

Monotone Comparative Statics: Geometric Approach

B.H. Strulovici · T.A. Weber

Published online: 29 December 2007
© Springer Science+Business Media, LLC 2007

Abstract We consider the comparative statics of solutions to parameterized optimization problems. A geometric method is developed for finding a vector field that, at each point in the parameter space, indicates a direction in which monotone comparative statics obtains. Given such a vector field, we provide sufficient conditions under which the problem can be reparameterized on the parameter space (or a subset thereof) in a way that guarantees monotone comparative statics. A key feature of our method is that it does not require the feasible set to be a lattice and works in the absence of the standard quasi-supermodularity and single-crossing assumptions on the objective function. We illustrate our approach with a variety of applications.

Keywords Change of parameters · Parameterized optimization problems · Single-crossing · Supermodularity

1 Introduction

In many model-based optimization problems, important insights can be derived by comparing solutions for different parameter values. The model's *parameters* are *ex-*

Communicated by D.G. Luenberger.

We are grateful to Kenneth Arrow, Darrell Duffie, David Luenberger, Paul Milgrom, John Quah, and Pete Veinott for helpful comments.

Research in part supported by a David Morgenthaler II Faculty Scholar Award.

B.H. Strulovici
Nuffield College, Oxford University, Oxford, UK
e-mail: bruno.strulovici@economics.ox.ac.uk

T.A. Weber (✉)
Department of Management Science and Engineering, 442 Terman Engineering Center,
Stanford University, Stanford, CA 94305-4026, USA
e-mail: webert@stanford.edu

ogenously specified and often can be varied for analysis purposes, while its *variables*, which constitute the building blocks for its predictions, are *endogenously* determined by (i) imposed model relations and (ii) parameter values. For instance, an economic model might be concerned with a firm's optimal production of widgets (e.g., in terms of capital and labor requirements), given both a production function relating output to inputs and a set of prices (e.g., the market price for widgets, the cost of capital, and an average wage rate). In the neoclassical tradition the model would then impose maximization of the firm's profit to determine optimal amounts of factor inputs as a function of their respective prices. More generally, if a model's predictions can be expressed as an optimal action in some finite-dimensional space, then *comparative statics* studies the direction in which the optimal action changes as a result of some disturbance in the values of the model's parameters [1]. Thus, in our example, normalizing the price of the firm's output to one, the optimal choice of inputs critically depends on the prices of the production factors which are this model's parameters.

Sensitivity and perturbation analysis has been the traditional mathematical approach to analyze changes of objective values and solutions with respect to parameters [2, 3]. The associated methods focus on establishing regularity properties and generally do not consider the *monotonicity* properties of solutions, which are of prime importance for decision makers in many applied optimization problems. In this paper we consider the possibility of guaranteeing monotone behavior of solutions through an appropriate reparameterization of the problem. The key question of *monotone comparative statics* is to determine under what conditions the model predictions vary monotonically with the parameters [4, 5]. General answers to this question for optimal actions chosen from feasible sets, which are usually assumed to be lattices satisfying a set-monotonicity requirement with respect to the parameters, are provided by [4–6]. They provide a necessary and sufficient condition for optimal actions to exhibit monotone comparative statics with respect to the parameters. For the special case of our neoclassical production decision problem, where the production function is independent of the factor prices, monotone comparative statics (i.e., inputs nonincreasing in prices) obtain if and only if the production function is supermodular, which—assuming twice continuous differentiability—amounts to requiring that all cross-partial derivatives of the production function are nonnegative.

The practical importance of monotone comparative statics that justifies its widespread use, particularly in economics, lies in the fact that robust insights can be obtained in the absence of an analytical solution to the model: the monotonicity of optimal actions in parameters is guaranteed if the system's objective function satisfies certain easy-to-check requirements. In addition, monotonicity of optimal actions can yield useful rules of thumb for decision makers and thus help in arriving at “optimally imperfect decisions” [7]. Clearly, in our production example (which is examined more closely in Sect. 5.2) it would be helpful for the firm to be able to translate price movements immediately (without any further computations) into appropriate input changes which at the very least vary in the right *direction*, even when its production function is not supermodular due to anticomplementarities between factor inputs. Unfortunately, the currently available theory on monotone comparative statics returns negative results in situations where the aforementioned characterization of monotone comparative statics by [6] fails. We argue that this failure is often due

to the fact that the parameterization of the problem is taken as given.¹ Indeed, our results indicate that it may be possible to achieve monotonicity of solutions in new parameters that are obtained by a one-to-one mapping from the original parameter space. In fact, for problems with a smooth analytical structure we show that, provided with sufficiently precise knowledge about the location of an optimal action in the action space, it is possible to find a reparameterization that achieves monotone comparative statics in any single component of the decision.

Our central goal is to provide a new method for achieving monotone comparative statics of solutions to parameterized optimization problems, first by relaxing and thereby generalizing the standard monotone comparative statics problem, and second, by providing a systematic way to reparameterize the problem such that monotone comparative statics can be achieved.² In addition to addressing cases where the standard results do not apply, our method can, by offering a new description of the parameter space, shed light on important relations between decision variables and parameters of economic problems.

The approach developed here builds on tools in differential geometry, and we thus require models with a smooth structure (e.g., a parameterized optimization problem with a twice continuously differentiable objective function), even though—as we are well aware—none of our statements fundamentally depends on the differentiable structure. All of our results can also be expected to hold if the problem is suitably discretized; nevertheless, we prefer to adopt a differentiable approach for ease of exposition, since then the tools of differential geometry can be applied seamlessly. We decompose the problem of changing problem parameters to achieve monotone comparative statics (MCS) into two parts. First, the decision maker needs to solve a *local* MCS problem by finding for each point t in the parameter space \mathcal{T} a direction vector $v(t)$ (i.e., an element of the tangent space of \mathcal{T} at t) that would increase the optimal action $x(t)$ if parameters were to be locally changed from t in the direction of $v(t)$. If the location of $x(t)$ is not known precisely, as is generally the case, then the direction $v(t)$ must be such that it induces local monotonicity with respect to all points in a subset $\mathcal{R}(t)$ of the action space \mathcal{X} which is known to contain the optimal action $x(t)$, given the possibility that any point in $\mathcal{R}(t)$ might turn out to be optimal. A full solution to this problem consists in a vector field $v(t)$, defined for all values of t in the parameter space. Second, given the vector field $v(t)$, the decision maker needs to find a reparameterization solving the *global* MCS problem. We show that this can always be achieved locally through “rectification” of the vector field. Under a few additional conditions rectification can also be achieved globally, leading to the desired global MCS reparameterization of the decision problem.

The paper proceeds as follows. In Sect. 2 we introduce the problem of obtaining monotone comparative statics (i.e., monotone dependence of solutions on parameters) for parameterized optimization problems, both from a local and a global view-

¹This is true in the literature, except for some rare cases where trivial reparameterizations such as a change of sign or other simple *ad hoc* reparameterizations are chosen under very special circumstances, e.g., by [8] in a network flow problem.

²Even though not explicitly developed, our methods apply equally to equilibrium problems, by replacing the first-order necessary optimality conditions of the optimization problem with the equations specifying equilibria.

point. When considered locally, obtaining monotone comparative statics corresponds to finding directions in the parameter space in which solutions to the optimization problem are nondecreasing in parameters. A solution to the local monotone comparative statics problem is provided in Sect. 3 using a vector field method. Subsequently, in Sect. 4, we take a more global perspective: having obtained a vector field of monotone comparative statics directions defined at each point of the parameter space, we demonstrate that it is possible (at least locally) to change the parameters of the optimization problem, i.e., to *reparameterize* it such that monotone comparative statics of the solutions of the reparameterized problem obtain. We show that the reparameterization can be global if a hyperplane can be found that is transverse to a vector field that solves the local MCS problem at each point of the parameter space. To illustrate our results we discuss a number of economic applications in Sect. 5 and conclude with a discussion and directions for further research in Sect. 6.

2 Problem Formulation

We consider a decision maker who, given a parameter value $t \in \mathcal{T}$, aims to select an element $x(t)$ of an action space \mathcal{X} so as to maximize her objective function $f : \mathcal{X} \times \mathcal{T} \rightarrow \mathbb{R}$. She thus tries to solve the *parameterized optimization problem*

$$x(t) \in \arg \max_{x \in \mathcal{X}} f(x, t), \tag{1}$$

where \mathcal{X} is a n -dimensional compact manifold³ and \mathcal{T} is an m -dimensional compact manifold. Provided that a solution to (1) exists,⁴ the decision maker is interested in the comparative statics of the maximizer $x(t)$ as $t \in \mathcal{T}$ varies. More generally, we assume that the decision maker is concerned with the behavior of the composition $\varphi \circ x(t) = \varphi(x(t))$, where $\varphi : \mathcal{X} \rightarrow \mathbb{R}^d$ (with $1 \leq d \leq n$) is an *evaluation function* that the decision maker uses to assess any solution $x(t)$ of (1). For example, if the decision maker is interested only in the comparative statics of the first component of the maximizer $x(t) = (x_1, \dots, x_n)(t)$, she can choose $\varphi(x) = x_1$. From Milgrom and Shannon’s [6] monotonicity theorem, we know that if $\varphi(x) = x$ and \mathcal{X} is a lattice, then $\varphi \circ x(t) = x(t)$ is increasing in t (on any given sublattice of X) if and only if f is quasi-supermodular⁵ in x and satisfies the single-crossing property⁶ in (x, t) .

³ If \mathcal{X} lies in a lower-dimensional submanifold of \mathbb{R}^n , the analysis can still be applied, but differential calculus should be understood on this submanifold, and openness should be understood relative to the submanifold, cf. Sect. 3.5.

⁴ If for any parameter $t \in \mathcal{T}$ the function $f(\cdot, t)$ is continuous and \mathcal{X} is bounded, a solution to the parameterized optimization problem (1) exists in the closure of \mathcal{X} by the Weierstrass theorem ([9], p. 540).

⁵ A real-valued function f defined on a lattice $\mathcal{X} \subset \mathbb{R}^n$ is quasi-supermodular if $f(x) \geq (>) f(x \wedge y)$ implies $f(x \vee y) \geq (>) f(y)$, for all x, y in \mathcal{X} , where $x \vee y = (\max\{x_1, y_1\}, \dots, \max\{x_n, y_n\})$ and $x \wedge y = (\min\{x_1, y_1\}, \dots, \min\{x_n, y_n\})$. As its name suggests, quasi-supermodularity is a weaker condition than supermodularity.

⁶ A real-valued function f defined on the product $\mathcal{X} \times \mathcal{T}$ of two partially ordered spaces has the single-crossing property if, whenever $x' > x$ and $t' > t$, $f(x', t) \geq (>) f(x, t)$ implies $f(x', t') \geq (>) f(x, t')$. The single-crossing property is a weaker condition than supermodularity in (x, t) .

Conversely, for any objective function f that does *not* satisfy these conditions, $x(t)$ cannot be nondecreasing on \mathcal{T} . Nevertheless, despite this negative result using the standard theory, it may be possible for the decision maker to at least find a *path* γ_t in the parameter space, so that starting at a given $t \in \mathcal{T}$ the function $\varphi \circ x$ is nondecreasing along γ_t . Monotone comparative statics may thus be obtained following certain directions in the parameter space. It is useful given any $t \in \mathcal{T}$ to state the decision maker’s MCS problem (at t) in precisely these terms.

MCS Problem (at t) Given a continuously differentiable evaluation function φ and a parameter value $t \in \mathcal{T}$, find a nonempty open interval $\mathcal{I}_t \subset \mathbb{R}$ with $0 \in \mathcal{I}_t$ and a path $\gamma_t : \mathcal{I}_t \rightarrow \mathcal{T}$, such that $\varphi \circ x(\gamma_t(\lambda))$ is nondecreasing⁷ for all $\lambda \in \mathcal{I}_t$ and $\gamma_t(0) = t$.

If a solution $(\mathcal{I}_t, \gamma_t)$ of the MCS problem at t is such that the path γ_t cannot be extended in \mathcal{T} , then we call the solution *maximal*. We can restrict our attention, without any loss of generality, to a maximal solution of the MCS problem at t . Monotone comparative statics relative to an evaluation function φ and a parameter starting value t obtain whenever the vector $\varphi \circ x$ is componentwise nondecreasing along an appropriate path γ_t in the parameter space \mathcal{T} . Along any such path, $x(\gamma_t(\lambda))$ solves (1) for all $\lambda \in \mathcal{I}_t$. A solution to the MCS problem for all $t \in \mathcal{T}$ results in a *global flow* $\theta(\lambda, t) = \gamma_t(\lambda)$, for which $\theta(0, t) = t$ and

$$\lambda \leq \mu \implies \varphi \circ x(\theta(\lambda, t)) \leq \varphi \circ x(\theta(\mu, t)), \tag{2}$$

for any $\lambda, \mu \in \mathcal{I}_t$. If it is possible to represent the global flow in the form

$$\theta(\lambda_t, \psi(t)) = t, \tag{3}$$

where λ_t is uniquely determined and $\psi : \mathcal{T} \rightarrow \mathcal{P}$ is a function that maps the parameter space to a fixed $(m - 1)$ -dimensional hypersurface \mathcal{P} , transverse⁸ to the vector field induced by the MCS paths γ_t , then using the new parameters $s(t) = (s_1, \dots, s_m)(t) = (\lambda_t, \pi(\psi(t)))$ guarantees monotone comparative statics of $\varphi \circ x(s)$ in $s_1 = \lambda$, at least locally, where π is a diffeomorphism from \mathcal{P} onto a subset of \mathbb{R}^{m-1} (details are provided in Sect. 4). Letting $\mathcal{F} = \bigcup_{t \in \mathcal{T}} \mathcal{I}_t \times \{t\}$, we formulate the global parameter-change problem accordingly.

Global MCS Reparameterization If the flow $\theta : \mathcal{F} \rightarrow \mathcal{T}$ solves the MCS problem everywhere in \mathcal{T} , find a new parameterization $s(t) = (s_1, \dots, s_m)(t) = (\lambda_t, \pi(\psi(t)))$ such that (3) is satisfied for all $t \in \mathcal{T}$.

A global MCS reparameterization provides the decision maker with new problem parameters $s = (s_1, \dots, s_m)$ that guarantee monotonicity of $\varphi(\hat{x}(s))$ in the first component $s_1 \in \mathcal{I}$, where

$$\hat{x}(s) = \arg \max_{x \in \mathcal{X}} \hat{f}(x, s) \tag{4}$$

and $\hat{f}(x, s)$ corresponds to the objective function $f(x, t)$ after the parameter change.

⁷A vector $v(\lambda)$ is nondecreasing in λ if and only if each of its components is nondecreasing in λ .

⁸See Assumption A6 for the precise definition.

3 Solving the MCS Problem at t

Our goal is to find directions in which solutions of the parameterized optimization problem (1) (or functions thereof) are increasing. For this we introduce a “pseudo-gradient” $W(x, t)$ which mimics the gradient matrix $\nabla_t x(t) = W(x(t), t)$ corresponding to all potential solutions $x \in \mathcal{R}(t)$ with respect to the parameters t on a set $\mathcal{R}(t)$ that is known to contain the actual solution $x(t)$. We refer to $\mathcal{R}(t)$ as a “reduced feasible set,” for it is a subset of the set of all feasible actions \mathcal{X} . The cardinality of $\mathcal{R}(t)$ is a measure of how much information the decision maker has about the location of the solution to (1) at t . If for a given $t \in \mathcal{T}$ all row-vectors $\nabla_t x_i(t)$, $i \in \{1, \dots, n\}$, of the pseudo-gradient lie in the same m -dimensional half-space for all points of the reduced feasible set (a subset of \mathcal{X}), then a direction $v(t) \in \mathbb{R}^m \setminus \{0\}$ exists in which monotone comparative statics obtain locally. If such a direction $v(t)$ can be found for all points t of the parameter space \mathcal{T} , then the resulting vector field $v : \mathcal{T} \rightarrow \mathbb{R}^m$ constitutes a solution to the MCS problem on \mathcal{T} . The flow induced by this vector field can be used to obtain an MCS reparameterization of the optimization problem (1), which is discussed in Sect. 4. In this section we first introduce a number of assumptions needed for the vector field method. We then provide techniques to implement the method, and relate the vector field method to classic supermodularity results. Finally, we provide important methods to deal with problems that contain equality and/or inequality constraints.

3.1 Assumptions

In order to use standard tools from differential geometry, we require that f be sufficiently smooth.

Assumption A1 (Smoothness) The objective function f is twice continuously differentiable in x and has continuous cross-derivatives with respect to each tuple (x_i, t_k) , for all $1 \leq i \leq n$ and all $1 \leq k \leq m$.

Let us denote by

$$H(x, t) = \nabla_{xx} f(x, t) = \left[\frac{\partial^2 f(x, t)}{\partial x_i \partial x_j} \right]_{i,j=1}^n$$

the Hessian matrix of $f(\cdot, t)$ evaluated at (x, t) and by

$$K(x, t) = \nabla_{xt} f(x, t) = \left[\frac{\partial^2 f(x, t)}{\partial x_i \partial t_k} \right]_{i,k=1}^{n,m}$$

the matrix of cross-derivatives of f between decision-variable and parameter components, evaluated at (x, t) . In order to bypass (at least for now) any difficulties arising from binding constraints at the optimal action, we will assume that the parameterized optimization problem (1) possesses a unique unconstrained optimum. Constrained optimization problems often can be restated equivalently so as to satisfy this assumption, and Sect. 3.5 is dedicated to this issue. We also emphasize that nothing in our method requires that the feasible set \mathcal{X} be a lattice (cf. also footnote 18).

Assumption A2 (Existence and Uniqueness) For each $t \in \mathcal{T}$, the parameterized optimization problem (1) has a unique solution $x(t)$.

In general, the set of maximizers is guaranteed to be nonempty and in the interior of \mathcal{X} if, in addition to being continuous, f is *coercive* relative to \mathcal{X} , in the sense that for any $t \in \mathcal{T}$ there exists a point $\check{x}(t) \in \mathcal{X}$ such that⁹ $f(\check{x}(t), t) \geq \sup f(\partial\mathcal{X}, t)$, (cf. [9], pp. 8, 540). If $f(\cdot, t)$ possesses multiple strict local extrema for some $t \in \mathcal{T}$, our results can be applied with respect to the comparative statics of each local maximum. Since \mathcal{X} is open, Fermat’s Lemma ([10], Vol. I, p. 215) implies that each strict local extremum $c(t)$ (and in particular the unique global maximum $x(t)$ guaranteed by Assumption A2) is a critical point of $f(\cdot, t)$, i.e., $\nabla_x f(c(t), t) = 0$. We denote by $\mathcal{C}(t)$ the *set of all critical points*¹⁰ of $f(\cdot, t)$ in \mathcal{X} at t ,

$$\mathcal{C}(t) = \{x \in \mathcal{X} : \nabla_x f(x, t) = 0\}.$$

Sard [11, 12] has shown that the corresponding *set of critical values* $f(\mathcal{C}(t), t)$, which is the image of $\mathcal{C}(t)$ under f , is of Lebesgue-measure zero.

Thus, if the decision maker can determine $\mathcal{C}(t)$, she might be able to find the optimal action $x(t)$ as the solution of a *reduced* optimization problem, $x(t) = \arg \max_{c \in \mathcal{C}(t)} f(c, t)$, using the first-order necessary optimality conditions. Assumption A2 also implies that for any $t \in \mathcal{T}$ at the unique global optimum $x(t)$ the Hessian matrix of $f(\cdot, t)$ is negative semidefinite, i.e., $x(t)$ satisfies the *second-order necessary optimality condition* $x(t) \in \mathcal{D}(t)$, where

$$\mathcal{D}(t) = \{x \in \mathcal{X} : H(x, t) \leq 0\}.$$

This allows the decision maker to further reduce the optimization problem combining the first-order and second-order necessary optimality conditions and solve

$$x(t) = \arg \max_{x \in \mathcal{R}(t)} f(x, t), \tag{5}$$

where we refer to $\mathcal{R}(t) \subseteq \mathcal{X}$ as a *reduced feasible set*; in this case $\mathcal{R}(t) = (\mathcal{C} \cap \mathcal{D})(t)$. More generally, we refer to *any* subset $\mathcal{R}(t)$ of \mathcal{X} which is guaranteed to contain the solution $x(t)$ of (1) as an *admissible reduced feasible set*. Any element of $\mathcal{R}(t)$ is called a *reduced-feasible action*. If the decision maker can determine an admissible reduced feasible set $\mathcal{R}(t)$ (e.g., by using first- and second-order necessary optimality conditions), so that she is able to solve the (reduced) parameterized optimization problem (5) on \mathcal{T} , then the MCS problem always has a trivial solution, as will become clear below (cf. Theorem 3.1). Unfortunately, in many practical applications, a closed-form solution of (1) is not possible, or the objective function is not perfectly known by the decision maker (see [13] and Sect. 6). In that case, by constructing a reduced feasible set $\mathcal{R}(t) \subseteq \mathcal{X}$ that is guaranteed to contain the optimal action $x(t)$

⁹ $\partial\mathcal{X}$ denotes the boundary of \mathcal{X} .

¹⁰Note that if $f(\cdot, t)$ has a critical point (i.e., $\mathcal{X} \cap \mathcal{C}(t) \neq \emptyset$) and is strictly concave on \mathcal{X} for all $t \in \mathcal{T}$, then Assumption A2 is automatically satisfied, since $f(\cdot, t)$ is necessarily single-peaked on \mathcal{X} .

(e.g., by using heuristics related to the special structure of the problem), the decision maker may still be able to solve the MCS problem without an explicit solution to the (equivalent) parameterized optimization problems (1) and (5). To obtain a solution to the MCS problem when the optimal action $x(t)$ can be only imperfectly localized in the set $\mathcal{R}(t) \subseteq \mathcal{X}$, we require that all critical points of $f(\cdot, t)$ in $\mathcal{R}(t)$ be nondegenerate (i.e., such that the Hessian matrix of $f(\cdot, t)$ is nonsingular there).

Assumption A3 (Nondegeneracy) For any $t \in \mathcal{T}$, the Hessian matrix $H(x, t)$ is nonsingular for all $x \in \mathcal{R}(t)$, for some reduced feasible set $\mathcal{R}(t) \subseteq \mathcal{X}$ which contains $x(t)$.

This assumption guarantees that the inverse $H^{-1}(x, t)$ is well defined and continuous at any point $(x, t) \in \mathcal{R}(t) \times \mathcal{T}$. Hence the expression $-(H^{-1}K)(x, t)$, evaluated at a point (x, t) possibly *different* from the optimal $(x(t), t)$, is well-defined. Assumption A3 is automatically satisfied if the objective function is strictly concave.

Lemma 3.1 Under Assumptions A1–A3, the unique optimal solution $x(t)$ of the parameterized optimization problem (1) is continuously differentiable on \mathcal{T} . The corresponding Jacobi matrix is given by

$$\nabla_t x(t) = \left[\frac{\partial x_i(t)}{\partial t_k} \right]_{i,k=1}^{n,m} = -(H^{-1}K)(x(t), t), \tag{6}$$

for all $t \in \mathcal{T}$.

Proof By Assumption A2, a unique interior solution $x(t)$ to the parameterized optimization problem (1) exists for all $t \in \mathcal{T}$, satisfying $H(x(t), t) \leq 0$ and

$$\nabla_x f(x(t), t) = 0. \tag{7}$$

By Assumption A1, we can differentiate (7) with respect to t (using the chain rule) and obtain

$$\nabla_t(\nabla_x f(x(t), t)) = (\nabla_{xx} f(x(t), t))(\nabla_t x(t)) + \nabla_{xt} f(x(t), t) = 0,$$

or equivalently

$$H(x(t), t)\nabla_t x(t) + K(x(t), t) = 0,$$

for all $t \in \mathcal{T}$. Since $H(x(t), t)$ is nonsingular by Assumption A3, we get expression (6) after left-multiplication with $H^{-1}(x(t), t)$ in the last equality. We now show that $\nabla_t x(t)$ is continuous. Since the maximizer $x(t)$ is unique and the objective function f continuous, we have that, as a consequence of Berge’s ([14], p. 116) maximum theorem, the maximizer $x(t)$ is continuous in t (for it is upper-semicontinuous and single-valued). By virtue of Assumption A1 and the nonsingularity of H , all entries of the matrix $(H^{-1}K)(x(t), t)$ are well defined and, as a composition of continuous functions, also continuous. Hence, the Jacobi matrix $\nabla_t x(t)$ on the left-hand side of (6) must also be continuous, which completes the proof. □

In order to study the monotonicity of $\varphi \circ x(t)$, we require some smoothness as well as *functional independence* ([10], Vol. I, p. 508) of the evaluation function.

Assumption A4 (Functional Independence) The evaluation function φ is continuously differentiable on \mathcal{X} and its Jacobi matrix

$$\Phi(x) = \nabla_x \varphi(x) = \left[\frac{\partial \varphi_l(x)}{\partial x_i} \right]_{l,i=1}^{d,n}$$

has (full) rank d for any (reduced-)feasible action x .

Assumption A4 is not critical for our results and can (except for the smoothness portion) be relaxed. Functional independence guarantees that the MCS problem is locally never trivial, since no two of φ 's components are collinear. If Assumptions A1–A4 hold, then for any (reduced-)feasible tuple (x, t) , we can define the $(d \times m)$ -matrix

$$W(x, t) = -\Phi(x)H^{-1}(x, t)K(x, t), \tag{8}$$

which we term the *pseudogradient of the MCS problem at (x, t)* . In analogy to Lemma 3.1, it is easy to show that the pseudo-gradient evaluated at any optimizing decision-parameter tuple $(x(t), t)$ describes the comparative statics of $\varphi \circ x(t)$ along paths parallel to the standard coordinates in the parameter space \mathcal{T} , i.e.,

$$\nabla_t \varphi(x(t)) = W(x(t), t). \tag{9}$$

We say that the pseudo-gradient of the MCS problem is *orientable* at (x, t) if the collection of all of its row vectors is a subset of a common half space of \mathbb{R}^m . If for a given $t \in \mathcal{T}$ the row vectors of $W(x, t)$ lie in a common half space of \mathbb{R}^m for all x in a set $\mathcal{Y} \subseteq \mathcal{X}$, then we say that the *pseudogradient is orientable on \mathcal{Y} at t* .

Example 3.1 If $m = n \geq 1$, $\phi(x) = x$, and $f(x, t) = x'Qx - t'x$, where Q is a symmetric negative definite matrix, then $\Phi = I$, $H = Q < 0$, and $K = -I$. The corresponding pseudogradient is $W = -\Phi H^{-1}K = Q^{-1}$. This pseudogradient is orientable, since its n row vectors are linearly independent and span \mathbb{R}^n . For example, one can take the sum of the row vectors as a normal vector of the half-space that contains all row vectors.

3.2 Vector Field Method

To find a path γ_t that solves the MCS problem at $t \in \mathcal{T}$, our method requires eliciting some information about the direction of the gradient $\nabla_t(\varphi(x(t)))$, given that the decision maker knows only that $x(t)$ lies in some reduced feasible set $\mathcal{R}(t) \subseteq \mathcal{X}$. More specifically, we need to determine a direction forming acute angles with the pseudo-gradient of the MCS problem. Such a direction exists if the pseudo-gradient is orientable.

Assumption A5 (Pseudogradient Orientability) There exists a continuously differentiable vector field $v(t) \in \mathbb{R}^m \setminus \{0\}$ such that, for each $t \in \mathcal{T}$,

$$W(x, t)v(t) \geq 0, \quad (10)$$

for all $x \in \mathcal{R}(t)$, where $\mathcal{R}(t) \subseteq \mathcal{X}$ is an admissible reduced feasible set.

Under Assumption A5, $v(t)$ defines a vector field on \mathcal{T} and a phase diagram with paths corresponding to the flow of this vector field.¹¹ The key result of this section is that $\varphi \circ x$ is nondecreasing along the paths.

Theorem 3.1 Under Assumptions A1–A5, let $\mathcal{I} \subset \mathbb{R}$ be an open interval and let $\gamma : \mathcal{I} \rightarrow \mathcal{T}$ be a differentiable path such that

$$\dot{\gamma}(\lambda) = v(\gamma(\lambda)),$$

for all $\lambda \in \mathcal{I}$. Then, $\varphi(x(\gamma(\lambda)))$ is nondecreasing for all $\lambda \in \mathcal{I}$.

Proof Since γ is differentiable on \mathcal{I} , Lemma 3.1 implies (using Assumptions A1 and A2) that

$$\nabla_t x(\gamma(\lambda)) = -(H^{-1}K)(x(\gamma(\lambda)), \gamma(\lambda)).$$

Thus, using Assumption A4 and (9), the gradient of $\varphi(x(\gamma(\lambda)))$ with respect to λ is given by

$$\nabla_\lambda \varphi(x(\gamma(\lambda))) = W(x(\gamma(\lambda)), \gamma(\lambda))\dot{\gamma}(\lambda) = W(x(\gamma(\lambda)), \gamma(\lambda))v(\gamma(\lambda)).$$

By virtue of inequality (10) in Assumption A5, the latter expression is nonnegative. \square

Under the assumptions of Theorem 3.1, we say that $\varphi(x(t))$ is nondecreasing along the trajectories of $v(t)$.

3.3 Implementation

Under Assumptions A1–A5 the vector field method can be implemented using the following techniques. First, for any $t \in \mathcal{T}$ determine a reduced feasible set $\mathcal{R}(t) \subseteq \mathcal{X}$ that is guaranteed to contain the interior solution $x(t)$ of the parameterized optimization problem (1). Second, to satisfy Assumption A5, verify that the pseudogradient $W(x, t)$ is orientable on $\mathcal{R}(t)$. Note that for any given $t \in \mathcal{T}$ it may be possible to achieve orientability by premultiplying the evaluation function φ by some diagonal $(d \times d)$ -matrix of the form $M(t) = \text{diag}(m_1, \dots, m_d)(t)$ where $m_l(t) \in \{-1, 1\}$ for all $l \in \{1, \dots, d\}$. Indeed if $\hat{\varphi}(x, t) = M(t)\varphi(x)$, then it is possible to find a matrix $M(t)$ such that the pseudogradient $\hat{W}(x, t)$ corresponding to the modified evaluation function $\hat{\varphi}$,

$$\hat{W}(x, t) = M(t)W(x, t) = -M(t)\Phi(x)H^{-1}(x, t)K(x, t),$$

¹¹The trajectories of a vector field $v(t)$ exist and are unique on the whole domain \mathcal{T} if the vector field is Lipschitz there ([15], pp. 74–77), and in particular when v is continuously differentiable on \mathcal{T} .

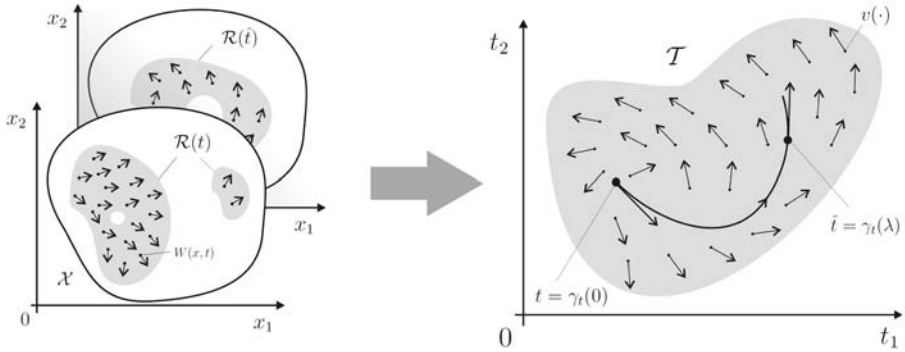


Fig. 1 Solution to the MCS Problem at t for $\varphi(x) = x_1$

is orientable at (x, t) . In particular, if a matrix M can be found that is independent of t , it may be advantageous for the analysis of the problem if the decision maker uses the evaluation function $\hat{\varphi}$ instead of φ . Third, find a vector field v that satisfies (10). To accomplish this, a systematic, algorithmic procedure to determine a vector field $v(t)$ that is “maximally aligned” with the pseudogradient $W(x, t)$ consists in solving the maximin problem¹²

$$v(t) \in \arg \max_{v \in \mathbb{R}^m: \|v\|=1} \left\{ \min_{x \in \mathcal{R}(t)} \left\{ \min_{1 \leq l \leq d} \langle W_l(x, t), v \rangle \right\} \right\}. \tag{11}$$

When the assumptions of the min-max theorem [16] are satisfied, any solution $v(t)$ to problem (11) also satisfies

$$v(t) \in \arg \min_{x \in \mathcal{R}(t)} \left\{ \max_{v \in \mathbb{R}^m: \|v\|=1} \left\{ \min_{1 \leq l \leq d} \langle W_l(x, t), v \rangle \right\} \right\}. \tag{12}$$

Motivated by the minimax formulation (12), since W is orientable by assumption, one obtains the set of candidate vector fields

$$\mathcal{V} = \left\{ \frac{W_1(\hat{x}^1(t), t)}{\|W_d(\hat{x}^1(t), t)\|}, \dots, \frac{W_d(\hat{x}^d(t), t)}{\|W_d(\hat{x}^d(t), t)\|} \right\}_{t \in \mathcal{T}}, \tag{13}$$

where for any $l \in \{1, \dots, d\}$ we have set

$$\hat{x}^l(t) = \arg \min_{x \in \mathcal{R}(t)} \|W_l(x, t)\|.$$

One can now check whether some element of \mathcal{V} is a suitable vector field on \mathcal{T} or possibly a subset thereof.

The above three steps can be iterated to tighten the reduced feasible set. It can also be useful to consider only subsets of the parameter space \mathcal{T} . Note that if an exact

¹² W_l is the l -th row vector of W , and $\langle \cdot, \cdot \rangle$ denotes the scalar product in the relevant Hilbert space.

solution to the maximization problem (5) is known for some $t \in \mathcal{T}$, then the MCS problem at t has a solution if and only if $W(x(t), t)$ is orientable at $(x(t), t)$.

Example 3.2 Consider a firm that has the option to invest in a number x of geographically dispersed markets¹³ (e.g., cities in the US) at an increasing convex cost $C(x) \geq 0$. For simplicity, the market price $p \in (0, 1)$ for the firm’s product is assumed to be the same in each market; it is announced nationally and is a parameter of the problem. Marginal production costs are zero. The demand in each market is $D(p) = 1 - p$, and the firm is risk averse with constant absolute risk aversion ρ . By investing in x markets the firm also reaps an increasing concave side benefit $B(x) \geq 0$ (e.g., through real-estate transactions).¹⁴ With probability $q \in (0, 1 - p)$ that is (at most) proportional to the quantity sold in each market, the firm incurs a unit loss in any market due to a liability claim. The firm’s expected payoffs are approximately¹⁵

$$\begin{aligned} \Pi(x, p, q) &= xpD(p) + B(x) - E[\tilde{L}|q] - \frac{\rho}{2}\text{Var}[\tilde{L}|q] - C(x) \\ &= \left(p(1 - p) - q - \frac{\rho q(1 - q)}{2} \right) x + B(x) - C(x), \end{aligned}$$

where the random variable $\tilde{L} \in \{0, 1, \dots, x\}$ represents the firm’s total losses. Maximizing profits we thus obtain $x^*(p, q) = \xi(p(1 - p) - q - \rho q(1 - q)/2)$, where ξ is the (increasing) inverse of $C' - B'$. Hence, the maximizer $x(t)$ is increasing in the parameter $t \in \mathcal{T} = \{(p, q) \in \mathbb{R}_{++}^2 : p + q < 1\}$ if and only if $p(1 - p) - q - \rho q(1 - q)/2$ is increasing. The corresponding pseudogradient at the optimum is

$$W(x(t)) = [-1 + \rho q - \rho/2, 1 - 2p] \xi'(p(1 - p) - q - \rho q(1 - q)/2),$$

so that, with $v(t) = (-p, q\rho/2)$, we obtain

$$\langle W(x(t)), v(t) \rangle = \frac{\rho \xi'(t)}{2} \left(\frac{2p}{\rho} + p + q - 4pq \right) \geq \frac{(4\rho - 1)\xi'(t)}{16} > 0,$$

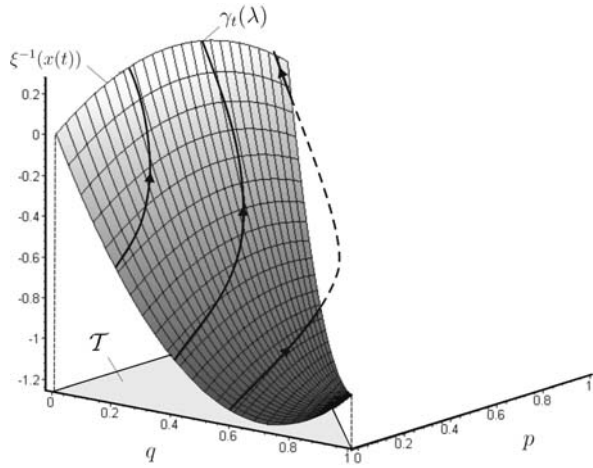
for all $t \in \mathcal{T}$ and $\rho > 1/4$. The vector field v solves the MCS problem on \mathcal{T} . In Example 4.1 we show how to obtain a global MCS reparameterization of the problem based on the vector field v , cf. Fig. 2. Let us remark that clearly in this example $\mathcal{R}(t) = \{x(t)\}$ if ξ^{-1} is known precisely (for any given B and C). However, our conclusions can be obtained without further specifying the firm’s payoffs. Note that it would have been possible to replace $p(1 - p) - q - \rho(1 - q)q/2$ by a scalar parameter λ , which, however, would decrease the resolution for a decision maker somewhat

¹³We allow x to take non-integer values.

¹⁴By imposing the Inada conditions $C'(0) < \infty$ and $B'(0) = \infty$ one can easily guarantee that the optimum is interior, i.e., $x(t) > 0$, so that we can without loss of generality set $\mathcal{X} = (0, \infty)$.

¹⁵The dependence of Π on the parameter ρ is not explicitly noted. In fact, in this problem the comparative statics with respect to ρ are obvious. We can thus use ρ itself in solving the MCS problem and finding an appropriate simple reparameterization, which illustrates an interesting “partial reparameterization” variant of our technique.

Fig. 2 Global MCS reparameterization in Examples 3.2 and 4.1



(λ is not the result of a one-to-one mapping from the parameter space), yet clearly provide trivial but precise monotone comparative statics.

3.4 Relation to Classic Supermodularity Results

We now derive a well-known supermodularity result as a particular case of Theorem 3.1. Under Assumption A1, we recall that $f(x, t)$ is *supermodular* in x if $\frac{\partial^2 f(x,t)}{\partial x_i \partial x_j}$ is nonnegative for all $(x, t) \in \mathcal{X} \times \mathcal{T}$ and $1 \leq i \neq j \leq n$. We call the function $f(x, t)$ *supermodular* in (x, t) if, in addition, $\frac{\partial^2 f}{\partial x_i \partial t_k}(x, t)$ is nonnegative for all $(x, t) \in \mathcal{X} \times \mathcal{T}$ and $1 \leq i \leq n, 1 \leq k \leq m$.

Corollary 3.1 *Suppose that Assumptions A1–A3 hold and that f is supermodular in (x, t) . Then, $x(t)$ is nondecreasing in t .*

Proof We show that $x(t)$ is nondecreasing in each component of $t = (t_1, \dots, t_m)$. Supermodularity in (x, t) implies that all components of $K(x, t)$ are nonnegative on $\mathcal{X} \times \mathcal{T}$. It also implies that $H_{ij}(x, t) \geq 0$ for all $i, j \in \{1, \dots, m\}$ with $i \neq j$ on $\mathcal{X} \times \mathcal{T}$. Without loss of generality we can restrict our attention to the reduced feasible set $\mathcal{R}(t) = (\mathcal{C} \cap \mathcal{D})(t)$. Hence $H_{ii}(x(t), t) \leq 0$ for all $i \in \{1, \dots, n\}$, for the Hessian matrix is negative definite at the optimum. Since $H_{ij}^{-1} = (-1)^{i+j} \det(H^{ji}) / \det(H)$, it is a simple linear algebra exercise to verify that $H_{ij}^{-1} \leq 0$ on $\mathcal{X} \times \mathcal{T}$ for all i, j in $\{1, \dots, n\}$.¹⁶ For any vector $v(t) > 0$, all entries of $-H^{-1}(x, t)K(x, t)v(t)$ are therefore nonnegative. As a result, Assumption A5 is satisfied for $\varphi(x) = x$ and $v(t) \equiv e_k$ where e_k is the k -th unit vector in the canonical basis of \mathbb{R}^m . An application of Theorem 3.1 with $\varphi(x) = x$ concludes the proof. \square

¹⁶The adjoint matrix H^{ij} is obtained by removing the i -th row and the j -th column from H .

An important case that is not currently dealt with in the monotone comparative statics literature is when f is supermodular in x but does not have the single-crossing property in (x, t) .¹⁷ In that context, Assumption A5 can be simplified as follows.

Assumption A5' For each $t \in \mathcal{T}$, there exists a vector $v(t) \in \mathbb{R}^m$ such that $K(x, t)v(t)$ is nonnegative for all $x \in \mathcal{R}(t)$, where $\mathcal{R}(t) \subseteq \mathcal{X}$ is an admissible reduced feasible set.

Corollary 3.2 *Suppose that Assumptions A1–A3 and A5' hold and that f is supermodular in x . Then, $x(t)$ is nondecreasing along the trajectories of v .*

Proof We modify the proof of Corollary 3.1. Supermodularity in x ensures that $H_{ij}^{-1} \leq 0$ on $\mathcal{X} \times \mathcal{T}$ for all $i, j \in \{1, \dots, n\}$. This together with Assumption A5' implies that all entries of $-H^{-1}(x, t)K(x, t)v(t)$ are nonnegative. A direct application of Theorem 3.1 with $\varphi(x) = x$ concludes the proof. \square

Corollary 3.2 applies to situations in which there are complementarities between the different decision variables, but not between decision variables and parameters. Since the maximizer under the original parameterization can be nonmonotonic, it is clear that Assumption A5' relaxes the tight single-crossing requirement put forward in [6] for the price of an MCS reparameterization of the problem.

3.5 Constrained Optimization Problems

Assumption A2 requires that the optimizer be in the *interior* of the feasible set \mathcal{X} . This assumption can be relaxed in different ways, *either* by reducing the dimensionality of the decision space (using a substitution approach for equality constraints), *or* by augmenting the dimensionality of the decision space (using a Lagrange-multiplier approach for equality and/or inequality constraints).

Substitution Approach Any equality constraints that are part of the definition of the feasible set \mathcal{X} in fact define a lower-dimensional set \mathcal{X}' that forms a submanifold of \mathcal{X} (with or without boundary). If the equality constraints can be solved globally for a number of decision variables, the parameterized optimization problem can be viewed as unconstrained on \mathcal{X}' after backsubstitution of these variables. More specifically, if \mathcal{X}' is diffeomorphic to an open subset of $\mathbb{R}^{\hat{n}}$ with $\hat{n} < n$, the problem can be seen as unconstrained on an open subset of $\mathbb{R}^{\hat{n}}$. To render our discussion precise, consider the problem¹⁸

$$\max_{x \in \mathcal{X}(t)} f(x, t), \quad (14)$$

¹⁷In particular, f is not supermodular in (x, t) .

¹⁸Note that in this formulation it is possible to have the feasible set depend on parameters. In contrast to standard MCS results obtained on lattices, we do not assume at the outset that $\mathcal{X}(t)$ is monotone in t with respect to the Veinott set order [6]. We are grateful to Pete Veinott for pointing out that his set order (originally termed “lower than” relation) was first introduced by him in a 1965 unpublished paper.

with

$$\mathcal{X}(t) = \{x \in \mathcal{Y} : g(x, t) = 0\},$$

where \mathcal{Y} is an open subset of \mathbb{R}^n , t belongs to an open set \mathcal{T} of \mathbb{R}^m , and g takes values in \mathbb{R}^k (for some $1 \leq k < n$) and is twice continuously differentiable. Suppose that the level set $g(x, t) = 0$ can be expressed explicitly as $(x_{n-k+1}, \dots, x_n)(t) = \tilde{g}(x_1, \dots, x_{n-k}, t)$, for $(x_1, \dots, x_{n-k}) \in \mathcal{X}'$, where \mathcal{X}' is the projection of \mathcal{Y} on the plane $\{(x_1, \dots, x_{n-k}, 0, \dots, 0) : (x_1, \dots, x_{n-k}) \in \mathbb{R}^{n-k}\} \subset \mathbb{R}^n$. The problem is then reduced to the following $(n - k)$ -variable unconstrained problem on \mathcal{X}' :

$$\begin{aligned} & \max_{(x_1, \dots, x_{n-k}) \in \mathcal{X}'} \tilde{f}(x_1, \dots, x_{n-k}, t) \\ & = \max_{(x_1, \dots, x_{n-k}) \in \mathcal{X}'} f(x_1, \dots, x_{n-k}, \tilde{g}(x_1, \dots, x_{n-k}, t), t). \end{aligned}$$

The application discussed in Sect. 5.3 provides an example of this transformation.

Lagrange-Multiplier Approach Equality constraints also can be approached with Lagrange multipliers. In the previous example, a necessary condition ([9], p. 255) for optimality is the existence of a k -dimensional vector v such that, at the optimum,

$$\nabla_x f(x, t) + v^T \nabla_x g(x, t) = 0.$$

Together with the k equations $g(x, t) = 0$, this determines the system of $n + k$ equations in $n + k + m$ variables x, t, v ,

$$G(x, t, v) = \begin{bmatrix} \nabla_x f(x, t) + v^T \nabla_x g(x, t) \\ g(x, t) \end{bmatrix} = 0.$$

The implicit function theorem implies that, if G_{xv} is invertible, then locally

$$\nabla_t(x, v)(t) = -[G_{xv}^{-1} G_t](x, v, t).$$

Even though the position of the optimal x and v in $\mathcal{X} \times \mathbb{R}^k$ is unknown, it might be possible to find directions in the parameter space such that $x(t)$ is nondecreasing in these directions. The following example illustrates this Lagrange-multiplier approach with equality constraints.

Example 3.3 Consider an economy with two goods $(x, y) \in \mathbb{R}_+^2$, with the production frontier $\{(x, y) \in \mathbb{R}_+^2 : g(x, y) = x^2 + y^2 - 1 = 0\}$ and a representative agent with utility $f(x, y) = u(x) + tv(y)$, where u, v are twice continuously differentiable, increasing and concave. We would like to determine the monotonicity properties of the optimizer $(x, y)(t)$ with respect to the parameter t . The constraint set is clearly not a lattice, hence classic supermodularity results do not apply directly.¹⁹ Using the

¹⁹However, classic results could be applied in conjunction with the substitution approach described earlier.

Lagrange-multiplier approach, we have

$$G(x, y, v, t) = \begin{bmatrix} u'(x) + 2vx \\ tv'(y) + 2vy \\ x^2 + y^2 - 1 \end{bmatrix},$$

which implies that

$$G_{x,y,v}(x, y, v, t) = \begin{bmatrix} u''(x) + 2v & 0 & 2x \\ 0 & tv''(y) + 2v & 2y \\ 2x & 2y & 0 \end{bmatrix}$$

and that

$$G_t(x, y, v, t) = \begin{bmatrix} 0 \\ v'(y) \\ 0 \end{bmatrix}.$$

An application of the implicit function theorem then yields

$$\frac{d}{dt} \begin{bmatrix} x \\ y \\ v \end{bmatrix} = \frac{v'(y)}{D} \begin{bmatrix} -4xy \\ 4x^2 \\ 2y(u''(x) + 2v) \end{bmatrix},$$

where $D = -4y^2(u''(x) + 2v) - 4x^2(tv''(y) + 2v)$. In this problem, the condition $\nabla f + v\nabla g = 0$ implies that v is negative, since the gradients of f and g both belong to the positive orthant of \mathbb{R}^2 . This, along with the concavity of u and v , implies that D is positive and that $\nabla_t x < 0 < \nabla_t y$, i.e., monotone comparative statics obtains.

Inequality constraints can be approached in a similar fashion. Consider again problem (14), this time with

$$\mathcal{X}(t) = \{x \in \mathcal{Y} : g(x, t) = 0, h(x, t) \leq 0\},$$

where \mathcal{Y} is an open subset of \mathbb{R}^n , t belongs to an open set \mathcal{T} of \mathbb{R}^m , g and h take values in respectively \mathbb{R}^k and \mathbb{R}^r (with $k < n$) and are both twice continuously differentiable. The Kuhn-Tucker necessary optimality conditions ([9], p. 284) imply the existence of adjoint variables v and μ in respectively \mathbb{R}^k and \mathbb{R}_+^r , such that

$$\nabla_x f + \langle v, \nabla_x g \rangle + \langle \mu, \nabla_x h \rangle = 0$$

and

$$\hat{\mu}_i(x, \mu, t) \equiv \mu_i(t)h_i(x, t) = 0,$$

for all $i \in \{1, \dots, r\}$. Letting

$$G(x, v, \mu, t) = \begin{bmatrix} \nabla_x f + v^T \nabla_x g + \mu^T \nabla_x h \\ g(x, t) \\ \hat{\mu}(x, \mu, t) \end{bmatrix},$$

a necessary optimality condition is that $G(x, v, \mu, t) = 0$. This defines a system of $n + k + r$ equations in $n + k + r + m$ variables, so that (if the relevant matrix is invertible) we can apply the implicit function theorem to compute $\nabla_t(x, v, \mu)$, and proceed as in the equality case. In some problems, it is possible to know in advance which inequality constraints are binding at the optimum. In this simple case, non-binding inequalities are ignored, while binding ones are treated as equality constraints. This approach is illustrated in the applications of Sects. 5.2 and 5.3.

4 Finding a Global MCS Reparameterization

Given a smooth solution $v : \mathcal{T} \rightarrow \mathbb{R}^m$ to the MCS Problem, it is interesting in practice to find an MCS reparameterization of the optimization problem (1). The idea is to start with the flow $\theta : \mathcal{F} \rightarrow \mathcal{T}$ induced by the vector field v and note that this flow is smooth and unique on what we refer to as the maximum “flow domain” $\mathcal{F} \subset \mathbb{R} \times \mathcal{T}$, beyond which the integral curves of the vector field cannot be extended. By taking a plane that is transverse (i.e., never collinear) to these integral curves, it is possible to construct new parameter coordinates under which monotone comparative statics obtain, at least locally.

4.1 Global Flows

As a consequence of the standard theory of ordinary differential equations (ODEs) [17], if the solution v is smooth, then integral curves to the vector field exist, are unique, and induce a smooth local flow θ .²⁰ Intuitively a flow describes how any point of the set \mathcal{T} is transported by a vector field v that is defined on \mathcal{T} . To formally define the well-known concept of a local flow, let us first introduce a *flow domain* $\mathcal{F} \subset \mathbb{R} \times \mathcal{T}$ with the property that, for any $t \in \mathcal{T}$, the set

$$\mathcal{F}^{(t)} = \{\lambda \in \mathbb{R} : (\lambda, t) \in \mathcal{F}\} \subset \mathbb{R}$$

is an open interval containing zero. A *local flow* on \mathcal{T} is a continuous map $\theta : \mathcal{F} \rightarrow \mathcal{T}$ that satisfies the two group laws,

$$\theta(0, t) = t, \tag{15}$$

for all $t \in \mathcal{T}$, and

$$\theta(\lambda, \theta(\mu, t)) = \theta(\lambda + \mu, t), \tag{16}$$

²⁰Existence and uniqueness of integral curves is also obtained when the vector field v merely satisfies a Lipschitz condition (cf. footnote 11).

for all $\lambda \in \mathcal{F}^{(t)}$ and $\mu \in \mathcal{F}^{(\theta(\lambda,t))}$ such that $\lambda + \mu \in \mathcal{F}^{(t)}$. In addition, for a given local flow θ , we define the *orbits*

$$\theta_\lambda(t) = \theta^{(t)}(\lambda) = \theta(\lambda, t),$$

for $(\lambda, t) \in \mathcal{F}$. A local flow (sometimes also referred to as a “one-parameter group action”) relates the vector field $v(t)$ for any $t \in \mathcal{T}$ to its orbits $\theta^{(t)}(\lambda)$. If the flow domain is such that the map θ is surjective (i.e., $\theta(\mathcal{F}) = \mathcal{T}$), then we call θ a *global flow*. The following result is standard, and will be of great use in the ensuing developments.

Theorem 4.1 (ODE Existence, Uniqueness, and Smoothness) *Let $v : \mathcal{T} \rightarrow \mathbb{R}^m$ be a smooth vector field. Consider the initial-value problem*

$$\nabla_\lambda \gamma_t(\lambda) = v(\gamma_t(\lambda)), \quad \gamma_t(\lambda_0) = t. \tag{17}$$

- (i) **EXISTENCE:** *For any $\lambda_0 \in \mathbb{R}$ there exist a nonempty open interval \mathcal{I} which contains λ_0 and an open set $\mathcal{U} \subset \mathcal{T}$, such that for any $t \in \mathcal{U}$ there is a smooth integral curve $\gamma_t : \mathcal{I} \rightarrow \mathcal{T}$ which solves (17) for all $\lambda \in \mathcal{I}$.*
- (ii) **UNIQUENESS:** *Any two smooth solutions to (17) agree on their common domain.*
- (iii) **SMOOTHNESS:** *Let $\mathcal{F} = \mathcal{I} \times \mathcal{U}$ as in (i). If we define the local flow $\theta : \mathcal{F} \rightarrow \mathcal{T}$ with $\theta(\lambda, t) = \gamma_t(\lambda)$, then θ is smooth.*

Proof See e.g., [18], pp. 452–459. □

In the following we use global flows corresponding to the solution vector field v of the MCS problem on \mathcal{T} (or a subset thereof) in order to find a global MCS reparameterization of problem (1) as previously indicated in Sect. 2

4.2 Coordinate Change in Parameter Space

By construction, if v is a solution to the MCS problem on \mathcal{T} , it is nonsingular everywhere. The lack of singular points allows us to give a canonical, local representation of v using a change of coordinates in \mathcal{T} . Consider a point t_0 of \mathcal{T} . If v is smooth, there exists an open interval $\mathcal{I} \subset \mathbb{R}$ containing the origin and an open subset \mathcal{U} of \mathcal{T} satisfying the conclusions of Theorem 4.1 (with $\lambda_0 = 0$). Moreover, since $v(t_0)$ is nonsingular, it uniquely determines an orthogonal hyperplane $\mathcal{H} \subset \mathbb{R}^m$ containing t_0 . Let

$$\mathcal{P} = \mathcal{U} \cap \mathcal{H} \subset \mathbb{R}^m.$$

Since \mathcal{H} is diffeomorphic to \mathbb{R}^{m-1} , \mathcal{P} can also be seen as an open subset of \mathbb{R}^{m-1} . To avoid confusion, let $\Pi = \pi(\mathcal{P})$ denote the image of \mathcal{P} under the diffeomorphism $\pi : \mathcal{H} \rightarrow \mathbb{R}^{m-1}$. Theorem 4.1 implies the existence of a smooth flow $\theta(\lambda, t)$, which we restrict to the domain $\mathcal{I} \times \mathcal{P}$. The flow can be reparameterized, with a slight abuse of notation, by $\theta(\lambda, \pi)$ on the product $\mathcal{S} = \mathcal{I} \times \Pi \subset \mathbb{R}^m$. Moreover, the corresponding range $\bar{\mathcal{T}} = \theta(\mathcal{S})$ is an open subset of \mathcal{T} by Theorem 4.1. Last, θ is one-to-one and smooth from \mathcal{S} to $\bar{\mathcal{T}}$, also by Theorem 4.1. We have therefore defined a local change of coordinates around t_0 : any t in the open neighborhood $\bar{\mathcal{T}}$ of t_0

can be uniquely expressed by a tuple $(\lambda_t, \pi(\psi(t))) \in \mathcal{S}$. The $(m - 1)$ -dimensional component $\psi(t)$ of t is the intersection of the trajectory going through t with the hyperplane \mathcal{H} . Moreover, since $t = \gamma_{\psi(t)}(\lambda_t)$, Theorem 3.1 implies that $\varphi \circ x$ is non-decreasing along the trajectories of the flow θ as λ increases. We have thus proved the following result.

Theorem 4.2 (Local MCS Reparameterization) *If v solves the MCS problem on \mathcal{T} , there exists a local MCS reparameterization of the form*

$$s = (s_1(t), s_2(t), \dots, s_m(t)) = (\lambda_t, \pi(\psi(t))), \tag{18}$$

around any $t_0 \in \mathcal{T}$, such that $\varphi(x(s))$ is nondecreasing in s_1 .

In practice, this result often can be applied globally, as the following example illustrates.

Example 4.1 In Example 3.2 we obtained a vector field of the form $v(t) = (-t_2, \alpha t_1)$ (for some $\alpha > 0$) as the solution to an MCS problem on some $\mathcal{T} \subset \mathbb{R}_+^2 \setminus \{0\}$. Since $v_2(t) = \alpha t_1 \neq 0$ on \mathcal{T} , we can choose the plane $\mathcal{P} = \{t \in \mathcal{T} : t = (t_1, 0)\}$, which is transverse to the vector field. The (global) flow of the vector field $v(t)$ is obtained by solving the initial-value problem (17), whence

$$\gamma_t(\lambda) = \theta_\lambda(t_1, t_2) = \left(t_1 \cos \sqrt{\alpha} \lambda - \frac{t_2 \sin \sqrt{\alpha} \lambda}{\sqrt{\alpha}}, t_1 \sqrt{\alpha} \sin \sqrt{\alpha} \lambda + t_2 \cos \sqrt{\alpha} \lambda \right).$$

Thus, for any $(s, 0) \in \mathcal{P}$, we obtain

$$\theta_\lambda(s, 0) = (s \cos \sqrt{\alpha} \lambda, s \sqrt{\alpha} \sin \sqrt{\alpha} \lambda),$$

for $\lambda \in (0, \pi/(2\sqrt{\alpha}))$ and $s > 0$. Hence, on any contractible²¹ compact subset $\bar{\mathcal{T}}$ of \mathcal{T} , we obtain the global reparameterization $t \mapsto (\lambda, s)$ with $\lambda = \frac{1}{\sqrt{\alpha}} \arctan \frac{t_2}{\sqrt{\alpha} t_1}$ and $s = \sqrt{t_1^2 + (t_2^2/\alpha)}$; cf. Fig. 2. In the context of Example 3.2, the key insight for the decision maker from the MCS reparameterization is that the optimal number of markets to invest in varies monotonically in the ratio p/q , i.e., the product price in relation to the risk of liability lawsuits.

We now formally generalize the reparameterization technique used in the local case by providing a general condition under which Theorem 4.2 holds globally.

Assumption A6 (Transverse Hypersurface) There exist subsets $\bar{\mathcal{T}}$ and \mathcal{P} of \mathcal{T} , with $\bar{\mathcal{T}}$ open and $\mathcal{P} \subset \bar{\mathcal{T}}$, such that:

- (i) There exists a convex, open subset Π of \mathbb{R}^{m-1} and a diffeomorphism π mapping \mathcal{P} onto Π .

²¹An m -dimensional open set with nonempty interior is contractible if it is homotopy equivalent (i.e., it can be deformed via a continuous transformation) to an m -dimensional open ball. Intuitively, contractible sets have no ‘‘holes.’’

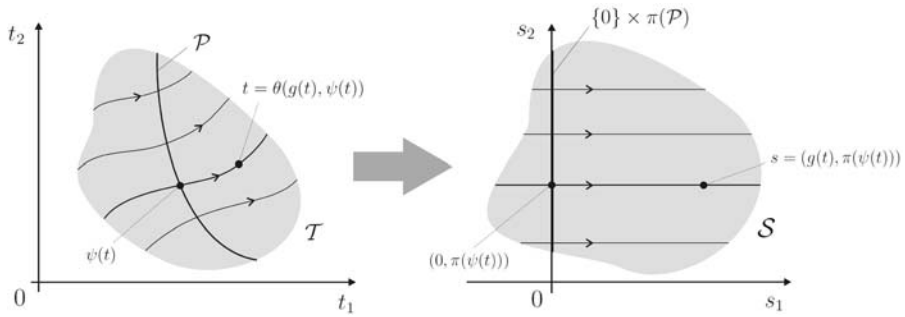


Fig. 3 Global MCS reparameterization

(ii) For each $t \in \tilde{T}$, $\gamma_t(\mathcal{F}^{(t)}) \cap \mathcal{P}$ is a singleton $\{\psi(t)\}$.

This last assumption ensures that trajectories of an MCS vector field v lead to a foliation of the subset \tilde{T} of \mathcal{T} . In other words, the existence of a set \mathcal{P} of points, each element of which can be associated with exactly one trajectory, allows projection of the set \tilde{T} onto \mathcal{P} and—via the length (from t to $\psi(t)$) of the trajectory (which could pass outside \tilde{T})—obtain a bijection between \tilde{T} and a set $\mathcal{S} \subset \mathbb{R}^m$. This bijection corresponds to the desired global MCS reparameterization containing the length of the MCS trajectories as one new parameter in which the solution to the reparameterized problem (4) (when evaluated with φ) varies monotonically.

Theorem 4.3 (Global MCS Reparameterization) *If the vector field v is a solution to the MCS problem on \tilde{T} and if Assumption A6 holds, then there exists a global MCS reparameterization of the form*

$$s = (s_1(t), s_2(t), \dots, s_m(t)) = (\lambda_t, \pi(\psi(t))) \tag{19}$$

such that $\varphi(x(s))$ is nondecreasing in s_1 .

Proof Each element t of \tilde{T} uniquely determines an element $\psi(t)$ of \mathcal{P} and a real λ_t such that $\gamma_{\psi(t)}(\lambda_t) = t$. The set $\mathcal{S} = \bigcup_{\psi \in \mathcal{P}} \mathcal{F}^{(\psi)} \times \{\pi(\psi)\}$ is an open subset of \mathbb{R}^m . Moreover, the mapping $t \mapsto s = (\lambda_t, \pi(\psi(t))) \in \mathcal{S}$ is one-to-one from \tilde{T} to \mathcal{S} , and smooth by Theorem 4.1. Last, since $t(s) = (\lambda, \psi)$ follows the trajectories of v as λ increases, Theorem 3.1 implies that $\varphi(x(s)) = \varphi(x(\lambda, \pi))$ is nondecreasing in $s_1 = \lambda$. □

When v has a potential²² $u : \mathcal{T} \rightarrow \mathbb{R}$, a good candidate for \mathcal{P} in Assumption A6 is any iso-potential that crosses all trajectories.²³ Although this need not always be the case, it is likely that iso-potentials will be diffeomorphic to an open subset of \mathbb{R}^{m-1} , and to a convex subset if one chooses \tilde{T} carefully. The following classic theorem

²²That is, v is the gradient of u .

²³In that case, trajectories cross \mathcal{P} only once, because the potential increases along them.

(essentially Poincaré’s lemma) gives a necessary and sufficient condition for the existence of a potential, which can be tested on any vector field v satisfying Assumption A5.

Theorem 4.4 (Existence of a Potential) *Let $\bar{T} \subset T$ be a contractible compact domain with nonempty interior and let v be a vector field on \bar{T} . Then, $\partial v_k / \partial t_l = \partial v_l / \partial t_k$, for all $k, l \in \{1, \dots, m\}$ with $k \neq l$, if and only if there exists a twice continuously differentiable potential $u : T \rightarrow \mathbb{R}$, i.e.,*

$$v(t) = \nabla_t u(t),$$

for all $t \in T$.

Proof See [10], Vol. II, p. 296. □

5 Applications

5.1 Optimal Capacity Choice and Product Distribution

Suppose that a firm can choose the quantity (or, equivalently, the quality) q of a product that will be provided to each of m geographically dispersed consumers. Each consumer $k \in \{1, \dots, m\}$ is located at a point $t_k \in \mathbb{R}$, representing his or her type. To deliver the product to consumer i the firm incurs a quadratic transportation (or, customization) cost $d(z - t_k)^2$, where $z \in \mathbb{R}$ is the location of a distribution center that the firm is able to determine freely, and d is a positive constant. The firm’s unit transportation cost from its factory (located at the origin) to the distribution center is given by the smooth convex function $C(z)$ with $C'(0) = C(0) = 0$. To keep our analysis simple, we assume that each consumer’s demand can be represented by a linear inverse demand function of the form $a - bq$, where a, b are positive constants.²⁴ The firm’s profit can therefore be written as

$$\Pi(q, z, t) = mq(a - bq) - dq \sum_{k=1}^m (t_k - z)^2 - mqC(z), \tag{20}$$

where $t = (t_1, \dots, t_m) \in \mathbb{R}^m$. The firm thus solves the parameterized optimization problem

$$\max_{(q, z) \in \mathbb{R}_+ \times \mathbb{R}} \Pi(q, z, t). \tag{21}$$

Provided that a is sufficiently large, it is easy to verify that problem (21) has a unique interior solution $(q^*(t), z^*(t))$ and satisfies Assumptions A1–A3. The firm’s MCS problem is: *how does the optimal per-consumer production quantity q^* vary with t ?*

²⁴This corresponds to a quadratic utility function $u_k(q) = aq - bq^2/2$ for all consumers $k \in \{1, \dots, m\}$. Allowing for demand heterogeneity with $u_k(q) = a_kq - b_kq^2/2$ leads to analogous results.

Starting from a market characterized by the parameter vector $t \in \mathbb{R}^m$, is there a direction in the parameter space in which the optimal production quantity q^* increases? We first determine the first-order necessary optimality conditions for (21),

$$\frac{\partial \Pi}{\partial q} = a - 2bq - d \sum_{k=1}^m (t_k - z)^2 - mC(z) = 0 \tag{22}$$

and

$$\frac{\partial \Pi}{\partial z} = 2dq \sum_{k=1}^m (t_k - z) - mC'(z) = 0. \tag{23}$$

Therefore, we can restrict $\mathcal{R}(t)$ to the subset of tuples $(q, z) \in \mathcal{X} = \mathbb{R}_+ \times \mathbb{R}$ that satisfy (22) and (23). To apply the method, we also compute Φ , H^{-1} and K . Since $\varphi(q, z) = q$, we have that $\Phi = (1, 0)$ and Assumption A4 is satisfied. Moreover,

$$H = \begin{bmatrix} -2b & 2d \sum_{k=1}^m (t_k - z) - mC'(z) \\ 2d \sum_{k=1}^m (t_k - z) - mC'(z) & -2mdq - mC''(z) \end{bmatrix}.$$

The first-order necessary optimality condition with respect to z , given in (23), simplifies the Hessian matrix H to

$$H = \begin{bmatrix} -2b & 0 \\ 0 & -2mdq - mC''(z) \end{bmatrix}.$$

It follows that²⁵

$$H^{-1} = \frac{1}{2bmq(2d + C''(z))} \begin{bmatrix} -2mdq - mC''(z) & 0 \\ 0 & -2b \end{bmatrix},$$

whence

$$-\Phi H^{-1} = \left(\frac{1}{2b}, 0 \right).$$

The first row of K is given by²⁶

$$\left(\frac{\partial^2 \Pi}{\partial q \partial t_1}, \dots, \frac{\partial^2 \Pi}{\partial q \partial t_m} \right) = 2d((z - t_1), \dots, (z - t_m)).$$

Therefore, in order to satisfy Assumption A5, we are looking for a vector $v(t) \in \mathbb{R}^m$ such that

$$\frac{d}{b} \langle (z - t_1, \dots, z - t_m), v(t) \rangle \geq 0,$$

²⁵Observe that the determinant is nonzero, since $C''(z) \geq 0$ by convexity of g .

²⁶There is no need to compute the second row, since K is left-multiplied by $-\Phi H^{-1}$, whose second component is zero.

for all $z \in \mathbb{R}$, or equivalently

$$\langle (ze - t), v(t) \rangle \geq 0, \tag{24}$$

for all $z \in \mathbb{R}$, where $e = (1, \dots, 1)/m$ is the unit vector of the first bisectrix²⁷ Δ in \mathbb{R}^m . It is easy to see that if $v(t)$ is orthogonal to e and has a nonnegative scalar product with $-t$, the condition is satisfied. The vector $v(t) = -t + m\langle t, e \rangle e$ is such that first, $\langle e, v(t) \rangle = 0$, since $\langle e, e \rangle = 1/m$; and second, $\langle -t, v(t) \rangle = \langle t, t \rangle - m(\langle t, e \rangle)^2 \geq 0$ by the Cauchy-Schwarz inequality. Moreover, the inequality is strict if t is not collinear with e . Last, observe that, when seen from t , $v(t)$ points directly²⁸ to the first bisectrix Δ of \mathbb{R}^m . We therefore conclude from Theorem 3.1 that $q^*(t)$ increases as t gets closer to Δ . In other words, as the consumer types become “closer,” the optimal product quantity increases. When the consumer types are identical (t collinear to e), the optimal production reaches its maximum. The problem can thus be reparameterized in the following way: define a cylinder \mathcal{P} around the first bisectrix Δ , for example

$$\mathcal{P} = \left\{ t \in \mathbb{R}^{m-1} : d(t, \Delta) = 1 \right\},$$

where d is the Euclidian distance from a point to a line. This cylinder is an $(m - 1)$ -dimensional manifold, which can be parameterized by $m - 1$ components. Moreover, \mathcal{P} is transverse to all trajectories, and is hit by all trajectories once, so that Assumption A6 is satisfied.²⁹ Therefore, we have a global reparameterization of \mathbb{R}^m where $m - 1$ components correspond to the position on the cylinder and determine a radius emanating from Δ , and the remaining component is a parametric representation of the radius. In this particular context, it is possible to construct a more efficient parameterization: let \mathcal{H} denote the $(m - 1)$ -dimensional hyperplane of \mathbb{R}^m orthogonal to the first bisectrix and going through the origin, and $(\epsilon_1, \epsilon_2, \dots, \epsilon_{m-1})$ be an orthogonal basis of \mathcal{H} . Then, $(e_1, e_2, \dots, e_m) = (\epsilon_1, \epsilon_2, \dots, \epsilon_{m-1}, e)$ is an orthogonal basis of \mathbb{R}^m . Moreover, if t is represented with respect to that basis, i.e., $t = s_1 e_1 + s_2 e_2 + \dots + s_m e_m$, we obtain a new parameterization of the parameter space such that $q^*(s_1, \dots, s_m)$ is nonincreasing in (s_1, \dots, s_{m-1}) (the smaller these coordinates, the closer is t to the first bisectrix).³⁰

It is worth observing that in order to solve the MCS problem, we relied on our knowledge of the first-order optimality condition for z . It is our second use of the crucial device $\mathcal{R}(t)$ which enables us to narrow down the domain on which Assumption A5 must be satisfied. On the other hand, we did not use the fact that the optimizer (q^*, z^*) satisfies the first-order optimality condition for q . Thus $\mathcal{R}(t)$ could have been larger without affecting our ability to construct the vector field v to satisfy Assumption A5. We also note that classic supermodularity is of no use in this

²⁷The first bisectrix is defined by the equation $t_1 = t_2 = \dots = t_m$.

²⁸The vector $v(t)$ points in the direction of the orthogonal projection of t on Δ .

²⁹To be rigorous, \mathcal{P} is diffeomorphic to $\mathcal{S}^{m-2} \times \mathbb{R}$, where \mathcal{S}^{m-2} is the unit sphere in \mathbb{R}^{m-1} . This parameterization is a generalization of cylindrical coordinates in \mathbb{R}^3 .

³⁰Moreover, it can be shown that $q^*(s)$ is independent of the last component, s_m .

problem, since q^* is not monotonic in any of the t_k 's. Last, observe a remarkable fact in our analysis of this example: we are able to obtain monotone comparative statics for q^* without solving explicitly for either q^* or z^* . In general, the method can be used to derive monotone comparative statics for any single decision variable, say, x_1 , while one is able to solve the optimization problem explicitly for some other variables, say, $x_k(t), x_{i+1}(t), \dots, x_n(t)$. In that case the reduced feasible set $\mathcal{R}(t)$ can be narrowed down to the set of all $x \in \mathcal{X}$ such that $x_k = x_k(t), \dots, x_n = x_n(t)$.

5.2 Neoclassical Production

Consider a firm's optimal choice of factor inputs, capital k and labor l , so as to maximize the objective function

$$f(x, t) = g(k, l) - rk - wl, \quad (25)$$

where $x = (k, l)$ and $t = (r, w)$ with r the rate of return of capital and w the average wage rate. As pointed out by [6], if g is not supermodular, comparative statics are not monotone in the original parameterization. To demonstrate the use of our method, we assume that g is twice continuously differentiable and that there exists a unique optimizer in the interior of \mathbb{R}_+^2 . Therefore, Assumptions A1–A3 are satisfied, with the Hessian and cross-derivative matrices

$$H = \begin{bmatrix} g_{kk} & g_{kl} \\ g_{kl} & g_{ll} \end{bmatrix} \quad \text{and} \quad K = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

respectively. The pseudogradient is therefore

$$W = \frac{1}{D} \begin{bmatrix} g_{ll} & -g_{kl} \\ -g_{kl} & g_{kk} \end{bmatrix},$$

where $D(k, l) = (g_{kk}g_{ll} - g_{kl}^2)(k, l)$ is the determinant of $H(k, l)$. Because of the strict concavity of g at the optimizer, we can restrict the reduced feasible set $\mathcal{R}(t)$ to the subset of \mathbb{R}_+^2 where H is negative definite, implying that $D(k, l)$ is positive.³¹ Notice that the pseudo-gradient and the reduced feasible set are independent of r and w .³² To simplify our exposition we drop the explicit dependence on t and refer to $\mathcal{R}(t)$ as \mathcal{R} . If one can find a vector v making a positive scalar product with $W(k, l)$ for all (k, l) , it will satisfy Assumption A5 for all values of r and w . The vector field will then consist of straight, parallel trajectories of direction v . This will generate a linear reparameterization of the problem (the basis of the new coordinate system consisting of v and any other vector not collinear to v), under which both k and l are nondecreasing in the first parameter coordinate. Before addressing the problem of monotone comparative statics for both k and l , let us consider the simpler problem

³¹The determinant D is positive at any maximizer of (25) as the product of the two negative eigenvalues of H .

³²Note that H , K and D are all independent of (r, w) .

of finding monotone comparative statics for k alone. That is, we consider the function $\varphi(k, l) = k$, which trivially satisfies Assumption A4. The pseudogradient then becomes

$$W_k(k, l) = \Phi(k, l)W(k, l),$$

with $\Phi(k, l) = (1, 0)$. This yields

$$W_k(k, l) = \frac{1}{D}(g_{ll}, -g_{kl}).$$

We are looking for a vector $v \in \mathbb{R}^2 \setminus \{0\}$ such that $\langle W_k(k, l), v \rangle \geq 0$, or equivalently

$$\langle (g_{ll}, -g_{kl}), v \rangle \geq 0. \tag{26}$$

Since g_{ll} is nonpositive, a solution is $v = (-1, 0)$. That is, $k(r, w)$ is nondecreasing in r . Notice that this result obtains without assumption on g except for smoothness.³³ In general, $k(r, w)$ is not monotonic in w : this would require $g_{kl} \geq 0$ for all (k, l) (as can be seen by substituting $v = (0, -1)$ in (26)). However, there may be other directions of v such that k is nondecreasing. Equation (26) can be rewritten as

$$v_1 g_{ll}(k, l) - v_2 g_{kl}(k, l) \geq 0,$$

for all $(k, l) \in \mathbb{R}_+^2$. Since $g_{ll} \leq 0$ on \mathcal{R} , this is equivalent to

$$v_1 \leq \min\{D, d\}v_2, \tag{27}$$

where $(d, D) = (\inf \delta, \sup \Delta)$, with

$$\delta = \left\{ \frac{g_{kl}}{g_{ll}}(k, l) : g_{kl} \geq 0, g_{ll} < 0, (k, l) \in \mathcal{R} \right\},$$

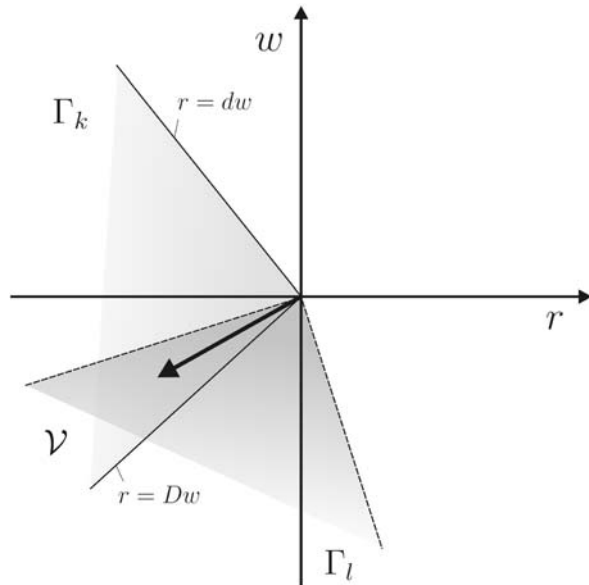
$$\Delta = \left\{ \frac{g_{kl}}{g_{ll}}(k, l) : g_{kl} \leq 0, g_{ll} < 0, (k, l) \in \mathcal{R} \right\},$$

as well as the conventions that $\inf\{\emptyset\} = +\infty$ and $\sup\{\emptyset\} = -\infty$. When $\Delta \neq \emptyset$, $D \geq 0$. Similarly, $d \leq 0$ if $\delta \neq \emptyset$. When g is supermodular, Δ is empty or reduced to the singleton $\{0\}$, so that $D \leq 0$. Moreover, $\delta \neq \emptyset$ implies $d \leq 0$, so that condition (27) is satisfied by any $v \in \mathbb{R}_-^2$, by virtue of the nonpositivity of $d \wedge D$. This proves that $k(r, w)$ is nonincreasing not only in r , but also in w , whenever g is supermodular. In general, relation (27) defines a convex cone $\Gamma_k \in \mathbb{R}^2$ based at the origin which always contains the negative real line $\mathbb{R}_- \times \{0\}$. Except when both $D = +\infty$ and $d = -\infty$, Γ_k has a nonempty interior. If $g_{kl} > 0$, $D = -\infty$, implying that Γ_k is a half-space that is located below the line $v_1 = dv_2$.

³³ Another way to see this is the following: the function $f(k, l, r, w) = g(k, l) - rk - wl$ is supermodular in (r, k) . To apply standard supermodularity results, define $F(k, r, w) = \max_{l \geq 0} f(k, l, r, w)$. F is supermodular in (k, r) and $\arg \max_{k \geq 0} F(k, r, w) = k(r, w)$. This implies that $k(r, w)$ is nondecreasing in r . We thank Paul Milgrom for this observation.

The optimizer $k(r, w)$ is nondecreasing in any direction of Γ_k . When Γ_k has a non-empty interior, it is possible to change coordinates in the parameter space by using two independent basis vectors in Γ_k . As pointed out earlier, this coordinate change is global, since Γ_k is independent of the particular values chosen for r and w . Similarly, $l(r, w)$ is nondecreasing in any direction located in the cone Γ_l based at the origin and containing the negative imaginary line $\{0\} \times \mathbb{R}_-$. Having constructed these two cones, we can now address the more challenging question, *is it possible to find directions in which both k and l increase?* The answer depends on $\mathcal{V} = \Gamma_k \cap \Gamma_l$. If \mathcal{V} is empty, we cannot construct any direction that jointly increases k and l . If \mathcal{V} is non-empty, then it is also a convex cone, whose elements are directions of joint increase. The intersection \mathcal{V} being empty does not prove the nonexistence of directions of joint increase. It just means that we do not have enough information on the optimizers to produce such directions. As our information gets richer, the set \mathcal{R} becomes narrower, which implies that the cones Γ_k and Γ_l become wider. When one has enough information, the cones are wide enough to intersect, yielding the desired directions of joint increase (cf. Fig. 4).³⁴ When \mathcal{V} is nonempty, the vector field can be chosen constant: $v(r, w) = v$ for some $v \in \mathcal{V}$. In that case, any straight line \mathcal{P} orthogonal to v satisfies Assumption A6: it is transverse and is hit exactly once by all trajectories. The reparameterization is then simple: take any vector e_2 on that line and let $e_1 = v$. Then, (e_1, e_2) is an orthogonal basis of \mathbb{R}^2 , such that if one expresses $t = (r, w)$ on that basis (that is, $(r, w) = s_1 e_1 + s_2 e_2$), then $(k(s), l(s))$ is nondecreasing in s_1 .

Fig. 4 Neoclassical production: monotonicity of $k(r, w)$ on $\Gamma_k \cap \Gamma_l$



³⁴The analysis for obtaining directions of joint decrease is naturally analogous, by considering opposite directions. Similarly, it is possible to obtain directions of increase in one parameter and decrease in the other.

5.3 Giffen Goods

In an economy with two goods, an agent wishes to maximize her increasing and concave utility by solving

$$\max_{(x,y) \in \mathbb{R}_+^2} u(x, y),$$

subject to

$$px + qy \leq w,$$

where x, y are the quantities of the two goods, p and q are their respective positive prices, and w represents the agent’s wealth. Without loss of generality, we select the second good to be the numéraire and correspondingly set $q = 1$. In addition, since both goods are desirable, the agent’s budget constraint is binding. As pointed out in Sect. 3.5, the agent’s problem can be restated as

$$\max_{x \in [0, w/p]} u(x, w - px).$$

In this formulation the problem has one decision variable and two parameters.³⁵ If u is smooth, compactness ensures the existence of an optimizer. We assume that for our starting values of w and p , the optimizer $x(w, p)$ is unique, located in $(0, \frac{w}{p})$, and that $u(\cdot, w, p)$ is locally strictly concave at $x(w, p)$ and everywhere twice continuously differentiable. This implies that Assumptions A1–A3 are satisfied. The first good is “normal” if $x(w, p)$ is nonincreasing in p , and Giffen or “inferior” if this monotonicity is sometimes violated.³⁶ Similarly, one would usually expect the consumption of any good to increase with the agent’s wealth. However, this monotonicity is also sometimes violated. The question then becomes, *under what conditions is a good normal, and how are price and wealth effects connected?* The Hessian and cross-derivative matrices are

$$H = u_{xx} - 2pu_{xy} + p^2u_{yy} \quad \text{and} \quad K = [u_{xy} - pu_{yy} - xu_{xy} + pxu_{yy} - u_y].$$

The pseudogradient W is therefore given by

$$W(x, w, p) = \frac{1}{D}[\alpha(x, w, p) - x\alpha - u_y(x, w - xp)],$$

where $D = -H > 0$ (by strict concavity of H at the optimizer) and $\alpha(x, w, p) = u_{xy} - pu_{yy}$. We also note that since $\varphi(x) = x$, Assumption A4 is trivially satisfied.

³⁵As mentioned earlier, we have therefore converted the initial problem—with two decision variables on a one-dimensional manifold of \mathbb{R}^2 (with empty interior but nonempty relative interior)—into a problem with one decision variable on a set with nonempty interior.

³⁶Classic examples include potatoes or bread. The gist of the argument goes as follows: when the price for bread increases, poorer people cannot afford buying “luxury goods” such as meat, and end up consuming more bread, which is still the cheapest good. Other goods violating this monotonicity are Veblen goods [19] or positional goods ([20], Chap. 3) which are such that the implied status of the owner increases with their price.

In order to meet Assumption A5, we are thus looking for a vector $v \in \mathbb{R}^2$ such that

$$\alpha(x, w, p)v_1 - (x\alpha(x, w, p) + u_y)v_2 \geq 0, \tag{28}$$

for all x in the reduced feasible set $\mathcal{R}(w, p)$. First, we observe that if u is supermodular and concave in its second variable, α is nonnegative, which implies, along with the nonnegativity of u_y , that any vector v in $\mathbb{R}_+ \times \mathbb{R}_-$ solves (28). This means that if the two goods are complements and if the utility function is concave in the second good, then the first good is normal.³⁷ In the general case, we show that there is a hierarchical relationship between wealth and price effects. Increasing wealth amounts to setting $v_1 > 0$ and $v_2 = 0$, so that the good is normal with respect to the wealth effect if and only if $\alpha \geq 0$. On the other hand, $\alpha \geq 0$ implies that $x\alpha + u_y \geq 0$. Since decreasing the price amounts to setting $v_2 > 0$ and $v_1 = 0$, the good is therefore normal with respect to the price effect if $\alpha \geq 0$. This shows that the following result holds for any smooth, nondecreasing utility function: if an augmentation in wealth increases the optimal consumption of a good, then a cut in its price also increases its optimal consumption. In general, the converse is not true.³⁸ If the optimal consumption $x(w, p)$ is known or constrained to belong to some subinterval $\mathcal{J} = (x_1, x_2) \subset (0, w/p)$,³⁹ the analysis can be refined. For example, suppose that $\min_{x \in \mathcal{J}} \{\alpha(x, w, p)\} \geq 0$ for all w, p . Then the good is normal with respect to both wealth and price effects (any $v \in \mathbb{R}_+ \times \mathbb{R}_-$ solves (28)). If α sometimes takes negative values but $\min_{x \in \mathcal{J}} \{x\alpha(x, w, p) + u_y(x, w - px)\} \geq 0$, then the good is normal with respect to price effects ($v \in \{0\} \times \mathbb{R}_-$ solves (28)). More generally, suppose that $\alpha(\cdot, w, p)$ changes sign only once on (x_1, x_2) , and that $\alpha(x_1, w, p) > 0$ for all w and p in an open neighborhood of initial values of wealth and price. The second condition means that the good is normal for low consumption, while the first condition means that the good becomes Giffen for high consumption values. Then, if the vectors

$$(\alpha(x_1, w, p), -x_1\alpha(x_1, w, p) + u_y(x_1, w - px_1))$$

and

$$(\alpha(x_2, w, p), -x_2\alpha(x_2, w, p) + u_y(x_2, w - px_2))$$

are in the same half-plane, there exists a normal vector $v(w, p)$ of the half-space whose scalar product with $W(x, w, p)$ is nonnegative for all $x \in (x_1, x_2)$. The situation is represented in Fig. 5.

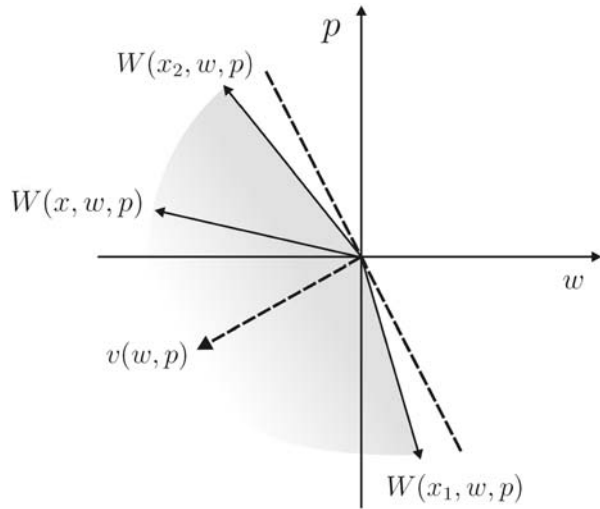
It is easy to verify that $v(w, p)$ can always be taken in the negative orthant \mathbb{R}_-^2 . Therefore we have the following result: if the pseudo-gradients of consumption boundaries x_1, x_2 lie in the same half-plane, and if the good behaves as a normal good for low consumption values and as a Giffen one for high consumption values

³⁷This result can also be shown by observing that the concavity of u in y implies the supermodularity of v in $(x, w, -p)$.

³⁸The result can be read in the opposite direction: if a good is inferior with respect to price effect, it is also inferior for wealth effect.

³⁹For example, minimal consumption could be imposed or supply could be limited.

Fig. 5 Giffen goods: solving the MCS problem for $x(w, p)$



(such as potatoes, cf. footnote 36), then there exists a way to increase optimal consumption of this good by reducing both wealth and price at the same time. It can also be shown that in the same situation, there is no way to increase consumption by raising both wealth and price at the same time.

5.4 Multiattribute Screening

A variation of our method can be used in the context of screening with multiple instruments. A firm faces customers of different types, distributed on an interval $\mathcal{X} \subset \mathbb{R}$ according to a positive density function g . The firm sells products whose m attributes are described by a vector $t \in \mathcal{T} \subset \mathbb{R}^m$. The goal of the firm is to propose a product line $\Gamma \subset \mathcal{T}$ and a price schedule $P : \Gamma \rightarrow \mathbb{R}_+$ (with $0 \in \Gamma$ and $P(0) = 0$) that maximizes its expected profit

$$\pi(P, \Gamma) = \int_{\mathcal{X}} [P(t(x)) - C(t(x))] g(x) dx,$$

where $C(t(x))$ is the cost of producing $t(x) \in \Gamma$ and $t(x)$ solves the type- x consumer’s utility maximization problem

$$t(x) \in \arg \max_{t \in \Gamma} \{u(x, t) - P(t)\}.$$

This general screening problem has been only solved for particular cases.⁴⁰ When $m = 1$ (only one instrument available), it is possible to directly compute the optimal price schedule $P(x)$ under some additional supermodularity assumptions on the

⁴⁰See [21–23] for the multiattribute, one-dimensional type case, and [24] for a recent account of the general multidimensional screening literature.

primitives of the problem. Assuming that u is smooth, we can define

$$\mu(x, t) = u(x, t) - u_x(x, t) \frac{1 - G(x)}{g(x)},$$

where G is the probability distribution of the density g . [25] has shown, based on a technique developed by [26], that if u and μ are twice differentiable and supermodular (i.e., $u_{xt} \geq 0$ and $\mu_{xt} \geq 0$ on $\mathcal{X} \times \mathcal{T}$), then it is possible to construct the optimal price schedule. Moreover, this optimal schedule leads to “perfect” screening (without bunching): each consumer of type x will buy a distinct product $t(x)$. Suppose now that there are $m \geq 2$ product attributes available. Our method can be extended to build product lines that perfectly screen consumers. Defining the pseudogradient W as the $(2 \times m)$ -matrix

$$W(x, t) = \begin{bmatrix} \nabla_x^T u_t(x, t) \\ \nabla_x^T \mu_t(x, t) \end{bmatrix},$$

suppose that there exists for all $t \in \mathcal{T}$ a nonzero vector $v(t)$ such that

$$W(x, t)v(t) \geq 0,$$

for all $x \in \mathcal{X}$ (i.e., Assumption A5 is satisfied). We can then define $\Gamma \subset \mathcal{T}$ to be the image of any smooth trajectory $\gamma : (0, 1) \rightarrow \mathcal{T}$ generated by the vector field v . This leads to a reparameterization of the utility u and the function μ when restricted to $\mathcal{X} \times \Gamma$ (that is, when customers are offered the product line Γ). Specifically, we define \tilde{u} and $\tilde{\mu}$ on $\mathcal{X} \times (0, 1)$ by $\tilde{u}(x, \lambda) = u(x, \gamma(\lambda))$ and $\tilde{\mu}(x, \lambda) = \mu(x, \gamma(\lambda))$. Using Lemma 3.1 and Theorem 3.1, we can show that \tilde{u} and $\tilde{\mu}$ are supermodular on $\mathcal{X} \times (0, 1)$. The aforementioned result then implies that it is possible to find the optimal price schedule on Γ , and that this schedule perfectly screens customers. This approach does not solve the original problem of maximizing the profit on \mathcal{T} , since we artificially restricted ourselves to the product line Γ . However, the method can be repeated for several distinct trajectories, and leads to a perfectly screening price schedule that maximizes the expected profit not only on a particular product line, but on a large class of product lines that spans the whole multiattribute space \mathcal{T} .

6 Discussion

In the available literature on monotone comparative statics, the parameterization of the optimization problem is essentially taken as given.⁴¹ The presently known criteria for the monotonicity of solutions hold, therefore, only with respect to the particular problem formulation given at the outset. Milgrom and Shannon’s [6] characterization of the monotonicity of solutions to (1) on lattices requires the objective function f to

⁴¹Note that the decision variables are also typically taken as given. Our method in principle allows for a change of the decision variables to obtain monotone comparative statics through an appropriate choice of the evaluation function φ .

be quasi-supermodular in x and to satisfy a single-crossing property in (x, t) .⁴² The supermodularity requirement on the objective function can thereby be interpreted in terms of “complementarity” of decision variables, a concept that dates back at least to Edgeworth [28] and whose origins are reviewed by [29]. Reference [30] demonstrates the power of complementarities and associated supermodularity properties in interpreting decision changes as monotone responses to exogenous shifts of economic conditions. Even though equilibria cannot be located exactly, complementarities allow one to make precise statements about the direction in which optimal decisions move as a consequence of parameter changes. In the absence of such complementarities, the presently available theory unfortunately guarantees the non-monotonicity of solutions, even though this non-monotonicity might just be a symptom of an unsuitable parameterization of the problem. This paper proposes a way to obtain an equivalent formulation of the optimization problem (1) using a new parameterization, such that—provided sufficient knowledge about the location of the solution—monotone comparative statics may be obtained. Finding a new parameterization of the problem amounts to creating a set of economic indicators which allow for monotonic decision making and thus easy rules of thumb (i.e., when the relevant indicator goes up, the optimal decision goes up, too). This seems especially useful in situations where the same optimization problem needs to be solved repeatedly for different parameter values.

Let us briefly mention at this point that our method naturally extends to equilibrium problems (cf. also [31]) specified by a relation

$$F(x, t) = 0,$$

where $F : \mathcal{X} \times \mathcal{T} \rightarrow \mathbb{R}^n$ is a continuously differentiable function, as can be seen by associating $\nabla_x f$ with F , so that $H = \nabla_x F$ and $K = \nabla_t F$. With these substitutions in place, all of our results hold essentially without modification. We also remind the reader that, as pointed out in Sect. 3.5, even though through Assumption A2 we require the existence of a unique interior global optimum, parameter-dependent constraints can be accommodated in a straightforward way by shifting the analysis to a submanifold in \mathcal{X} or by augmenting the space of decision variables by Lagrange multipliers corresponding to the binding constraints.

Sometimes our method may also be useful for reducing the number of parameters without any losses. To show this, let us first note that, clearly, an “ideal” parameterization of problem (1) for the case $m \geq n$ associates exactly one parameter t_i with each component of the decision variable x_i , and is such that x_i remains unchanged in response to a change of parameter t_j (with $j \neq i$). If the location of the optimal solution is perfectly known, a reparameterization with these “ideal” properties can always be obtained by setting $s_i = x_i(t_i)$ for $i \in \{1, \dots, n\}$ and simply discarding all other $m - n$ parameters, t_{n+1}, \dots, t_m . Unfortunately, lack of knowledge about the location of the optimizer (up to a monotone transformation) usually makes this trivial solution impracticable. Nevertheless, it may sometimes be possible to reduce the

⁴²Reference [27] applies these results to expected-utility maximization problems under uncertainty and finds necessary and sufficient conditions on the model primitives in that context.

number of parameters by finding directions $v(t)$ for which $\langle W(x, t), v(t) \rangle = 0$ for all x in an admissible reduced feasible set $\mathcal{R}(t)$. If such a direction can be found, the solution does not depend on λ in the associated global MCS reparameterization and λ may thus be discarded from the set of new parameters. By repeating this process it may be possible to eliminate further parameters.

The tradeoff between the decision maker's knowledge about the location of the optimal action and her ability to find directions that guarantee monotone behavior of the optimizer (i.e., to solve the (local) MCS problem) is related to "partially specified problems", as discussed by [13]. One type of partially specified optimization problems possesses an objective function of the form $f(x, t) = g(x, t) + \delta(x)$, where δ is any affine mapping from $\mathcal{X} \subset \mathbb{R}$ to \mathbb{R} . Monotonicity of optimal solutions $x_\delta(t)$ to the problem (1) for *any* affine δ is then equivalent to the (otherwise unknown) function g being supermodular on $\mathcal{X} \times \mathcal{T}$. The key idea in this approach is that the class of perturbations δ is large enough relative to g and \mathcal{X} to allow for any location of the optimizer in \mathcal{X} . A variation of our method bypasses this definitive result when the function g fails to be supermodular: it might be possible to transform the parameter space so as to "supermodularize" the function g . If $g(x, t)$ is not supermodular in (x, t) , we can build trajectories $\gamma : (0, 1) \rightarrow \mathcal{T}$ in the parameter space such that $g(x, \gamma(\lambda))$ is supermodular in (x, λ) . We have also used this approach in Sect. 5.4 and Corollary 3.2. While supermodularization of functions is just a particular application of our method, its repeated use in this paper suggests its potential benefits in numerous other settings, such as for supermodularizing noncooperative games.

Further research could proceed to relax some of the differentiable structure imposed to obtain our results. Systematic MCS reparameterizations can be expected to generalize naturally to an analysis on lattices. The problem is to find a rule on $\mathcal{W} \subset \mathcal{T} \times \mathcal{T}$ such that $(t, t') \in \mathcal{W}$ implies $\phi(x(t')) \geq \phi(x(t))$. In particular, suppose that we can build trajectories $\{\mathcal{T}_i\}$ in \mathcal{T} such that $(x, t) \mapsto f(x, t)$ has the single-crossing property⁴³ on $\mathcal{X} \times \mathcal{T}_i$ for all $i \in \mathcal{I}$. If f is in addition (quasi-)supermodular in x , standard results as in [13] apply, showing that $x(t)$ is nondecreasing along the trajectories. The problem is of course to construct such trajectories. Our method for doing so is based on differential calculus, but there may be other ways to build trajectories (e.g., through discretization of our results), or at least to find rules in the parameter space, in order to achieve monotone comparative statics.

References

1. Samuelson, P.A.: The stability of equilibrium: comparative statics and dynamics. *Econometrica* **9**(2), 97–120 (1941)
2. Bonnans, J.F., Shapiro, A.: *Perturbation Analysis of Optimization Problems*. Springer, New York (2000)
3. Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation I*. Springer, New York (2006)
4. Topkis, D.M.: *Ordered optimal solutions*. Doctoral Dissertation, Stanford University, Stanford, CA (1968)

⁴³Any trajectory is totally ordered, with the order implied by the parametric description of the curve.

5. Topkis, D.M.: Supermodularity and Complementarity. Princeton University Press, Princeton (1998)
6. Milgrom, P., Shannon, C.: Monotone comparative statics. *Econometrica* **62**(1), 157–180 (1994)
7. Baumol, W.J., Quandt, R.E.: Rules of thumb and optimally imperfect decisions. *Am. Econ. Rev.* **54**(2), 23–46 (1964)
8. Granot, F., Veinott, A.F.: Substitutes, complements and ripples in network flows. *Math. Oper. Res.* **10**(3), 471–497 (1985)
9. Bertsekas, D.P.: *Nonlinear Programming*. Athena Scientific, Belmont (1995)
10. Zorich, V.A.: *Mathematical Analysis*, vols. I, II. Springer, New York (2004)
11. Sard, A.: The measure of the critical points of differentiable maps. *Bull. Am. Math. Soc.* **48**, 883–890 (1942)
12. Sard, A.: Images of critical sets. *Ann. Math.* **68**(2), 247–259 (1958)
13. Milgrom, P.: Comparing optima: do simplifying assumptions affect conclusions? *J. Political Econ.* **102**(3), 607–615 (1994)
14. Berge, C.: *Topological Spaces*. Oliver and Boyd, Edinburgh (1963). Reprinted by Dover Publications, Mineola, in 1997
15. Khalil, H.K.: *Nonlinear Systems*. Macmillan, New York (1992)
16. Kakutani, S.: A generalization of Brouwer’s fixed point theorem. *Duke Math. J.* **8**, 457–459 (1941)
17. Arnold, V.I.: *Ordinary Differential Equations*. MIT Press, Cambridge (1973)
18. Lee, J.M.: *Introduction to Smooth Manifolds*. Springer, New York (2003)
19. Veblen, T.: *The Theory of the Leisure Class* (1899). Reprinted by Penguin Books, New York, in 1994
20. Hirsch, F.: *Social Limits to Growth*. Harvard University Press, Cambridge (1976)
21. Roberts, K.W.S.: Welfare considerations of nonlinear pricing. *Econ. J.* **89**(353), 66–83 (1979)
22. Mirman, L.J., Sibley, D.S.: Optimal nonlinear prices for multiproduct monopolies. *Bell J. Econ.* **11**(2), 659–670 (1980)
23. Matthews, S., Moore, J.: Monopoly provision of quality and warranties: an exploration of the theory of multidimensional screening. *Econometrica* **55**(2), 441–467 (1987)
24. Rochet, J.-C., Stole, L.A.: The economics of multidimensional screening. In: Dewatripont, M., Hansen, L.-P., Turnovsky, S.J. (eds.) *Advances in Economics and Econometrics: Theory and Applications: Eighth World Congress*, vol. I, pp. 150–197. Cambridge University Press, New York (2003)
25. Mussa, M., Rosen, S.: Monopoly and product quality. *J. Econ. Theory* **18**(2), 301–317 (1978)
26. Mirrlees, J.A.: An exploration in the theory of optimal income taxation. *Rev. Econ. Stud.* **38**(2), 175–208 (1971)
27. Athey, S.: Monotone comparative statics under uncertainty. *Q. J. Econ.* **117**(1), 187–223 (2002)
28. Edgeworth, F.Y.: *The Pure Theory of Monopoly* (1897). Reprinted in: Edgeworth, F.Y. *Papers Relating to Political Economy*, Macmillan, London, 1925
29. Samuelson, P.A.: Complementarity: an essay on the 40th anniversary of the Hicks-Allen revolution in demand theory. *J. Econ. Lit.* **12**(4), 1255–1289 (1974)
30. Milgrom, P., Roberts, J.: The economics of modern manufacturing: technology, strategy, and organization. *Am. Econ. Rev.* **80**(3), 511–528 (1990)
31. Milgrom, P., Roberts, J.: Comparing equilibria. *Am. Econ. Rev.* **84**(3), 441–459 (1994)