chapter. In his celebrated and controversial analysis of this problem, Michael Spence (1971) argued that more highly educated workers might be paid higher wages even if education did not contribute to productivity, provided that abler workers found it relatively more enjoyable or less personally costly to acquire education. Let θ denote the marginal revenue product or “ability” of a worker. If a firm that is a competitor in the labor market were perfectly informed about θ , it would be willing to pay a wage of up to θ to hire the worker. Let $U(e, w, \theta)$ be the utility of a worker of ability θ who obtains education level e and enjoys earnings (“wages”) w . Spence formalized the assumption that abler workers find it cheaper to acquire education as the mathematical assumption that U_e/U_w is increasing in θ .

Suppose that employers cannot observe θ but can observe each worker's level of education e . Finally, suppose that the market offers wages as a function of educational attainment given by $w(e)$, so that each worker knows that by acquiring education level e , s/he can enjoy earnings of $w(e)$. Theorem 3.1 gives the a key result for the analysis of this model: no matter how education affects the wage the worker receives, abler workers will always acquire more education. That is, the optimal choice $e^*(\theta|w)$ is monotone nondecreasing in θ for all w . With this result in hand, it is then straightforward to show that if education is costly ($\partial U/\partial e < 0$), then workers with more education are paid (weakly) higher wages in equilibrium, even though education is unproductive.

To see a second example, recall the savings-growth model described in the introduction to this chapter. Koopmans (1965) introduced and used “recursive utility” for the analysis, where the utility of the infinite consumption stream $c = (c_0, c_1, \dots)$ is hypothesized to be equal to

$U(c_0, c_1, \dots) = W(c_0, u_1)$ where $u_t \equiv U(c_t, c_{t+1}, \dots)$. This utility is to be maximized subject to the constraints (1) that wealth w_t must at all times be nonnegative, (2) that periodic consumption c_t plus savings s_t must be equal to current wealth w_t , and (3) that there is a function F that relates today's savings and tomorrow's wealth according to $w_{t+1} = F(s_t)$. The initial wealth w_0 is a datum of the problem.

Let $V(w_t)$ denote the optimal level of u_t , given w_t . Then, in each period the agent solves $\max_{0 \leq s_t \leq w_t} W(w_t - s_t, V(F(s_t)))$. If $\sigma(\cdot | F)$ denotes the optimal savings plan, so that in each period the optimal savings is given by $s_t = \sigma(w_t | F)$, then we can use Theorem 3.2 to show that savings increases in current wealth (and thus, an increase in initial wealth increases savings in all periods) if $-W_1/W_2$ is nondecreasing in w_t . This is equivalent to saying that the marginal rate of substitution of current for future consumption is decreasing in w_t : wealthier individuals need less future consumption to compensate them for giving up current consumption, and so they shift consumption towards the future.

We can generalize Theorem 3.2 to perform comparative statics analysis on sets of optimizers, as in our analysis of separable contexts. The following theorem has a stronger comparative statics conclusion, applying to sets of optimizers as discussed in detail in Chapter 2.2:

Theorem 3.2'. Assume the following about the problem (3.1):

(S3.2') No restrictions on S .

(V3.2') The function $V(x, y, \theta)$ satisfies $V_2 \neq 0$, and the indifference set $\{(x, y) | V(x, y, \theta) = k\}$ is a closed curve for all (θ, k) .

(G3.2') No restrictions imposed on g .

Then, $X^*(\theta; g)$ is nondecreasing in θ for all functions g if and only if V satisfies the single crossing condition.

The necessity part of Theorem 3.2 shows that single crossing of indifference curves cannot be relaxed without violating the comparative statics conclusion for some g . This is suggestive of why the single crossing condition is so prevalent in information economics: it cannot be relaxed without violating monotone comparative statics conclusions. However, the theorem states that the single crossing condition is necessary when g can be any function at all. What happens if there is more structure on the problem? Can the single crossing condition be relaxed? The next section studies criticality, finding exactly how narrow the context G must be before we can begin to relax our global single crossing condition without necessarily violating the comparative statics conclusion.

3.2 Critical Sufficient Conditions

In Chapter 1, we studied two different theorems which are relevant for the analysis of (3.1), Theorems 1.3 and 1.4. In both theorems, the SCC was a critical sufficient conditions for

comparative statics conclusions. The main differences between Theorems 1.3 and 1.4 lie in the additional monotonicity restrictions of (V4), and the much smaller class of functions g described in (G4). (G3) allows any function g , while (G4) restricts attention to functions g which can be written $g(x) = a x + b$. However, the conclusion of both theorems is the same: the single crossing condition is a critical sufficient condition for comparative statics conclusions to hold throughout the relevant context.

In this section, which is analogous to Chapter 2.3, we analyze the following questions: For what contexts is the single crossing condition a critical sufficient condition for comparative statics? How rich does the set G have to be so that the single crossing condition is required in order to guarantee a comparative statics result throughout the context?

Like Chapter 2.3, the answer will generally be that if regions of x values can be ruled out as potential optima, the behavior of our function on that region will be irrelevant. However, the way in which g enters the function V is slightly more complicated than in separable contexts. Consider the case where $V(x, g(x), \theta)$ is quasi-concave and differentiable. Then, the optimal solution to (3.1) is determined by the first order conditions, as follows:

$$-\frac{V_1}{V_2}(x, g(x), \theta) = \frac{\partial}{\partial x} g(x)$$

Thus, a given \hat{x} is a potential solution to (3.1) in given context if and only if there exists a function \hat{g} in G such that this equality is satisfied at $x = \hat{x}$. This discussion motivates our definition of a “two-parameter family.” A two parameter family is rich enough so that, given any x , slope a , and intercept b , there exists an element of the family so that $g(x) = b$ and $g'(x) = a$. When this is true, we can always find a g in the family so that the first order conditions are satisfies at \hat{x} .

Definition. The family of functions $\{g_t(\cdot) | t \in T\}$ of functions from $S \rightarrow \mathbb{R}$ (where S is a convex subset of \mathbb{R}) is a **full two-parameter family** if, for all $x, a, b \in \mathbb{R}$, there exists a $t \in T$ such that $g_t(x) = b$ and $\frac{\partial}{\partial x} g_t(x) = a$. It is a **positive (negative) two-parameter family** if, for all $x, a, b \in \mathbb{R}$ such that $a > (<) 0$, there is some parameter value $t \in T$ such that $g_t(x) = b$ and $\frac{\partial}{\partial x} g_t(x) = a$.

It turns out that, when the problem is quasiconcave (that is, the preferred sets of V are convex), then a full two-parameter family of functions g is rich enough so that the single crossing conditions is a necessary condition for the comparative statics result to hold throughout the context. If the family of functions g is *not* a full two-parameter family, then the single crossing condition is *not* necessary for the comparative statics result to be robust in the context.

Theorem 3.3. Assume the following about the problem (3.1):

(S3.3) S is a convex set.

(V3.3) The function $V(x, y, \theta)$ is quasi-concave and differentiable in its first two arguments and satisfies (CC)

(G3.3) $\{g_t(x)\}$ is a family of functions which are concave and differentiable in x .

(a) If, in addition, the family G satisfies

(G3.3a) $\{g_t(x)\}$ is a full two parameter family.

then, $x^H(\theta; g_t)$ is nondecreasing in θ for all t if and only if V satisfies the single crossing condition.

(b) If, instead, the family G satisfies

(G3.3b) $\{g_t(x)\}$ is **not** a full two parameter family,

then there exists a V satisfying (V14) such that V does not satisfy the single crossing condition, but $x^H(\theta; g_t)$ is nondecreasing in θ for all t .

We can further extend Theorem 3.3 to problems where some monotonicity is assumed, for example, if the direct effect of x on V is always negative and the indirect effect always positive, or vice versa.

Theorem 3.3'. Assume the following about the problem (3.1):

(S3.3) S is a convex set.

(V3.3') The function $V(x, y, \theta)$ is quasi-concave and differentiable in its first two arguments and satisfies (CC) Further, $V_1 < (>) 0$.

(G3.3) $\{g_t(x)\}$ is a family of functions which are concave and differentiable in x .

(a) If, in addition, the family G satisfies

(G3.3a') $\{g_t(x)\}$ is a positive (negative) two parameter family.

then, $x^H(\theta; g_t)$ is nondecreasing in θ for all t if and only if V satisfies the single crossing condition.

(b) If, instead, the family G satisfies

(G3.3b') $\{g_t(x)\}$ is **not** a positive (negative) two parameter family,

then there exists a V satisfying (V14') such that V does not satisfy the single crossing condition, but $x^H(\theta; g_t)$ is nondecreasing in θ for all t .

Finally, it is interesting to note that Theorems 3.2 and 3.3 can be used in conjunction with one another to show that comparative statics results which hold in problems with a great deal of structure can be extended to problems where much less structure is present. We formalize this idea with the following corollary:

Theorem 3.4. Assume the following about the problem (3.1):

(S3.4) S is a convex set.

(V3.4) The function $V(x, y, \theta)$ is quasi-concave and differentiable in its first two arguments, satisfies $V_2 \neq 0$, and the indifference set $\{(x, y) | V(x, y, \theta) = k\}$ is a closed curve for all (θ, k) .

(G3.4) For each t , $g_t(x)$ can be written as $g_t(x) = ax + b$ for $a, b \in \mathbb{R}$.

Then, $\mathbf{x}^H(\theta; g_t)$ is nondecreasing in θ for all t if and only if $\mathbf{x}^H(\theta; h)$ is nondecreasing in θ for all functions $h: \mathfrak{R} \rightarrow \mathfrak{R}$.

The corollary states that, under the assumptions of the theorem, it is *equivalent* to check that a comparative statics result holds for all linear functions of the form in (G15), and to check that the result holds for *all* functions (no matter how ill-behaved)!

This fact has been exploited by Milgrom (1995), who analyzed how comparative statics results from the classic Arrow (1971) model of investment under uncertainty can be extended to establish a number of new results in a seemingly more complicated model of a firm's decisions under uncertainty (Sandmo, 1971). We will now show how Theorem 3.4 can be applied to this problem.

The model of Arrow concerns an investor with initial wealth W , who chooses the amount of money (x) to invest in a risky security. The remainder of the investor's wealth ($W - x$) is invested in a safe asset. The risky security provides a random rate of $r + \mu_r$, while the safe asset provides a return R . The probability distribution over the random rate of return is denoted $F(r; \theta)$, where the distribution is parameterized by θ . The investor is risk averse, with utility for income given by $U(I, \theta)$, where the parameter θ might also affect the utility function. Thus, the investor's expected profits are given as follows: $\int_r U((W - x)R + x(r + \mu_r), \theta) dF(r; \theta)$. We can think of this function as specifying a benefit to x , increasing the expected wealth according to $\mu_r - R$, and a cost, which is increasing risk. Thus, we may rewrite the problem in the form of (3.1), where $V(x, y, \theta) = \int_r U(y + x \cdot r, \theta) dF(r; \theta)$, and in this case $g(x) = (W - x)R + x\mu_r$. Thus, the agent maximizes $V(x, g(x), \theta)$.

Arrow's investment model has been studied extensively. For example, suppose that the utility function satisfies decreasing absolute risk aversion in income. Then, the investor will decrease his investment in the risk asset (1) if a shift in θ decreases the riskiness of the security, and (3.1) if a shift in θ increases the agent's initial wealth (this could be modeled by letting $V(x, y, \theta) = U(y + x \cdot r, \theta) = \hat{U}(\theta \cdot R + y + x \cdot r)$). Many other questions have been asked of this model as well, such as questions about the effects of tax rates, risk aversion, or other kinds of changes in the probability distribution.

At about the same time that Arrow developed his model of investment, Sandmo was studying the extension of the theory of the firm to the case of uncertainty. He investigated the choice of a risk-averse entrepreneur who chose quantity to maximize expected utility. We generalize his work here, allowing the firm to face the inverse demand function $P(x) + r$, where r is a random shock to demand, and to have the cost function $C(x)$. Thus, we can place this problem in the framework we developed for Arrow's model, using the same function $V(x, y, \theta)$ but now letting $g(x) = x \cdot P(x) - C(x)$. The entrepreneur's objective function is then given by $V(x, g(x), \theta) = \int_r U((P(x) + r)x - C(x), \theta) dF(r; \theta)$. As discussed above, we can use this formulation to ask comparative statics questions about wealth effects, tax rates, changes in risk attitudes or in the mean or variance or riskiness of r , or any combination of these, as well as many other

comparative statics results. However, before the analysis of Milgrom (1995), many fewer comparative statics results were known about Sandmo's models than about Arrow's, and those that were known were derived separately.

To see why, consider the consequence of Theorem 3.4: any comparative statics result which holds for Arrow's problem, will also hold in Sandmo's! This is true because the family of functions given by $g(x) = (W-x)R + x\mu_r$ is a positive two-parameter family. Given any x , a , and b , we can set $\mu_r - R = a$ and set $WR = b$. When comparative statics results are derived for Arrow's problem, typically no *ex ante* restrictions are placed on the parameter values, so that the de facto approach is to ensure that the result will hold throughout the positive two-parameter family described by $g(x)$. But, by Theorem 3.4, the optimal choice of x is nondecreasing in θ for all g in the positive two-parameter family, *if and only if* the optimal choice of x is nondecreasing in θ for all functions h . By formulating the problem so that we see that Arrow's model is just a very restrictive version of Sandmo's, we are able to show that all of the existing results about Arrow's model can be applied immediately to Sandmo's.

In effect, our analysis has identified that the special structure of Arrow's model, the linear structure, was not critical for the comparative statics analysis. Therefore, all of the results from his model remain true when the restrictive linearity assumption is dropped.

3.3 Strict Comparative Statics

As in Chapter 2.3.2, we will briefly provide the conditions for strict monotone comparative statics predictions. Again, differentiability of the objective function is required for our theorem, and it is necessary to place restrictions directly on the function V_1/V_2 as opposed to appealing to the single crossing condition. That is because V_1/V_2 might be strictly increasing *almost everywhere*, but not everywhere, and still satisfy a strict single crossing condition. Unfortunately, almost everywhere is not enough to guarantee that the strict comparative statics result holds. Thus, we have the following theorem (Edlin and Shannon, 1996):

Theorem 3.5. Assume the following about the problem (3.1):

(S3.5) *S is convex.*

(V3.5) *The function $V(x,y,\theta)$ is continuously differentiable in its first two arguments and continuous in its third argument, $V_2 \neq 0$, and the indifference set $\{(x,y) | V(x,y,\theta)=k\}$ is a closed curve for all (θ,k) .*

(G3.5) *The function $g(x)$ is continuously differentiable.*

(I3.5) *The highest solution $\mathbf{x}^H(\cdot;g)$ is interior.*

Then, $\mathbf{x}^H(\cdot;g)$ is strictly increasing for all functions g if and only if V_1/V_2 is everywhere increasing in θ .

This theorem has applications in signaling and screening contexts. For example, in the signaling literature it is often useful to distinguish between “separating,” where the principal can

distinguish between different types of agents (this corresponds to an agent's action being *strictly* increasing in his type), and "pooling," where more than one type of agent might take the same action. Theorem 3.5 gives sufficient conditions to rule out pooling: if $g(x)$ is continuously differentiable and the solution is interior, then the standard "strict" single crossing assumption made in the literature is sufficient to rule out pooling equilibria.

3.4 Problems

Note: these are not debugged

1. Consider the function $V(x,y,\theta) = h(x,\theta) + y$. Show that V has the single crossing property if and only if h has nondecreasing differences. Use this fact and Theorem 3.2 to prove Theorem 2.2 for the special case where function h is continuously differentiable.
2. In the problem of maximizing $(p - c)Q(p,\theta)$ mentioned earlier in this chapter, we suggested the formulation: $V(x,y,\theta) = yQ(x,\theta)$ and $f(x) = x - c$. For this formulation, what restrictions does the single crossing condition impose on the function Q ? Suppose we had instead used the formulation $V(x,y,\theta) = xQ(x,\theta) - cQ(y,\theta)$. What restrictions are imposed on Q in this case? Finally, suppose we had taken $V(x,y,\theta) = (x - c)Q(1/y,\theta)$. What are the corresponding restrictions on Q ?
3. Here is a variation on one of the implications of Theorem 3.2. Suppose that the single crossing condition (1) of Theorem 3.2 is strengthened to require that V_1/V_2 is *increasing* in θ . Show that the solution $x^*(\theta | f, K)$ has the following property stronger than monotonicity in the strong set order: if $\theta' > \theta$ with $x \in x^*(\theta | g, K)$ and $x' \in x^*(\theta' | g, K)$, then $x' \geq x$. (This is the statement that every selection from $x^*(\theta | g, K)$ is monotone nondecreasing in θ .)
4. A worker facing a tax rate of τ and a wage rate of w must decide how many hours to work. Her utility for goods and leisure is given by $U(G,L)$, where the price of goods is taken to be unity. In addition, the worker has an exogenous periodic investment income of I . Thus, by working x hours, the worker can afford to consume goods in the amount $G = (1-\tau)(I+wx)$. Under what condition on U is it always true (regardless of I and w) that an increase in the tax rate reduces hours worked? Would your answer change if the income were not a linear function of hours worked (for example if the worker might have to choose between working 40 hours for \$300, 45 for \$500, or 0 hours for \$0, with no other hour-income options available)?
5. A common result in inventory theory is that the cost of inventories to serve demand at rate R is αR , where the constant of proportionality $\alpha > 0$ depends on various prices and service parameters. Suppose that a firm can avoid inventory costs by communicating in advance with its customers, but that customers may differ in their individual costs of communication. There are Q customers in all and the cost of communicating in advance with a fraction $x \in [0,1]$ of the customers is $QC(x)$. The remaining $(1-x)Q$ customers are served from inventory. The total communication and inventory cost of serving all customers when a

fraction x is served by communication is $QC(x) + \alpha[(1-x)Q]$. Use Theorem 3.2 to show that as Q increases, the optimal value of x declines.

6. In several countries companies are required by law to have their accounts and financial statement audited by a certified accounting firm. In such circumstances, it is commonly believed that auditing by a major, reputable firm makes it unlikely that the Revenue Service will conduct a second audit, although such belief typically has no legal grounds.

Suppose that a firm may choose among a continuum of auditors indexed by their severity: $x = 0$ indicates an extremely lenient auditor, while $x = 1$ indicates a reputable, strict auditor. Denote by θ the "degree of truthfulness" of the company's accounts: $\theta = 0$ indicates severe accounting malpractice, and $\theta = 1$ indicates truthful accounts.

Let $p(x)$ ($p' < 0$) be the probability of an IRS audit if the firm chooses auditor x , and assume that auditor x refuses to issue an audit report with probability $x(1 - \theta)$.

Finally, assume that the company's tax liability is given by $t(\theta)$, and in the event of an IRS audit, the company will be required to pay a fine equal to $F(\theta)$.

- a. Show that companies keeping more truthful accounts will choose less lenient auditors.
- b. More generally, if $q(x, \theta)$ denotes the probability that auditor x will refuse to issue an audit report to a firm with "truthfulness parameter" θ , what is the critically sufficient condition for the above conclusion to hold robustly (w.r.t. to the exact specification of $p(x)$)?
7. In several countries companies are required by law to have their accounts and financial statement audited by a certified accounting firm. In such circumstances, it is commonly believed that auditing by a major, reputable firm makes it unlikely that the Revenue Service will conduct a second audit, although such belief typically has no legal grounds.

Suppose that a firm may choose among a continuum of auditors indexed by their severity: $x = 0$ indicates an extremely lenient auditor, while $x = 1$ indicates a reputable, strict auditor. Denote by θ the "degree of truthfulness" of the company's accounts: $\theta = 0$ indicates severe accounting malpractice, and $\theta = 1$ indicates truthful accounts.

Let $p(x)$ ($p' < 0$) be the probability of an IRS audit if the firm chooses auditor x , and assume that auditor x refuses to issue an audit report with probability $x(1 - \theta)$.

Finally, assume that the company's tax liability is given by $t(\theta)$, and in the event of an IRS audit, the company will be required to pay a fine equal to $F(\theta)$.

Show that companies keeping more truthful accounts will choose less lenient auditors. More generally, if $q(x, \theta)$ denotes the probability that auditor x will refuse to issue an audit report to a firm with "truthfulness parameter" θ , what is the critically sufficient condition for the above conclusion to hold robustly (w.r. to the exact specification of $p(x)$)?

8. A firm has to allocate resources between productivity enhancements and advertising. Assume that the firm's output (say, for fixed factor inputs) as a function of productivity enhancement investment k is given by $f(k)$. Assume also that the price that consumers are willing to pay for

one unit of the firm's output is given by $P(k, \theta)$, where k now denotes the size of the advertising budget and θ is a quality parameter.

a. Under what condition will a higher-quality firm invest a higher portion of its total resources K in advertising? Interpret the condition.

b. Now assume that the firm's output is also a function of the quality parameter. What is the critical sufficient condition now? Interpret.

9. A traveling salesman is visiting a small town in which there are only two shops. The salesman has already visited the first shop, and knows that if he spends α percent of the (working) day with the owner, he can close a sale worth $q(\alpha)$ in commissions. However, he has never visited the second shop; given the information he has gathered, with probability π he will be able to make a sale worth $q_H(1-\alpha)$ spending $1 - \alpha$ percent of his time there, and he will make no sale with probability $1 - \pi$. Assume both q and q_H are increasing.

Suppose the salesman's utility function is parameterized by θ . Under what condition will the salesman spend more time in the first (known) shop if θ increases?

Suppose that the salesman has exponential utility. How does the condition you have derived relate to his risk-aversion coefficient?

10. In certain industries firms compete mainly by offering unique or superior services and/or by bundling related or ancillary goods with the products they manufacture and sell. Airline companies and PC vendors may be a relatively good example.

Assume that, in a given industry, goods must be sold (by government decree, say), at a price of p . A firm has constant marginal costs c and faces a demand function given by $Q(p,y)$, where y denotes the amount of a related good that the firm may decide to bundle (free of charge) with each unit of its own product. By incurring an additional unit cost of k , the firm can bundle $b(k)$ units of the ancillary good.

If $Q_1 < 0$ (demand is downward sloping in price) and $Q_2 > 0$ (demand increases with the amount of ancillary good bundled), provide and interpret a simple condition that ensures that more expensive goods will be bundled with higher amounts of the ancillary good.

4 Multidimensional Comparative Statics on Product Sets

Economic models of maximizing behavior frequently involve multiple choice variables. In classical price theory, for example, firms pick the quantities of several inputs to maximize profits and consumers choose consumption levels of several goods. More generally, consumers and firms make other kinds of choices as well. For example, a firm's decision makers choose not only input and output quantities but also technologies, locations, business partners and so on. Choice problems with multiple decision variables raise several new kinds of questions about comparative statics that were missing (or received only limited treatment) from our analysis of problems with a single decision variable in Part I.

The first of these is a robustness question: when there are multiple choice variables, under what conditions do we obtain robust comparative statics on any one of the elements of the choice? As with the robustness issues addressed in earlier chapters, this question has roots in traditional price theory. For example, in textbook treatments of price theory, when one investigates whether two inputs – say labor and capital – are “complements” or “substitutes,” one usually seeks the weakest condition on the production function which implies that which an increase of the price of one input reduces or increases the demand for the second input *regardless of the second input's price* (at least within some range). When the firm's choices are not limited to be simply quantities of perfectly divisible inputs and outputs purchased on competitive markets, the corresponding analysis would specify that the firm chooses levels of x , y and z to maximize $f(x, y, z) + g(x) + h(y) + k(z, \theta)$ subject to $(x, y, z) \in S$, where f is the production function and g , h and k are the negatives of the cost functions for the three inputs. We then seek the weakest conditions on f and k which imply that $x^*(\theta; g, h, k)$ is nondecreasing or nonincreasing in θ for all functions g and h (and perhaps all sets S) in some suitably large class. We showed in Chapter 2 that such problems could sometimes be analyzed using the methods of that chapter by reformulating the problem as a sequence of univariate problems, but the argument was rather complicated and was limited in scope. In this chapter we will analyze multivariate choice directly, focusing on the additional issues which arise in the case of multiple inputs and obtaining more direct results.

A second kind of question that arises only in multivariate problems concerns *comovements* of choice variables: we seek the weakest conditions under which we can draw a robust conclusion that all the components of a maximum move together in the same direction in response to certain changes in exogenous parameters. The key condition for robust conclusions of comovement is complementarity among the choice elements. In the classical theory of the firm, an increase in demand or a fall in an input price will lead the firm to increase its purchases of land, capital, and labor, provided these are complements in the sense that a fall in the price of any one of them increases the demand for each of the others. Suppose, however, that the firm's choices are more complicated. For example, it might choose the intensity of the communications it has with its customers regarding their future demands, the average level of finished goods inventories it will hold as buffers, and the range of product varieties it will offer. The question of comovements is answered by a characterization of complementarities that does not require the choice variables to be simple inputs purchased on competitive markets. In the example, if more communication, lower inventories and more variety are complementary (as they are according to one recent

analysis¹), then a fall in the costs of communication would lead to a systematic comovement in the choices, with more communication, lower inventories and more variety being offered.

Third, there are questions about how particular multivariate structures can be mapped into univariate ones to allow application of the results of Chapters 2 and 3. In our study of producer theory in Chapter 2, we analyzed an objective of the form $f(x, y, \theta) + g(x) + h(y)$ by treating the problem as a two-stage optimization problem, letting $\hat{f}(x, \theta) = \max_y f(x, y, \theta) + h(y)$ and then studying the resulting value function, $\hat{f}(x, \theta) + g(x)$. The multivariate theory provides some hints about when the value function will satisfy useful properties (such as increasing differences).

The final kind of question concerns multiplier effects, like those represented by the LeChatelier principle investigated in Chapter 2. One asks how the introduction of additional choice variables affects the magnitude of the response to a changing parameter and about the robustness of any such conclusion.

Provided that we frame our answers using appropriate concepts, the answers to these questions in general settings prove to be identical to the answers from price theory, and even to the answers from much more specific price theoretic models in which functional forms are narrowly prescribed. To develop the relevant intuition, we begin with such an example that employs a highly specific and tractable functional form.

Consider a profit-maximizing firm that makes two varieties of a single product and whose costs are a convex quadratic function of the two output levels x and y . A price premium of Δ is earned on one of the varieties. The firm's profits are given by

$$px + (p + \Delta)y - C(x, y),$$

$$\text{where } C(x, y) = \frac{1}{2}\alpha x^2 + \frac{1}{2}\beta y^2 + \gamma xy.$$

Convexity of costs requires that $\alpha > 0$, $\beta > 0$, and $\alpha\beta - \gamma^2 > 0$.

The first comparative statics question we shall ask is: What is the effect on x^* of a change in the price p ? Intuitively, the direct effect of an increase in the price p is to encourage an increase in the output of each variety, and hence in x . However, increasing p also has an indirect effect: it leads to a change in y , which affects the costs of producing the quantity x of the first variety. If we wish to allow $C(x, y)$ to be *any* member of the class of quadratic cost functions, what are the “robust” conclusions about how a change in the price p affects the optimal level of x ?

In this problem, adding a quadratic function $h(y)$ to the objective is equivalent changing β and Δ . So asking that the comparative statics result be robust to changes in a quadratic cost function amounts to asking that the comparative statics conclusion about x does not depend on those two parameters.

¹ Milgrom and Roberts (1988). A similar analysis using a different inventory model is offered in section 4.1 of this chapter.

The functional forms we have chosen make this problem easy to solve in closed form: the optimal choices for x and y are $x^* = \frac{-\gamma\Delta + (\beta - \gamma)p}{\alpha\beta - \gamma^2}$ and $y^* = \frac{\alpha\Delta + (\alpha - \gamma)p}{\alpha\beta - \gamma^2}$. We observe that x^* is a nondecreasing function of the price parameter p for all values of Δ and all $\beta > 0$ if and only if $\gamma \leq 0$.² Indeed, when this condition is satisfied, both x^* and y^* are nondecreasing functions of p on the whole range where costs are convex and so the expressions of this paragraph are valid.

Intuitively, the robustness condition arises for the following reason. An increase in the price tends to have the “direct” effect of increasing both x^* and y^* by raising the marginal benefit of each choice. When $\gamma \leq 0$, the indirect effect works in the same direction: any increase in y reduces the marginal cost C_x and so encourages a further increase in x . Similarly, any increase in x reduces the marginal cost C_y and so encourages a further increase in y . The direct and indirect effects are thus mutually reinforcing, each pointing to increases in x and y . Consequently, the relative magnitudes of the effects do not matter. In contrast, if $\gamma > 0$, then the indirect effect is in opposition to the direct, and the relative magnitudes do matter. As a result, no robust conclusion is possible: the two choice variables may move together for some parameter values and in opposite directions for others.

What we wish to emphasize is that, unlike the formal argument used to derive x^* and y^* , the general intuition does not rely either on the assumption that costs are quadratic or even on the assumption that the cost function is convex. What it does rely on is the condition that $C_{xy} \leq 0$, that is, that the cost function has decreasing differences. This is also the condition that defines what it means for two products to be “cost complements.” For it is “complementarity” that causes the direct and indirect effects to line up, resulting in unambiguous results. We shall later see that, indeed, the intuition can be matched with a formal argument establishing that decreasing differences is a necessary and sufficient condition for this kind of robust comparative static result.

Notice, too, that the condition that $C_{xy} \leq 0$ is equivalent in the quadratic case to the condition that $\gamma \leq 0$: The same “decreasing differences” condition that is necessary and sufficient when we allow general functions h to be added to the objective is also necessary and sufficient in this case where only quadratic functions that preserve convexity are considered. One useful lesson of our analysis will be to demonstrate that this irrelevance of the quadratic form of the cost function is no coincidence: one can frequently identify necessary and sufficient conditions for robust comparative statics that are still necessary even after one has imposed severe restrictions on the functional forms of the objective.

The same condition that $C_{xy} \leq 0$ also provides answers to our second and third questions about robust comparative statics. In response to the second kind of question, because the condition identifies the two product varieties as “cost complements,” the optimal output levels rise together with an increase in the price of both, providing a robust prediction about comovements in response to a change in the price p . In response to the third kind, let us create the dynamic

² There is no optimal value for x if $\beta \leq 0$.

programming value function by maximizing the original objective with respect to y for each fixed value of x . That exercise leads to the following:

$$\hat{f}(x, p) = \max_y px + py - C(x, y) = \frac{1}{2} \frac{p^2}{\beta} + (1 - \frac{\gamma}{\beta})px - \frac{1}{2}(\alpha - \frac{\gamma^2}{\beta})x^2.$$

Notice that this value function has increasing differences in p and x if $\gamma \leq 0$. This is useful information even if one does not require robustness with respect to β . If one wants the conclusion to hold for all $\beta > 0$, then the condition that $\gamma \leq 0$ is necessary as well. In this example, as in the general theory, the same condition arises repeatedly as the critical sufficient condition for the kind of robust comparative statics we shall be seeking.

The last kind of question – about multiplier effects and the LeChatelier principle – does not impose any restrictions on the parameter in this example. The case studied here is a special case of the two variable conclusion already treated in Chapter 2, since the functional form is such that either $C_{xy} \geq 0$ or $C_{xy} \leq 0$ for all x, y (depending on the sign of γ). At the end of this chapter, we will show how the analysis of Chapter 2 generalizes to the multivariate case.

In what follows, we first study optimization problems that are unconstrained or for which the constraint set is a product space, such as \mathbb{R}^N . We show that a critical sufficient condition for certain comparative statics conclusions will be that the objective function is “supermodular,” that is, that it satisfies increasing differences in each pair of choices, and also between each choice and the exogenous parameter. We then generalize the analysis to include comparative statics on sets of optimizers, and we also analyze constrained optimization. In Chapter 5, we analyze more general problems, where the choice set might include objects which are not real numbers, such as sets.

4.1 Supermodularity: A Sufficient Condition for Comparative Statics with a Unique Optimum

In the next two sections, we will make several simplifying assumptions to keep the analysis simple and to introduce the main ideas without getting bogged down in notation and definitions. In particular, we study unconstrained maximization problems in which the agent chooses from a “product set” such as \mathbb{R}^N , and in which there exists a unique optimum. Although the generalization to constrained optimization will require some additional work, the assumption of that the optimizer is unique does not affect our results. In particular, once the relevant definitions are in place to allow treatment of sets of optimizers, we will find that *all* of the theorems developed in these two sections apply equally well when there are multiple optima.

To describe product sets and their elements more precisely, we need a little bit of notation. With this notation, we will then extend our definition of increasing differences to multivariate functions and state our first multivariate comparative statics theorem.

Definition: A set X is called a product set if there exist sets X_1, \dots, X_N such that $X = X_1 \times \dots \times X_N$. X is a product set in \mathbb{R}^N if $X_n \subseteq \mathbb{R}$, $n = 1, \dots, N$.

The unit square in \mathbb{R}^2 , $[0,1] \times [0,1]$, is an example of a product set. Note that we do not require that X_n be an interval. In particular, for example, X_n might be some subset of the integers, representing indivisibility of some choice variable. The crucial property of X is that the possible values of any component of a choice vector \mathbf{x} be independent of the levels of the other components: If $\mathbf{x} \in X$ and $\mathbf{y} = (x_1, \dots, x_n + \Delta, \dots, x_N)$, then $\mathbf{y} \in X$ if and only if each component lies in its respective domain, that is, if $x_n + \Delta \in X$.

All the product sets we will consider in this Chapter will be ones where the component sets X_n are *chains*, that is, there is an order \geq_n on each X_n such that for each x' and x'' in X_n , either $x' \geq_n x''$ or $x'' \geq_n x'$. These orders then induce a (partial) order \geq over the product set X according to which $\mathbf{x}' \geq \mathbf{x}''$ if and only if $x'_n \geq x''_n$ for all n . For example, the familiar partial order over \mathbb{R}^N is obtained this way. This induced order over \mathbb{R}^N is not a complete order: for example, $(1,3)$ is neither greater nor less than $(2,2)$ in this order. For pairs that are ranked, when $\mathbf{x}' \geq \mathbf{x}''$, we will say that \mathbf{x}' is “larger than” or “bigger than” or “higher than” \mathbf{x}'' . We write $\mathbf{x}' > \mathbf{x}''$ if and only if $\mathbf{x}' \geq \mathbf{x}''$ and $\mathbf{x}' \neq \mathbf{x}''$, and $\mathbf{x}' \gg \mathbf{x}''$ if and only if $x'_n > x''_n$ for all n .

Recall the definition of increasing differences from Chapter 2. This definition involved functions of two arguments; however, it is easily adapted to a multi-dimensional setting by requiring that the relevant conditions be satisfied for a pair of variables, holding fixed all other arguments.

Definition: Let X be a product set, and let $f: X \rightarrow \mathbb{R}$. Then f has **increasing differences** in $(x_n; x_m)$, $n \neq m$, if and only if for all $x_n' \in X_n$ and $x_n'' \in X_n$ with $x_n' > x_n''$, and for all x_j , $j \neq n, m$, $f(x_1, \dots, x_n', \dots, x_N) - f(x_1, \dots, x_n'', \dots, x_N)$ is nondecreasing in x_m .

The increasing difference property for functions on \mathbb{R}^N is defined for pairs of variables and therefore has the same characteristics as those we already identified chapter 1. First, the condition is symmetric: If f has increasing differences in $(x_n; x_m)$ then it also has increasing differences in $(x_m; x_n)$. If f is differentiable in x_n , then f has increasing differences in $(x_n; x_m)$ if and only if $\frac{\partial}{\partial x_n} f(\cdot)$ is a nondecreasing function of x_m , and if f is twice continuously differentiable, then f has increasing differences in $(x_n; x_m)$ if and only if $\frac{\partial^2}{\partial x_n \partial x_m} f \geq 0$. If both f and g have increasing differences in $(x_n; x_m)$, then $f+g$ has increasing differences as well.

Intuitively, increasing differences in $(x_n; x_m)$ means that raising the level of x_m weakly increases the return to raising x_n . As we discussed in Chapter 2, if F is a smooth production function, then increasing differences in inputs x_n and x_m means that the marginal product of each of these inputs is a nondecreasing function of the level of the other.

We use the word *supermodular* to describe a situation when a function has increasing differences in each pair of variables.

³ Supermodularity is also sometimes called “Edgeworth complementarity” because Edgeworth used the same condition in his treatment of consumer theory.

*Definition: If f has increasing differences in $(x_n; x_m)$, for every $n \neq m$, then the function f is supermodular.*⁴

In Chapter 5, we will generalize the concept of supermodularity to functions defined on sets X which are *not* product sets. For the purposes of this chapter, the term will simply be a convenient way to describe functions which are characterized by complementarity between each pair of variables.

We are now ready to state our first multivariate comparative statics theorem, which concerns the following problem (where $f: X \times \Theta \rightarrow \mathbb{R}$, X is a product set, and $\Theta \subset \mathbb{R}$):

$$\max_{\mathbf{x} \in X} f(\mathbf{x}, \theta) \quad (4.1)$$

Denote the set of solutions of (4.1) by $X^*(\theta)$. In this section and the next, we make the following simplifying assumption:

$$\text{For all } \theta \in \Theta, X^*(\theta) \text{ has exactly one element.} \quad (\text{A4.1})$$

Since we will show in Section 4.3 that the results of these sections hold even without (A4.1) once we develop a suitable ordering over sets, we omit this condition from the statement of our theorems. The following theorem gives a sufficient condition for monotone comparative statics in an unconstrained problem.

Theorem 4.1: Consider problem (4.1). Suppose that f is supermodular. Then $X^(\theta)$ is nondecreasing in θ .*

This theorem is a corollary of Theorem 4.4 in Section 4.3, so we will not provide a separate proof. An intuition for the result is suggested by the example at the beginning of the Chapter: the increase in θ (weakly) raises the return to increasing each of the choice variables x_n , because there are increasing differences with respect to $(x_n; \theta)$. Thus, the direct effect is to increase each of the choice variables. Further, raising any of the choice variables (weakly) increases the return to raising each of the others, so that the “indirect” or “cross” effects all reinforce the direct effect.

To make this intuition more concrete and to illustrate some techniques that are useful in applications, we offer the following examples.

4.1.1 Application: Flexible Manufacturing

One of the changes observed in the several major manufacturing industries (especially the auto industry) in the 1980s was a switch to more flexible manufacturing methods and numerous other accompanying changes. The following model of some of those changes is a simplified version of one developed by Milgrom and Roberts (1990).

We focus attention on three decisions that a firm makes in organizing its production

⁴ Supermodularity is also sometimes called “Edgeworth complementarity” because Edgeworth used the same condition in his treatment of consumer theory.

operations: the number of varieties of its product (n), the average run size (x), and the flexibility of its equipment and methods (y). We omit from our analysis any effect that changing costs has on the firm's market share and scale of operation, which are affected by competitors' adjustments as well, focusing our modeling attention only on the choices outlined above. Thus we fix the firm's output level at q and assume that the mean price the firm receives for its products is $P(n)$, which depends on the number of varieties produced. The firm's revenues are therefore $qP(n)$. The cost it incurs in our model is the sum of several terms:

$$C(q) + hn \frac{x}{2} + \frac{q}{x} s(y) + K(y, \theta).$$

The $C(q)$ term is the direct cost of production in each period. The next two terms are derived from the “*economic order quantity*” inventory model. A firm that sells a total volume of q units and produces lots (each of a single variety) of size x maintains an average inventory (per variety) of $x/2$ and has q/x set-ups, so its inventory holding costs are proportional to $nx/2$ and its set-up costs are proportional to q/x . In the cost formula, h and $s(y)$ are the holding costs per unit and the costs per set-up, respectively. The costs of set-up depend on the “flexibility” of equipment and methods y . Changes in this firm's decisions arise because we assume that the marginal cost of flexibility is falling, that is, K_y is decreasing in θ (and θ is increasing over time as technology improves).

In applying Theorem 4.1 and its close cousins it is extremely useful to redefine variables so that the directions of the anticipated changes line up. In this example, we expect flexibility and product variety to increase but run sizes and inventories to decrease, so we define a new variable $\hat{x} = -x$ and use it instead of x in the analysis. With this change, the problem is as follows:

$$\max_{n, \hat{x}, y} qP(n) - C(q) + hn \frac{\hat{x}}{2} + \frac{q}{\hat{x}} s(y) - K(y, \theta)$$

Regarding the objective as a function of the vector (\hat{x}, y, n, θ) , it is easy to check that each term in the sum is supermodular, so that the whole objective is supermodular. The only assumptions required for this conclusion are the definitional assumptions that K_y is decreasing in θ and that $s(\cdot)$ is nonincreasing. According to Theorem 4.1, we may conclude without further assumptions that the choice variables are nondecreasing functions of the parameter θ , that is, that a reduction in the marginal cost of flexibility leads to increases in flexibility, reductions in run size and inventory per product (and a concomitant increase in the number of set-ups), and an increase in product variety.

Another widely reported change in the 1980s was an increased rate of product innovation, which would cause the inventory holding cost h to rise due to increased costs of obsolescence. Regarding the objective as a function of (\hat{x}, y, n, h) , one can again check that the function is supermodular. The theorem implies that increases in h , like increases in θ , lead to increases in the optimal values of the choice variables. Later, we will see that the LeChatelier principle applies to such problems, so that, say, holding fixed the run sizes and product variety in the model leads to smaller changes in the optimal level of flexibility y than would be found when these

complementary changes are allowed.

Notice that the comparative statics conclusions we have just derived conclusion are “robust”: they do not require any assumption about the form or properties of the pricing function $P(\cdot)$ and they restrict $s(\cdot)$ only by the innocuous requirement that it be nonincreasing. Moreover, the conclusions apply even if the objective is not concave. This is especially important in this application because, omitting the revenue term about which no assumptions have been made, the remaining terms do not define a concave objective, regardless of the specification of $s(\cdot)$ and $K(\cdot, \cdot)$. (To verify that, note that by maximizing over \hat{x} with the other variables fixed, the objective can be reduced to $qP(n) - \sqrt{\frac{1}{2}nhqs(y)} - K(y, \theta)$.) Ignoring the revenue term, this maximand is strictly *convex* in n , which would not be possible if the original objective were concave in the choice variables.) Unlike the implicit function approach, the new approach does not require special attention to nonconvexities and even allows us to omit the unnatural assumption that the number of varieties is a continuous variable.

4.1.2 Application: Incentive Theory

A second application of Theorem 4.1 is drawn from principal-agent theory.⁵ In this example, we ask how the formal incentives provided to an agent change with factors such as the precision with which an agent’s output can be measured.

Assume that an agent exerts effort e_i on behalf of a principal in each of several tasks $i=1, \dots, N$. For example, a sales agent may seek out new customers, provide service to existing customers, and gather information about changing competitor strategies and new market opportunities for her employer, who is the principal. We suppose that the effort costs incurred by the agent are given by a strictly convex, quadratic function: $C(e_1, \dots, e_N)$. The principal pays the agent a salary plus bonuses or commissions of $\sum_i p_i x_i$. In this formula, $x_i = e_i + \varepsilon_i$ is the measured performance in task i : it is equal to the agent’s effort in the task plus a normally distributed error term with mean zero and variance $\sigma_i^2 = 1/\theta_i$. Also, p_i is the “piece rate” paid for performance in task i . The parameter θ_i represents the *precision* with which performance is measured.

With this specification, the variance of the agent’s total compensation is $\sum_i p_i^2 / \theta_i$. If the agent is risk averse with a constant coefficient of absolute risk aversion r , then she incurs a risk premium of $\frac{1}{2}r\sum_i p_i^2 / \theta_i$ on account of this variance. The total certainty equivalent payoff of the principal and agent together consists of the benefits enjoyed by the principal minus the costs incurred by the agent minus the agent’s risk premium:

$$TCE(\mathbf{e}, \mathbf{p}, \boldsymbol{\theta}) \equiv \sum_i b_i e_i - C(\mathbf{e}) - \frac{1}{2}r \sum_i p_i^2 / \theta_i .$$

A contract consists of a specification of the agent’s base salary, the vector piece rates \mathbf{p} , and a

⁵ The particular model given here is a reduced form based on the model in Holmstrom and Milgrom (1994).

suggested vector of efforts \mathbf{e} . In order to motivate the agent to follow the suggestion, the piece rates and effort levels need to be consistent. Since a self-interested agent would maximize its certainty equivalent, which is equal to $\sum_i p_i e_i - C(\mathbf{e}) - \frac{1}{2} r \sum_i p_i^2 / \theta_i$, a contract is feasible only if the pair (\mathbf{e}, \mathbf{p}) satisfies the *incentive constraint* that \mathbf{e} maximizes the agent's certainty equivalent. The incentive constraint is represented by the first order conditions $p_i = C_i(\mathbf{e})$ for $i=1,\dots,N$. The solution to the first order condition defines the *effort supply* function, $\mathbf{e}(\mathbf{p})$. Since the cost function is quadratic, the incentive constraint is a system of linear equations, so the effort supply function is also linear. We assume that the various kinds of effort are substitutes for the agent in the sense of price theory, that is, $\partial e_i / \partial p_j \leq 0$ for all $i \neq j$.

The interesting conclusion of this exercise is that, once one substitutes the supply function into the previous stated objective, the resulting new objective $\pi(\mathbf{p}, \boldsymbol{\theta}) \equiv \text{TCE}(\mathbf{e}(\mathbf{p}), \mathbf{p}, \boldsymbol{\theta})$ given below is supermodular:

$$\pi(\mathbf{p}, \boldsymbol{\theta}) = \sum_i b_i e_i(\mathbf{p}) - C(\mathbf{e}(\mathbf{p})) - \frac{1}{2} r \sum_i p_i^2 / \theta_i.$$

To verify supermodularity, we need to check that the mixed partial derivatives of π are non-negative. First, inspection of π verifies that increases in each θ_i (weakly) increase the marginal returns to p_j . To compute the mixed partial derivatives with respect to p_i and p_j , observe that:

$$\frac{\partial \pi}{\partial p_j} = \sum_i (b_i - C_i(\mathbf{e}(\mathbf{p}))) \frac{\partial e_i}{\partial p_j} - r p_j / \theta_j = \sum_i (b_i - p_i) \frac{\partial e_i}{\partial p_j} - r p_j / \theta_j,$$

where the last equality is obtained by substituting for C_i from the incentive constraint. Next, for $j \neq i$,

$$\frac{\partial^2 \pi}{\partial p_j \partial p_i} = -\frac{\partial e_i}{\partial p_j} + \frac{\partial^2 e_i}{\partial p_j \partial p_i} = -\frac{\partial e_i}{\partial p_j} \geq 0,$$

where the second equality is from the linearity of the supply function and the inequality follows by our assumption that efforts are substitutes in the supply function.

The conclusions of the analysis are that (1) in this model, when efforts towards different activities are substitutes in the effort supply function, piece rates for different activities are mutually complementary instruments, and (2) an increase in the precision of measuring any *one* kind of performance leads to an increase in the piece rates paid for *all* kinds of performance.

4.2 Comparison with the Implicit Functions Approach

Our analysis thus far has not included the assumption that the objective function is concave, or differentiable, or that the optimizers are on the interior of the domain, or that the domain is a convex set. So the results hold even when taking first-order conditions is impossible, or when their solutions do not correspond to (local) optima. We will show in the next section that, when suitably interpreted, they continue to hold when there are multiple optima for some relevant parameter values.

Still, it is interesting to compare the assumptions and conclusions of the above theorems with those involved in applying multidimensional implicit function methods. Suppose then that X is a convex set in \Re^N and that f is twice continuously differentiable on the interior of $X \times \Theta$. Suppose that for each value of θ there is a unique maximizer of $f(\mathbf{x}, \theta)$, denoted $\mathbf{x}^*(\theta)$, and that this optimum lies on the interior of X . Then the first-order conditions, which are

$$\begin{aligned}\frac{\partial}{\partial x_1} f(x_1, \dots, x_N, \theta) &= 0 \\ &\vdots \\ \frac{\partial}{\partial x_N} f(x_1, \dots, x_N, \theta) &= 0,\end{aligned}$$

must be met at $(\mathbf{x}^*(\theta), \theta)$ for each θ . That is, (letting subscripts of f denote partial derivatives), $h_n(\theta) \equiv f_n(\mathbf{x}^*(\theta), \theta) = 0$ for all θ and all $n = 1, \dots, N$. Assume further that $\mathbf{x}^*(\cdot)$ is differentiable. Then, for each θ , we can calculate the derivative of h_n , $n = 1, \dots, N$, as

$$\frac{dh_n}{d\theta} = \frac{\partial f_n}{\partial x_1} \frac{\partial x_1}{\partial \theta} + \dots + \frac{\partial f_n}{\partial x_N} \frac{\partial x_N}{\partial \theta} + \frac{\partial f_n}{\partial \theta} = f_{n1} \frac{\partial x_1}{\partial \theta} + \dots + f_{nN} \frac{\partial x_N}{\partial \theta} + f_{n\theta}$$

and, when evaluated at $(\mathbf{x}^*(\theta), \theta)$, these expressions all must be equal to zero. This gives a system of N equations in $N+1$ variables. Further, the second order condition requires that, in a neighborhood of $\mathbf{x}^*(\theta)$, the $N \times N$ matrix

$$[f_{ij}] = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]$$

must be negative semidefinite. For simplicity, let us make the standard regularity assumption that the matrix is actually negative definite. Then, in particular, the matrix of second derivatives is nonsingular, and so we can solve for the vector of derivatives of the \mathbf{x}^* function:

$$\begin{bmatrix} \frac{\partial \mathbf{x}_1^*}{\partial \theta} \\ \dots \\ \frac{\partial \mathbf{x}_n^*}{\partial \theta} \\ \dots \\ \frac{\partial \mathbf{x}_N^*}{\partial \theta} \end{bmatrix} = \begin{bmatrix} f_{11} & \dots & f_{1N} \\ \dots & f_{nn} & \dots \\ f_{N1} & \dots & f_{NN} \end{bmatrix}^{-1} \begin{bmatrix} -f_{1\theta} \\ \dots \\ -f_{n\theta} \\ \dots \\ -f_{N\theta} \end{bmatrix}$$

We can then try to determine the signs of the elements of the vector ($\frac{\partial}{\partial \theta} \partial x_n^* / \partial \theta$) using what we know about the matrix $[f_{ij}]$ and the vector $(-f_{n\theta})$ and any other conditions we may invoke.

For example, for $N = 2$, the formulae for the derivatives of \mathbf{x}^* are

$$\begin{aligned}\frac{\partial x_1^*}{\partial \theta} &= \frac{-f_{1\theta} f_{22} + f_{2\theta} f_{12}}{f_{11} f_{22} - f_{12} f_{21}} \\ \frac{\partial x_2^*}{\partial \theta} &= \frac{-f_{2\theta} f_{11} + f_{1\theta} f_{21}}{f_{11} f_{22} - f_{12} f_{21}}\end{aligned}$$

where the expressions are evaluated at $(\mathbf{x}^*(\theta), \theta)$. To derive comparative statics results, we want to know when both are non-negative, so that the solution is nondecreasing.

From the regularity assumption, we know that the denominator in these expressions is positive and that f_{11} and f_{22} are negative. As well, assuming f is twice continuously differentiable, we know that $f_{12} = f_{21}$. Assuming that $f_{1\theta}$ is non-negative is then sufficient to ensure that the first term in the numerator of the first expression is non-negative. Then it is sufficient for the optimal value of x_1 to be nondecreasing in θ that the second term in the numerator be non-negative, which means that $f_{2\theta}$ and f_{12} have to be of the same sign. Similarly, if $f_{2\theta}$ is non-negative, then it is sufficient for x_2^* to be non-decreasing that $f_{1\theta}$ and f_{21} be of the same sign. Both these can hold then only if $f_{12} = f_{21}$ is non-negative. Thus, unsurprisingly, the smooth analog of nondecreasing differences is also sufficient in this case.

Although we will not pursue it further here, it is possible to mimic our analysis of Chapter 2, comparing the implicit functions approach to the monotone methods approach as they apply to the various contexts we study in this Chapter. As we found in Chapter 2, it turns out that the additional assumptions required for the application of the implicit function theorem will not allow us to weaken the *critical* sufficient conditions we have identified.

4.3 Critical Sufficient Conditions

To derive critical sufficient conditions, we will want to consider problems in which the choice may be constrained. Formally, we constrain the choice to lie in some subset S of the product set X , and further we treat S as a parameter of the problem. There are potentially many ways to consider robustness in the multidimensional case, and we will discuss several. For simplicity, we first study contexts which correspond in form to those of classical producer theory but which are generalized enough to have broader application. Throughout this chapter, we restrict ourselves to problems in which the constraint set S is a product set and ask that our results be robust to variations in S : That is, we treat S as one of the factors which might vary across our context of interest.

In the classical producer theory and related applications, one is interested in models in which the individual returns to the choice variables have individual parameters. In producer theory,

these parameters are the input prices. This structure is captured more generally by the following problem:

$$\max_{\mathbf{x} \in S} f(\mathbf{x}) + \sum_n g_n(x_n, \theta_n) \quad (4.2)$$

We denote the optimal choice of \mathbf{x} in (4.2) by $X^*(\Theta, S)$. In the interpretation suggested by the classical theory of the firm, \mathbf{x} is an input vector, $f(\mathbf{x})$ represents the net revenues (i.e., in the competitive model, $f(\mathbf{x}) = p \cdot F(\mathbf{x})$, where F is a production function) and the function g_n gives the costs (expressed as a negative number) of using x_n units of input n , where these costs depend on a parameter θ_n . We are interested in how a change in one of the exogenous parameters, say θ_n , affects the whole vector of choices.

Theorem 4.1 applies directly to this problem. Thus, if each g_n satisfies increasing differences and f is supermodular (so that the marginal product of each input is an increasing function of the level of use of each other input), then the optimal choice of \mathbf{x} will be nondecreasing in θ . In the theory of the firm interpretation, g_n having increasing differences in (x_n, θ_n) means that an increase in θ_n raises the incremental cost incurred in increasing x_n . For example, if x_n is purchased in a competitive market at price $p_n = -\theta_n$, then $g_n(x_n, \theta_n) = \theta_n \cdot x_n$. Thus, if all the inputs are pair-wise complements, and if the price of one input falls, the firm will respond with higher levels of all inputs.

However, the requirement that *every pair* of choices satisfy the increasing differences condition may seem quite strong and one may wish to ask whether there is some weaker sufficient condition that leads to the same monotonicity conclusion. We begin to answer this question by specifying a context throughout which we wish the comparative statics result to be valid. First, we would like to include the case where inputs are purchased on a competitive market, and since we want to ask questions about what happens when prices vary, we include all possible input prices (although, as we will show later, we can restrict ourselves to the case where the production function is nondecreasing and the prices are positive). Second, we wish to allow the constraint set S to be any product set. This allows us to consider cases where inputs are available only in discrete increments, as well as cases where the inputs are continuous. It also allows us to bound the use of some inputs, to incorporate limits in the availability of a given resource.

The following result gives the critical sufficient conditions for monotonicity of the optimal choice of \mathbf{x} in θ_n , in the context just described. It establishes that even when the g_n functions are restricted to be of a linear form, as they are in price theory applications, supermodularity is a critical sufficient condition.

Theorem 4.2: Assume the following about problem (4.2):

(S4.2) X and $S \subseteq X$ are product sets.

(F4.2) No restrictions on $f: X \times \Theta \rightarrow \mathbb{R}$ ($\Theta \subset \mathbb{R}$).

(G4.2) Each $g_n: \mathbb{R}^2 \rightarrow \mathbb{R}$, $n = 1, \dots, N$, has the form $g_n(x_n; \theta_n) = x_n \cdot \theta_n$.

Then $X^*(\boldsymbol{\theta}, S)$ is nondecreasing in θ_n for all $n=1,..,N$, all S and all g_n , if and only if f is supermodular.

Proof: The direct part of the theorem is a corollary of Theorem 4.1, since the objective function in (4.2) is supermodular if f is. It suffices, therefore, to focus on the reverse implication.

Suppose that f fails increasing differences in (x_m, x_n) . Then there exists a product set $S = \{(x_m'', x_n'', \mathbf{x}_{mn}), (x_m', x_n', \mathbf{x}_{mn}), (x_m'', x_n', \mathbf{x}_{mn}), (x_m'', x_n', \mathbf{x}_{mn})\}$ where $x_m'' > x_m'$ and $x_n'' > x_n'$, and the following inequality holds (suppressing the \mathbf{x}_{mn} in the notation):

$$\Delta_i \Delta_j f \equiv f(x_i, x_j) + f(x'_i, x'_j) - f(x'_i, x_j) - f(x_i, x'_j) < 0$$

Let $f_{HH} = f(x_m'', x_n'')$, $f_{LL} = f(x_m', x_n')$, $f_{HL} = f(x_m'', x_n')$, and $f_{LH} = f(x_m', x_n'')$. Further, let $\Delta x_m = x_m'' - x_m'$, and $\Delta x_n = x_n'' - x_n'$. Then define $g_i(x_i, \theta_i) = x_i \cdot \theta_i$ for $i=m$, and the following parameters:

$$\theta_n = \frac{1}{2\Delta x_n} [f_{LL} + f_{HL} - f_{HH} - f_{LH}]$$

Pick any $\theta_m'' > \theta_m'$ such that

$$\theta'_m < \frac{1}{2\Delta x_m} [f_{LL} + f_{LH} - f_{HH} - f_{HL}] < \theta''_m$$

Now, let $\boldsymbol{\theta}_L = (\theta_m', \theta_n)$ and $\boldsymbol{\theta}_H = (\theta'_m, \theta_n)$. It is straightforward to verify that $X^*(\boldsymbol{\theta}_L, S) = \{(x_m'', x_n')\}$, while $X^*(\boldsymbol{\theta}_H, S) = \{(x_m', x_n'')\}$, violating hypothesis (i).

The theorem asserts first that if f is supermodular, then the whole vector of choice variables increases as the parameter goes up, *regardless of the constraint set S* (provided that it is a product set). In the application to the theory of the firm, this is analogous to the condition that the inputs being chosen are complements, even if the inputs are indivisible or subject to other individual restrictions on the quantities that can be used. The theorem further asserts that if we want this conclusion to be robust with respect to variations in the set S , then no weaker condition than supermodularity will do.

A disadvantage of Theorem 4.2 for applications in price theory is that it requires that we consider *any* prices for the inputs, including negative ones. This is no problem for the part of the theorem that asserts that supermodularity is sufficient for monotonicity. It is problematic for the reverse conclusion, however, because a condition that is necessary only if we insist on robustness to negative input prices is not of great economic interest. For this reason, the following modification of Theorem 4.7 is useful.

Theorem 4.3: Assume the following about problem (4.2):

(S4.3) X and $S \subseteq X$ are product sets.

(F4.3) $f: X \rightarrow \mathbb{R}$ is nondecreasing in \mathbf{x} .

(G4.3) Each $g_n: \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$, $n = 1, \dots, N$, has the form $g_n(x_n; \theta_n) = x_n \cdot \theta_n$, where $\theta_n \leq 0$.

Then $X^*(\theta, S)$ is nondecreasing in θ_n for all $n = 1, \dots, N$, all S satisfying (S4.3), and all g_n satisfying (G4.3), if and only if f is supermodular.

Proof: If f is nondecreasing, the parameters defined in the proof of Theorem 4.2 are negative. Thus, they generate the required counterexample.

Thus, if inputs are always productive, supermodularity of the “production function” is a critical sufficient condition for all the inputs to increase when the (non-negative) competitive price of any one of them falls. Still, it is important to realize that both Theorem 4.2 and Theorem 4.3 rely on the inclusion of the quantifier “for all product sets S ” in the comparative statics conclusion. This quantifier ensures that all possible input combinations are potentially relevant to the firm, which is important for the argument that f must be supermodular over its entire domain, X .

Now consider the case of a firm that is a perfect competitor in its output market but possibly a monopsonist in input markets. Denote the vector of inputs by \mathbf{x} , the output price by p , the production function by $F(\mathbf{x})$, and the cost of each input by $h_n(x_n, \theta_n)$, where θ_n is a parameter reflecting the supply conditions for input n . Define $f(\mathbf{x}) \equiv pF(\mathbf{x})$ and $g_n(x_n, \theta_n) = -h_n(x_n, \theta_n)$. Then (4.2) is the expression for the firm’s profits. We are again interested in critical sufficient conditions for use of all the inputs to increase with a change in one of the parameters θ_n . Suppose that an increase in θ_n decreases the cost of increasing x_n : formally, this is represented by the assumption that g_n has increasing differences. (Note we can handle the case where increasing θ_n always decreases the cost of increasing x_n by replacing the parameter by its negative or inverse. Thus, the substance of the assumption of increasing differences is that the effect of changing the parameter is uniform, always increasing or always decreasing the incremental cost of the input.) As the following generalization of Theorem 4.3 shows, the critical sufficient condition for X^* to be nondecreasing in each θ_n is the same as in the competitive case, namely, supermodularity of f . So the critical sufficient condition is the same in both the special linear case of competitive input markets and in the general case.

As in Chapter 2, however, the family of functions from which each g_n is drawn must be suitably rich, otherwise supermodularity will not be a necessary condition. For example, if the marginal cost of one input never varies, there may be values of that input which are never optimal, and so behavior of the function f will be irrelevant for those values. It turns out that, just as in Chapter 2, the relevant notion of richness is the full one-parameter family (thus, for any potential value, a , of the marginal cost or benefit to an individual choice x_n , there is some member of the family of functions, call it h_t , such that $h_t'(x)=a$). The families of functions described by (G4.2) are full one-parameter families; thus, any set which contains all linear functions will also be a full one-parameter family.

Theorem 4.4: Assume the following about problem (4.2):

(S4.4) X and $S \subseteq X$ are product sets.

(F4.4) No restrictions on $f: X \rightarrow \mathfrak{R}$.

(G4.4) For $n = 1, \dots, N$, $G^n = \{g_{nt}\}$ is a family of functions so that, for each $g_{nt} \in G^n$, $g_{nt}: \mathfrak{R}^2 \rightarrow \mathfrak{R}$ and $g_{nt}(x_n; \theta_n)$ has increasing differences in $(x_n; \theta_n)$. Further, G^n is a full one-parameter family.

Then $X^*(\Theta, S)$ is nondecreasing in θ_n for all S and all $g_{nt} \in G^n$, ($n=1, \dots, N$), if and only if f is supermodular.

Clearly, condition (G4.4) is satisfied when each G^n contains all functions that satisfy increasing differences. The condition allows for smaller families of functions as well, so long as they are rich enough to be full one-parameter families. Note that if f is nondecreasing — a natural assumption in many economic problems — then (analogously with Theorem 4.3) we can limit consideration to cost functions that are also nondecreasing.

Theorem 4.5: Assume the following about problem (4.2):

(S4.5) X and $S \subseteq X$ are product sets.

(F4.5) $f: X \rightarrow \mathfrak{R}$ is nondecreasing.

(G4.5) For $n = 1, \dots, N$, $G^n = \{g_{nt}\}$ is a family of functions so that, for each $g_{nt} \in G^n$, $g_{nt}: \mathfrak{R}^2 \rightarrow \mathfrak{R}$ and $g_{nt}(x_n; \theta_n)$ is nonincreasing in x_n and has increasing differences in $(x_n; \theta_n)$. Further, G^n is a negative one-parameter family.

Then $X^*(\Theta, S)$ is nondecreasing in θ_n for all S and all $g_{nt} \in G^n$, ($n=1, \dots, N$), if and only if f is supermodular.

As in Chapters 2 and 3, it is possible to combine Theorem 4.2 and 4.4 to show that a comparative statics conclusion from a problem with a lot of additional structure implies a result in a much more general class of problems.

Theorem 4.6: Assume the following about problem (4.2):

(S4.6) X and $S \subseteq X$ are product sets.

(F4.6) No restrictions on $f: X \rightarrow \mathfrak{R}$.

Consider $G = \{G^1, \dots, G^N\}$ such that each G^n satisfies (G4.2), and $\hat{G} = \{\hat{G}^1, \dots, \hat{G}^N\}$ such that each \hat{G}^n satisfies (G4.4). Then $X^*(\Theta, S)$ is nondecreasing in θ_n , all S , and all $g_{nt} \in G^n$, ($n=1, \dots, N$), if and only if $X^*(\Theta, S)$ is nondecreasing in θ_n , all S , and all $g_{nt} \in \hat{G}^n$, ($n=1, \dots, N$).

Thus, if we show robust monotone comparative statics for a family of models defined by linear cost functions, we have the same result for a much more general case. In the case of the comparative statics of producer theory, any qualitative comparative statics prediction which has been proved for the case of perfectly competitive input markets generalizes immediately to the case of nonlinear input cost functions. Clearly, the same sort of result is possible combining

Theorems 4.4 and 4.6.

4.4 Maximization Problems with Multiple Optima

Thus far, we have maintained the assumption that there is a unique optimum at each relevant value of the parameter. Ensuring uniqueness typically means imposing conditions that are economically unwarranted. For example, assuming strict concavity of the objective function involves assuming that each choice variable is divisible (so that the choice set is convex) and also rules out any form of increasing returns. Despite the inherent unattractiveness of such assumptions, they have been commonly made in economic analyses. One frequently used justification is a claim that uniqueness is needed in order to draw comparative statics conclusions. We have already seen in the single dimensional case that this need not be so: Comparative statics results can be obtained when there are multiple optima by looking at the behavior of the set of optima, and in fact there is little connection between uniqueness of the solution to an optimization problem and monotonicity. The situation is similar when there are multiple choice variables. In parallel to Chapter 2, we will again extend the usual order on the choice space – in this case \mathbb{R}^N – to an order on subsets of the space and then use that definition to characterize monotonicity of the set of optimizers.

In order to state and prove the theorems in this section, it will be useful to define increasing differences between two groups of variables. To describe increasing differences for a function $f(\mathbf{x})$ in two subsets of the components of \mathbf{x} , we need more notation. Given \mathbf{x}, \mathbf{y} and \mathbf{z} in \mathbb{R}^n , let I and J be disjoint subsets of indices in $\{1, \dots, n\}$, and let K be the set of all remaining indices, that is, $K \equiv \{1, \dots, n\} \setminus (I \cup J)$. Then we denote by $(\mathbf{x}_I, \mathbf{y}_J, \mathbf{z}_K)$ the vector \mathbf{w} whose n -th element is $w_n = x_n$ if $n \in I$, $w_n = y_n$ if $n \in J$ and $w_n = z_n$ if $n \in K$.

Definition: Let $f: X_1 \times \dots \times X_N \rightarrow \mathbb{R}$, let I and J be disjoint subsets of $\{1, \dots, N\}$. Then f has increasing differences in $(\mathbf{x}_I; \mathbf{x}_J)$ if and only if f has increasing differences in $(x_i; x_j)$ for all $i \in I$ and all $j \in J$.

So, given a function $f(\mathbf{x}; \theta)$, we formalize the idea that the parameter θ increases the returns to each component of \mathbf{x} by saying that f has increasing differences in $(\mathbf{x}; \theta)$. Further, if we wish to indicate that each element of $\{x_1, x_3, x_5\}$ increases the returns to each element of $\{x_2, x_4, x_6\}$, we let $I = \{1, 3, 5\}$, let $J = \{2, 4, 6\}$, and say that f has increasing differences in $(\mathbf{x}_I; \mathbf{x}_J)$.

It will also be convenient to define notation for the (component-wise) maximum of two vectors and the (component-wise) minimum of two vectors. For any $\mathbf{x}' \in \mathbb{R}^n$ and $\mathbf{x}'' \in \mathbb{R}^n$, let $\max\{\mathbf{x}', \mathbf{x}''\} \equiv (\max\{x_1', x_1''\}, \dots, \max\{x_N', x_N''\})$ and $\min\{\mathbf{x}', \mathbf{x}''\} \equiv (\min\{x_1', x_1''\}, \dots, \min\{x_N', x_N''\})$. These are just the vectors whose components are the element-by-element maxima and minima. This construction is illustrated in Figure 4.1. Note that if neither $\mathbf{x}' \geq \mathbf{x}''$ nor $\mathbf{x}' \leq \mathbf{x}''$, so that the two points are not ordered with respect to one another, then $\max\{\mathbf{x}', \mathbf{x}''\} > \mathbf{x}' > \min\{\mathbf{x}', \mathbf{x}''\}$ and $\max\{\mathbf{x}', \mathbf{x}''\} > \mathbf{x}'' > \min\{\mathbf{x}', \mathbf{x}''\}$. That is, the component-wise maximum of two vectors lies strictly above each vector, and the component-wise minimum lies strictly below both vectors.

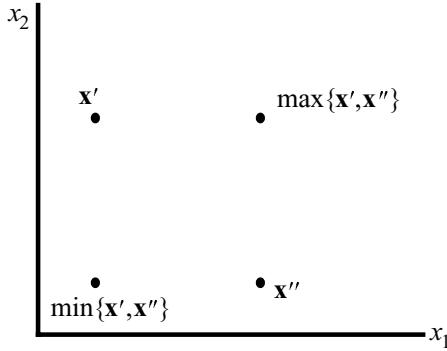


Figure 4 .1: The vectors \mathbf{x}' and \mathbf{x}'' lie between the vectors $\max\{\mathbf{x}', \mathbf{x}''\}$ and $\min\{\mathbf{x}', \mathbf{x}''\}$ (in the componentwise order).

In our treatment of comparative statics with multiple optima and one dimensional choice variables, we sometimes emphasized the comparative statics of the largest and smallest maximizers. However, in multiple dimensions, a set of maximizers might consist of two vectors which are not ordered, so that neither is larger than the other (such as \mathbf{x}' and \mathbf{x}'' in Figure 4.1). Together, the next two results give conditions under which a set of maximizers does in fact have a lowest and highest element.

Theorem 4.7: Let $f: X \times \Theta \rightarrow \mathbb{R}$, where $\Theta \subset \mathbb{R}$ and X is a product set. Let $X^*(\theta) \equiv \operatorname{argmax}_{\mathbf{x} \in X} f(\mathbf{x}, \theta)$. Suppose that f is supermodular. Then, for each $\mathbf{x}' \in X^*(\theta)$ and $\mathbf{x}'' \in X^*(\theta)$, $\max\{\mathbf{x}', \mathbf{x}''\} \in X^*(\theta)$ and $\min\{\mathbf{x}', \mathbf{x}''\} \in X^*(\theta)$.

Proof: Suppress θ in the notation. If \mathbf{x}' and \mathbf{x}'' are ordered with respect to one another, there is nothing to prove. So assume that neither $\mathbf{x}' \geq \mathbf{x}''$ nor $\mathbf{x}'' \geq \mathbf{x}'$, so that $\max\{\mathbf{x}', \mathbf{x}''\} > \mathbf{x}' > \min\{\mathbf{x}', \mathbf{x}''\}$ and $\max\{\mathbf{x}', \mathbf{x}''\} > \mathbf{x}'' > \min\{\mathbf{x}', \mathbf{x}''\}$. Then

$$\begin{aligned} 0 &\geq f(\max\{\mathbf{x}', \mathbf{x}''\}) - f(\mathbf{x}') \\ &\geq f(\mathbf{x}'') - f(\min\{\mathbf{x}', \mathbf{x}''\}) \\ &\geq 0 \end{aligned}$$

Thus, the function's value is the same at all four points and is maximized at each of them. The first of these inequalities is simply the assumption that \mathbf{x}' is maximizing, the second follows from increasing differences, and the third is the assumption that \mathbf{x}'' is a maximizer. To see just how the second equality is obtained, let I be the set of indices in $\{1, \dots, N\}$ for which $x'_n > x''_n$ and J be those for which $x'_n < x''_n$. These sets are nonempty and disjoint. The points $\max\{\mathbf{x}', \mathbf{x}''\}$ and \mathbf{x}' differ only in those components whose indeces are in J , as do the points \mathbf{x}'' and $\min\{\mathbf{x}', \mathbf{x}''\}$. Then the inequality is an implication of f having increasing differences in $(\mathbf{x}_J; \mathbf{x}_I)$, which, as noted, follows from the assumption that f has increasing differences in each pair of its arguments.

The Theorem indicates that the set of maximizers of a supermodular function has a special

structure. One possibility is that there is a unique maximizer. If there are multiple maximizers, then one of two possibilities occurs. First, all the optimizers may be ordered with respect to one another in terms of the component-wise order on \Re^N : the maximizers are then said to form a *chain*. Alternatively, if there is a pair of optimizers \mathbf{x}' and \mathbf{x}'' that are not ordered, then there is another, distinct optimizer, $\max\{\mathbf{x}', \mathbf{x}''\}$, that is larger than both \mathbf{x}' and \mathbf{x}'' , and another one, $\min\{\mathbf{x}', \mathbf{x}''\}$, that is smaller than both. Note, however, that the conditions of the Theorem give no guarantee that $X^*(\theta)$ is nonempty: For this we generally need that f is upper-semicontinuous in \mathbf{x} and that X is compact. If $X^*(\theta)$ is empty for some $\theta \in \Theta$, however, then the conclusion of the Theorem holds trivially.

If the set of maximizers is nonempty and compact, then we can build from Theorem 4.7 to conclude further that there is a largest and a smallest vector in $X^*(\theta)$ in terms of the component-wise order on \Re^N .

Theorem 4.8: Let $f: X \times \Theta \rightarrow \Re$, where $\Theta \subset \Re$ and X is a product set in \Re^N . Let $X^*(\theta) \equiv \operatorname{argmax}_{\mathbf{x} \in X} f(\mathbf{x}, \theta)$. Suppose that f is supermodular and that, for each $\theta \in \Theta$, $X^*(\theta) \equiv \operatorname{argmax}_{\mathbf{x} \in X} f(\mathbf{x}, \theta)$ is nonempty and compact. Then for each θ there exists $\mathbf{x}^H(\theta) \in X^*(\theta)$ and $\mathbf{x}^L(\theta) \in X^*(\theta)$ such that, for all $\mathbf{x} \in X^*(\theta)$, $\mathbf{x}^L(\theta) \leq \mathbf{x} \leq \mathbf{x}^H(\theta)$.

Proof: The proof is by construction. Fix θ and suppress it henceforth in the notation. Let \mathbf{z}^k be any solution of $\operatorname{Min}\{\mathbf{x}_k | \mathbf{x} \in X^*\}$ and let $\mathbf{x}^L = \operatorname{Min}(\mathbf{z}^1, \dots, \mathbf{z}^N)$. By construction, $\mathbf{x}^L \leq \mathbf{x}$ for all $\mathbf{x} \in X^*$. By Theorem 4.7, $\mathbf{x}^L \in X^*$.

The argument for the existence \mathbf{x}^H is similar.

Thus, for supermodular functions, the set of maximizers has a least element and a greatest element with respect to the component-wise partial ordering on \Re^N . All the solutions of the optimization problem then lie in the n -dimensional interval $[\mathbf{x}^L(\theta), \mathbf{x}^H(\theta)] \equiv \{\mathbf{x} | \mathbf{x}^L(\theta) \leq \mathbf{x} \leq \mathbf{x}^H(\theta)\}$.

With the preceding result, we could begin to perform comparative statics analyses about the lowest and highest maximizer of a function, similar to Chapter 1. However, it will turn out that our analysis of changes in the full set of maximizers will be identical to analyses of the behavior of the lowest and highest solutions. To perform analyses of changes in sets of optimizers, we need to have an ordering on multivariate sets. In Chapter 2 we introduced Veinott's strong set order on \Re for this purpose. The corresponding order on \Re^N is defined as follows:

Definition: Let $A, B \subset \Re^N$. Then A is higher than B in the strong order set order, written $A \geq_S B$, if and only if for any $\mathbf{a} \in A$ and any $\mathbf{b} \in B$, $\max\{\mathbf{a}, \mathbf{b}\} \in A$ and $\min\{\mathbf{a}, \mathbf{b}\} \in B$. A function $X: \Theta \rightarrow 2^S$, $S \subseteq \Re^N$ from an ordered set Θ into the set of all subsets of \Re^N is nondecreasing if and only if for all $\theta, \theta' \in \Theta$ such that $\theta > \theta'$, $X(\theta) \geq_S X(\theta')$.

When $N = 1$ this is the same as the definition given in Chapter 2. In multiple dimensions, the

definition involves some new and subtle elements. For example, the relation is not reflexive: the set $S=\{(01),(1,0)\}$ does not satisfy $S \geq_S S$. (Sets that do satisfy this condition are called *sublattices* and will play a special role in the more general theory developed in the next chapter.)

The figure below illustrates pairs of sets that are comparable in the strong set order. In the disk shown in the figure, the set of points above and to the right of the point a is higher in the strong order than the set of points below and to the left of a ; and the set of points on and to the right of the vertical line is higher than the set on and to the left of the line; and the set of points in the disk that are above and to the right of the downward-sloping line are higher than the set below and to the left of the line.

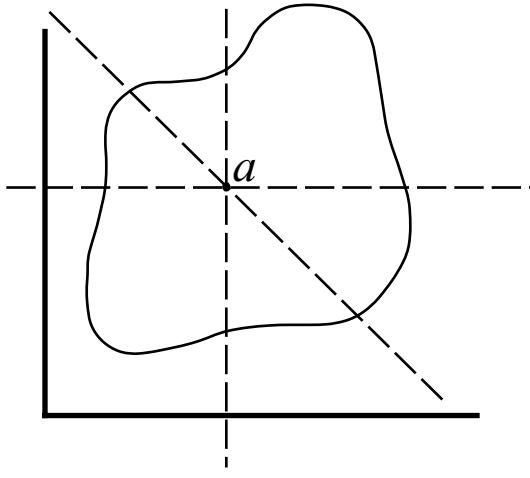


Figure 4.2 Illustrations of the multivariate strong set order.

In general, given any set S such that $S \geq_S S$, and any sets A and B such that the indicator function $\mathbf{1}_A(\mathbf{x})$ is nondecreasing in \mathbf{x} and $\mathbf{1}_B(\mathbf{x})$ is nonincreasing in \mathbf{x} , $A \cap S \geq_S B \cap S$. Further, the function $g: \Re^4 \rightarrow 2^{\Re^2}$, $g(a,b,c,d) = [a,b] \times [c,d]$, is nondecreasing in each of its arguments.

It is immediate from the definition that, for any set A , $\emptyset \geq_S A \geq_S \emptyset$. This means that we can formulate our comparative statics results in terms of the set of maximizers without concerning ourselves with conditions that guarantee that this set is non-empty: If ever the set is empty, then the monotonicity of the set in the parameter holds trivially.

We can now state our first comparative statics theorem which deals with the behavior of the bounds of the set of optimizers, as well as the set of optimizers, as the parameter changes.

Theorem 4.9: Suppose $f: X \times \Theta \rightarrow \Re$, where $\Theta \subset \Re$ and X is a product set in \Re^N . Let $X^*(\theta) \equiv \operatorname{argmax}_{\mathbf{x} \in X} f(\mathbf{x}, \theta)$.

(i) If f is supermodular, then $X^*(\theta)$ is nondecreasing in θ .

(ii) If f is supermodular and, in addition, $X^*(\theta)$ is nonempty and compact for each $\theta \in \Theta$, then $\mathbf{x}^L(\theta)$ and $\mathbf{x}^H(\theta)$ exist and are nondecreasing in θ .

Proof: (i) Suppose that $\theta > \theta'$ and let $\mathbf{x} \in X^*(\theta)$ and $\mathbf{x}' \in X^*(\theta')$. Then,

$$0 > f(\max\{\mathbf{x}, \mathbf{x}'\}, \theta) - f(\mathbf{x}, \theta) \quad \text{because } \mathbf{x} \in X^*(\theta)$$

$$\begin{aligned}
&\geq f(\mathbf{x}', \theta) - f(\min\{\mathbf{x}, \mathbf{x}'\}, \theta) && \text{by supermodularity} \\
&\geq f(\mathbf{x}', \theta') - f(\min\{\mathbf{x}, \mathbf{x}'\}, \theta') && \text{by supermodularity} \\
&\geq 0 && \text{because } \mathbf{x}' \in \mathbf{x}^*(\theta').
\end{aligned}$$

This implies that all the inequalities hold as equalities. Hence, $\max\{\mathbf{x}, \mathbf{x}'\} \in X^*(\theta)$ and $\min\{\mathbf{x}, \mathbf{x}'\} \in X^*(\theta')$, as desired.

(ii) By Theorem 4.8, $\mathbf{x}^L(\theta)$ and $\mathbf{x}^H(\theta)$ exist. Since $X^*(\theta)$ is nondecreasing by (i), we know that $\max\{\mathbf{x}^H(\theta), \mathbf{x}^H(\theta')\} \in X^*(\theta)$, and $\min\{\mathbf{x}^H(\theta), \mathbf{x}^H(\theta')\} \in X^*(\theta')$. Since $\max\{\mathbf{x}^H(\theta), \mathbf{x}^H(\theta')\} \geq \mathbf{x}^H(\theta)$, it must be that $\max\{\mathbf{x}^H(\theta), \mathbf{x}^H(\theta')\} = \mathbf{x}^H(\theta)$, otherwise we would have a contradiction of the definition of $\mathbf{x}^H(\theta)$ as the highest element of $X^*(\theta)$. Then, since $\mathbf{x}^H(\theta) = \max\{\mathbf{x}^H(\theta), \mathbf{x}^H(\theta')\} \geq \mathbf{x}^H(\theta')$, we are done.

Comparing the sets of optimizers at two values of the parameter θ , the Theorem implies that when there is a unique maximizer at each value, the maximizer is a nondecreasing function of θ : Higher values of the parameter lead to higher values of all the choice variables. If there is a unique maximizer at the lower value of θ and multiple optima at the higher value, then all the optimizers at the higher parameter value are larger than the single maximizer at the lower value of the parameter. Similarly, when there are multiple maximizers at θ and a unique one, \mathbf{x}' , at $\theta' > \theta$, then $\mathbf{x}' \geq \mathbf{x}^H(\theta)$. In addition, we are guaranteed the existence of a monotone nondecreasing selection from $X^*(\theta)$, that is, a function $\mathbf{x}^*: \Theta \rightarrow X$ such that, for all θ , $\mathbf{x}^*(\theta) \in X^*(\theta)$ and, for $\theta' > \theta$, $\mathbf{x}^*(\theta') \geq \mathbf{x}^*(\theta)$. In particular, \mathbf{x}^H and \mathbf{x}^L are both such monotone selections.

As an example, consider the following function on $\{0,1\} \times \{0,1\} \times [0,1]$:

$$f(x, y, \theta) = \begin{cases} -x - y & \theta < \frac{1}{4} \\ 0 & \frac{1}{4} \leq \theta < \frac{1}{2} \\ x & \frac{1}{2} \leq \theta < \frac{3}{4} \\ x + y & \frac{3}{4} \leq \theta \end{cases}$$

It is easy to see that, for fixed θ , f has increasing differences in (x, y) . Further, it has increasing differences in $(x; \theta)$ and $(y; \theta)$: for example, $f(1, y, \theta) - f(0, y, \theta)$ is -1 for $\theta < 1/4$, 0 for $\theta \in [1/4, 1/2)$, and 1 for $\theta \geq 1/2$, and this is an increasing function of θ . Then

$$X^*(\theta) = \begin{cases} (0, 0) & \theta < \frac{1}{4} \\ \{0, 1\} \setminus \{0, 1\} & \frac{1}{4} \leq \theta < \frac{1}{2} \\ \{(1, 0), (1, 1)\} & \frac{1}{2} \leq \theta < \frac{3}{4} \\ \{1, 1\} & \frac{3}{4} \leq \theta \end{cases}$$

Thus $X^*(\theta)$ is increasing in the strong set order, and \mathbf{x}^H and \mathbf{x}^L are nondecreasing functions of θ .

4.5 Critical Sufficient Conditions and Comparative Statics with Respect to Changes in the Constraint Set

This section treats problems when a single parameter interacts with all the choice variables, and further when the constraint set may vary with the parameters of the problem. For example, a change in government regulations might enforce stricter limits on emissions or a change in technology might expand the set of possible input-output vectors.

To give this situation a general formulation, we parameterize the constraint set by τ , so that $S(\tau)$ is a function mapping \mathfrak{R} to subsets of X . In this section, we will maintain the assumption that $S(\tau)$ is a product set for all τ , but this will be relaxed in Chapter 5. Thus we consider the problem

$$\max_{\mathbf{x} \in S(\tau)} f(\mathbf{x}, \theta) + \sum_{n=1}^N g_n(x_n) \quad (4.3)$$

Note that each g_n function depends only on x_n and not on either the parameter or any of the other elements of the choice variable. Define $X^*(\theta, \tau)$ to be the (possibly empty) set of solutions to this problem for a particular pair (θ, τ) .

To illustrate, suppose that the government imposes an upper limit on the firm's uses of certain inputs. Then the firm may choose inputs to solve a problem as follows:

$$\max_{x_i \leq \tau_i, i = 1, \dots, n} \theta f(x_1, \dots, x_n) - \sum_{i=1}^n w_i x_i$$

In this problem, which is a special case of problem (4.3), θ is the price of output and the w_i 's are input prices. The constraint set in this problem is nondecreasing in the strong set order as a function of the vector parameter τ . The comparative statics of such problems are the subject of the next theorem.

Theorem 4.10 Assume the following about problem (4.3):

(S4.10) X and $S(\tau) \subseteq X$ are product sets for all τ

(F4.10) No restrictions on $f: X \times \Theta \rightarrow \mathfrak{R}$, where $\Theta \subset \mathfrak{R}$.

(G4.10) No restrictions on $g_n: \mathfrak{R} \rightarrow \mathfrak{R}$, $n = 1, \dots, N$.

Then $X^(\theta, \tau)$ is nondecreasing in θ and τ for all g_n and all functions $S(\tau)$ which are nondecreasing in the strong set order if and only if f is supermodular.*

Proof: We know that supermodularity implies the comparative statics conclusion (1) by Theorem 4.4. Now, we treat the reverse implication. Thus, fix x_{-1} and let $x_1 > x_1'$ and $\theta > \theta'$. Let $S(\tau) = \{x_1, x_1'\} \times \{x_{-1}\}$ and let $g_{-1} = 0$ and g_1 such that $f(x_1, x_{-1}, \theta') + g_1(x_1) = f(x_1', x_{-1}, \theta') + g_1(x_1')$. Then $X^*(\theta', \tau) = \{x_1, x_1'\}$. The comparative statics conclusion then requires that $x_1 \in X^*(\theta, \tau)$, which requires that $f(x_1, x_{-1}, \theta) + g_1(x_1) \geq f(x_1', x_{-1}, \theta) + g_1(x_1')$. Hence, $f(x_1, x_1, \theta) - f(x_1', x_{-1}, \theta) \geq f(x_1, x_{-1}, \theta) - f(x_1', x_{-1}, \theta')$, which proves increasing

differences in (x_1, θ) , or generally in (x_n, θ) .

For increasing differences between x_1 and x_2 , we fix x_{-12} and θ and suppress those arguments, for example by writing $f(x_1, x_2)$ for $f(x_1, x_2, x_{-12}, \theta)$. Let $x_1 > x_1'$ and $x_2 > x_2'$, $S(\tau) = \{x_1, x_1'\} \times \{x_2\}$ and $S(\tau') = \{x_1, x_1'\} \times \{x_2'\}$. Notice that $S(\tau) \geq S(\tau')$. Let $g_{-1} = 0$ and g_1 so that $f(x_1, x_2') + g_1(x_1) = f(x_1', x_2') + g(x_1')$. Then $X^*(\tau, g) = \{x_1, x_1'\}$. Hence, by the strong set order, $x_1 \in X^*(\tau, g)$. Therefore, $f(x_1, x_2) + g_1(x_1) \geq f(x_1', x_2) + g(x_1')$. Thus, $f(x_1, x_2) - f(x_1', x_2) \geq f(x_1, x_2') - f(x_1', x_2')$, as was to be proved.

Remark: The requirement that the optimum be monotonic in τ for *all possible nondecreasing functions* $S(\tau)$ is a very restrictive one. In the proof of the theorem, this requirement ensures that there are parameter changes that can force the optimal value of any component, say x_1 , to move from any level to any other level. If less robustness than this is required in the application of interest, then the theorem leaves open the possibility that supermodularity may be a stronger condition than necessary.

We now return to comparing the new approach with the older one based on implicit function methods. To facilitate that, suppose that the following conditions, which permit application of the implicit function theorem, are satisfied: convexity of the underlying choice space, convexity of the constraint set, and smoothness and strong concavity⁶ of the objective function. We know that supermodularity of the objective function is sufficient for monotone comparative statics even without all these extra assumptions, and we know that in their absence it is also necessary if monotonicity is to hold for all g_n functions. Is supermodularity still necessary when the additional assumptions described above are made? Or, do the extra assumptions made in the standard analysis allow us to obtain a weaker set of critical sufficient conditions? The following theorem indicates that even with these additional assumptions and even if we also add the assumption that the g_n functions are linear, supermodularity is still a necessary condition for the desired conclusion.

Theorem 4.11: Assume the following about problem (4.3):

(S4.11) X is a convex product set in \mathbb{R}^N , $S(\tau) \subset X$ is a convex product set for all τ .

(F4.11) $f: X \times \Theta \rightarrow \mathbb{R}$ ($\Theta \subset \mathbb{R}$) is twice continuously differentiable in \mathbf{x} and strongly concave for all θ .

(G4.11) Each $g_n: \mathbb{R} \rightarrow \mathbb{R}$, $n = 1, \dots, N$, is linear.

Then $X^(\theta, \tau)$ is nondecreasing in θ and τ for all g_n ($n=1, \dots, N$) and all $S(\tau)$ which are nondecreasing in the strong set order if and only if f is supermodular.*

Proof: We use Theorem 2.5 to show that f must have nondecreasing differences in each (x_n, θ) -pair and also in each (x_n, x_m) -pair (note that for each n , the family of all functions

⁶ A function is *strongly concave* if its matrix of mixed partial derivatives exists everywhere and is everywhere negative definite.

g_n , which satisfy (G4.11) is a full one-parameter family). For the first of these conclusions, fix x_{-n} and set $S(\tau) = \{\mathbf{z} \in X \mid z_{-n} = x_{-n}\}$. Then, the problem is reduced to one of choosing a single real variable, x_n , and the conclusion of Theorem 2.5 implies that f has nondecreasing differences in (x_n, θ) . To show that there are nondecreasing differences in (x_n, x_m) , fix θ and x_{-nm} . Let $x_m > x_m'$ and define $S(\tau') = \{z \in X \mid z_{-nm} = x_{-nm}, z_m = x_m'\}$ and $S(\tau) = \{z \in X \mid z_{-nm} = x_{-nm}, z_m = x_m\}$. Notice that $S(\tau)$ and $S(\tau')$ differ only in the way they fix x_m . Consequently, the condition that $X^*(\tau, p) \geq X^*(\tau', p)$ in the original problem with x_m as a choice variable is equivalent to the condition that $X^*(x_m, p) \geq X^*(x_m', p)$ in the related problem where x_m is treated as an exogenous parameter. By Theorem 2.5, this latter conclusion holds for all p (and all x_{-nm}) if and only if f has nondecreasing differences in x_n and x_m .

4.6 Multiplicatively Separable Problems

4.7 Further Applications

4.7.1 The LeChatlier Principle with Three or More Inputs

4.8 Problems

1. Show that $f: \mathbb{R}^N \rightarrow \mathbb{R}$ has increasing differences in each pair $(x_i; x_j)$, $i \neq j$, if and only if it has increasing differences in $(\mathbf{x}_I; \mathbf{x}_J)$ for each disjoint pair of subsets I and J of $\{1, \dots, N\}$.
2. Let us revisit the classical theory of the firm. (i) Suppose there are two inputs, capital (k) and labor (l) with production function $F(k, l)$ and profit objective $pF(k, l) - rk - wl$. Use a single direct application of Theorem 4.2 or 4.3 to show that if $F_{kl} \leq 0$, then $k^*(w)$ is nondecreasing and $l^*(w)$ is nonincreasing. (ii) Suppose the input vector to the production function is $\mathbf{x} = (x_1, \dots, x_n)$ and output is given by $F(\mathbf{x})$. Further suppose that some inputs are indivisible and others are in limited supply, so that the firm operates subject to the constraints $x_1 \in K_1, \dots, x_n \in K_n$. The firm's objective function is then $pF(\mathbf{x}) - \mathbf{w} \cdot \mathbf{x}$, where \mathbf{w} now denotes the vector of input prices. Show that if $F_{ij} \leq 0$ for all $1 \leq i < j \leq n$, then regardless of the F_{ii} terms and the sets K_i , $x^*(\mathbf{w}, p)$ is nonincreasing in \mathbf{w} and nondecreasing in p . (iii) Let the set-up be the same as in part (ii) but suppose that there is a technology parameter θ which affects the productivity of the first input, so that the firm's objective function is $pF(\theta x_1, \mathbf{x}_{-1}) - \mathbf{w} \cdot \mathbf{x}$. Determine how changes in the technology θ affect the demand for inputs. In particular, what additional condition on F is needed to conclude that an increase in θ raises the demand for input 1, that is, that $\mathbf{x}_1(\theta)$ is nondecreasing? What additional condition is needed to conclude that $\mathbf{x}_1(\theta)$ is monotone nondecreasing?
3. A manufacturing firm makes n products on a single group of machines. The total demand for all its products is D per day. To maintain an average inventory of x per product, it must produce the product in run lengths of x , where x is a positive integer. Each run requires a set-up costing c in wages and production downtime. Holding cost for inventories is h per unit per day. More expensive machines are easier and cheaper to set up, but the cost of a machine is $K(c)$, which is a decreasing function of c . If changes in demand make it desirable to increase variety n without affecting total demand, what are the optimal adjustments in x and c ? What assumptions do you need to make about $K(\cdot)$?
4. Let w be the wage paid to a certain class of employees and t be the expenditure on training afforded to them, which is assumed to be a fixed cost paid once per employee. Let $L(w)$ be the anticipated length of the job tenure and assume that $L(\cdot)$ is increasing: employees paid higher wages quit less often. Let d be the discretion allowed to employees and let $\theta P(t, d)$ be the revenue product of the employee in the job. Assuming an interest rate of zero, formulate a plausible objective for the profit maximizing firm in this context, where the firm's choice variables are t , w , and d . Analyze how an increase in θ affects the optimal values of the choice variables. What assumptions do you need to make to obtain a determinate answer?⁷

⁷This problem is loosely based on a draft paper by Susan Athey, Joshua Gans, Scott Schaefer and Scott Stern.

