Generalized Monotonicity Analysis*

Bruno H. Strulovici[†] Thomas A. Weber[‡]

Abstract

Complex economic models often lack the structure for the application of standard techniques in monotone comparative statics. Generalized Monotonicity Analysis (GMA) extends the available methods in several directions. First, it provides a way of finding parameter moves that yield monotonicity of model solutions. Second, it allows studying the monotonicity of functions or subsets of variables. Third, GMA naturally provides bounds on the sensitivity of variables to parameter changes. Fourth, GMA may be used to derive conditions under which monotonicity obtains with respect to functions of parameters, corresponding to imposed parameter moves. Fifth, GMA contributes insights into the theory of comparative statics, for example, with respect to dealing with constraints or exploiting additional information about the model structure. Several applications of GMA are presented, including constrained optimization, non-supermodular games, aggregation, robust inference, and monotone comparative dynamics.

Keywords: Aggregation, Comparative Statics, Comparative Dynamics, Monotone Comparative Statics, Parameterized Equations, Parameter Transformation, Quantitative Monotonicity Analysis, Robust Inference, Supermodular Games.

JEL-Classification: C61, C43, C72, D11.

*We are very grateful to David Greenstreet, Paul Milgrom, John Quah, Kevin Roberts, and an anonymous reviewer for helpful comments and discussion.

†Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2600. E-mail: b-strulovici@northwestern.edu.

‡Department of Management Science and Engineering, 442 Terman Engineering Center, Stanford University, Stanford, CA 94305-4026. E-mail: webert@stanford.edu.

1 Introduction

The comparison of model predictions, typically described in terms of solutions to an optimization or equilibrium problem, for different parameter values is the subject of comparative statics (Hicks 1939; Samuelson 1941). Of particular interest is the monotonicity of solutions in model parameters, giving rise to *monotone comparative statics* (MCS). Until the appearance of ordinal methods, the standard tool of MCS was, beyond direct computation of solutions, the implicit function theorem (IFT). The introduction of ordinal methods, particularly by Topkis (1968), Milgrom and Shannon (1994), and Athey (2002), has allowed researchers to characterize MCS in terms of sufficient (and in some sense necessary) conditions on the primitives of the model.

The standard approach to MCS takes the formulation of the problem (which represents the model) as given. It cannot deliver positive results when monotonicity does not obtain with respect to the original variables and parameters. However, interesting monotonicity properties may generally be found once one broadens the horizon to the monotonicity of *functions of variables* with respect to *functions of parameters.* This paper presents a simple method, called *Generalized Monotonicity Analysis* (GMA), to systematically uncover and analyze monotonicity properties of solutions to optimization or equilibrium problems. In GMA, parameters are allowed to move simultaneously in the parameter space, and monotonicity is investigated for functions of model solutions. The approach, which is fundamentally geometric, also sheds new light on existing comparative-statics results. GMA introduces several concepts, such as 'pseudo-gradients' and 'monotonicity directions,' which may help applied economists to disentangle various issues that arise when trying to analyze the monotonicity properties of their models, even when model complexity rules out the use of standard MCS tools. The GMA approach raises questions beyond standard comparative-statics analysis. For example, one may ask what structure is required for the variables of a problem to be monotonic in some parameter aggregate (Section 4.3). In problems where parameters represent heterogeneity across some agents, one may ask what notion of "increased homogeneity" improves their incentives (Section 4.2). More generally, Section 3 shows that GMA discloses *all* monotonicity properties relating solutions of a problem to its parameters given the available information (Theorem 1).

GMA extends comparative statics in yet another direction, namely the study of the *rates* at which solutions vary with parameters. This application of GMA is referred to as *quantitative monotonicity analysis.* As one of the many potential applications of quantitative monotonicity analysis, we employ it in Section 4.4 to obtain rate constraints in a well-known problem of robust inference concerned with the empirical determination of firms' productivity levels.

The gist of GMA can be described as follows. Consider a model described by the parameterized equation

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

where $x \in \mathbb{R}^n$ is the variable, $t \in \mathbb{R}^m$ is the parameter, and f takes values in $\mathbb{R}^n$. For any t, let $x(t)$ denote the solution to the above system, which, for this introductory exposition, is assumed to be unique. For each t, let $R(t) \subset \mathbb{R}^n$ denote an *information set* known to contain the solution: $x(t) \in R(t)$. Assume, still for this introduction, that the Jacobian matrix $D_x f$ of f with respect to x is everywhere invertible. For each nonzero vector v of $\mathbb{R}^m$, consider the vector

$$W(x,t,v) = -[D_x f]^{-1}(x,t)(\partial f/\partial v),$$

where $\partial f/\partial v$ is the directional derivative of f in direction v. We call this vector a *pseudo-gradient* along direction v, because it coincides with the *actual* gradient of $x(t)$ with respect to v if one substitutes $x = x(t)$ in the expression for $W(x,t,v)$. The family $\bar{W}(t,v|R) = \{W(x,t,v)\}_{x\in R(t)}$ of all pseudo-gradients on the information set contains essential information about solution sensitivity to parameters: even though the exact location of $x(t)$ may be unknown, the reaction of $x(t)$ to a parameter change in direction v is contained in $\bar{W}(t,v|R)$. The key, then, is to determine properties of $\bar{W}(t,v|R)$ that permit statements about which parameter moves will affect solutions monotonically. The simplest case is when $\bar{W}(t,v|R)$ is a singleton. Suppose, for example, that $\bar{W}(t,e_j|R) = \{w_j(t)\}$, where e_j is the j-th basis vector of $\mathbb{R}^m$. In general, $w_j(t)$ may have some negative components: moving from t along the arbitrary direction e_j does not necessarily increase all components of $x(t)$. However, suppose that we can find real numbers $\lambda_j(t)$ such that $\sum_j \lambda_j(t)w_j(t)$ is a positive vector. Then, moving from parameter t in direction $v(t) = \sum_j \lambda_j(t)e_j$ increases all components of the solution $x(t)$, at least locally. We call such directions of parameter changes *monotonicity directions* at t. It turns out that the set $V(t|R)$ of all monotonicity directions at t is a convex cone,[1] which is referred to as the *GMA-cone* at t. In general, $\bar{W}(t,v|R)$ contains multiple elements. One can show that the smaller the set $\bar{W}(t,v|R)$ (in the inclusion order) and the larger the GMA-cone $V(t|R)$

[1]The cone may be empty. In that case, given one's information about the sensitivity of x to the parameters, it is impossible to construct parameter moves that guarantee an increase in the solution (cf. Theorem 1).

at t, the easier it becomes to find parameter moves along which solutions are monotonic. Selecting a vector field of monotonicity directions (contained in $V(t|R)$ at each t) yields trajectories along which a model solution (or smooth selections thereof, in the general case of a set-valued solution) is necessarily monotonic. — Instead of considering the monotonicity of a model solution directly, GMA can be applied to a *function* of the solution, such as an average (discussed in Section 4.3). Lastly, we stress that the method permits the natural use of additional information about solutions through the information set $R(t)$. For example, if the equation $f(x,t) = 0$ represents the first-order necessary optimality condition of a maximization problem, then $R(t)$ may consist of all x that also satisfy the second-order condition, $D_x f(x,t) \leq 0$. A smaller information set $R(t)$ results in smaller pseudo-gradient sets $\bar{W}(t, v|R)$, and hence in larger GMA-cones $V(t|R)$ (Theorem 2). Note that instead of its straightforward use to merely guarantee the monotonicity of model solutions along parameter paths, GMA can also be employed to *influence* the *rate* at which solutions vary with parameters, a more quantitative approach to MCS developed in Section 3.5.

When analyzing monotonicity properties of solutions to a given problem, results naturally depend on the concept of monotonicity that is being used. For example, a comparative-statics analysis based on ordinal methods usually describes the monotonicity of solutions in the strong set order.[2] In noncooperative games with strategic complementarities, equilibrium monotonicity is typically limited to the smallest and largest equilibria (Milgrom and Roberts 1990). However, Echenique (2002) shows that, in such games, monotonicity results can be extended to all *stable* equilibria. Allowing for less structured (e.g., non-supermodular) environments, we allow for any *differentiable* (or at least right-differentiable) selection of the solution set $X(t)$ and propose a systematic way to investigate properties of such a selection. GMA can generate statements of the following nature.

> *Each differentiable solution is nondecreasing as parameters are moved in this direction or along that trajectory.*

The differential nature of such statements may or may not restrict the strength of the results, and can sometimes be relaxed to yield more general statements. The following five

[2] $X(s) \leq X(t)$ in the strong set order (Veinott 1989) if and only if $(x, y) \in X(s) \times X(t)$ implies that $(x \wedge y, x \vee y) \in X(s) \times X(t)$, where $x \wedge y$ denotes the componentwise minimum and $x \vee y$ the componentwise maximum of the two vectors x and y.

points further clarify how GMA can be applied. First, as mentioned earlier, GMA may be used as an exploratory step to elicit monotonicity properties of a complex problem. Directly applying ordinal methods to a complex problem may often require much ingenuity or expertise, or may simply fail if one restricts attention to the original variables and parameters. In contrast, the method proposed here is based on elementary differential calculus and provides a helpful start to explore monotonicity. The results of this first step may then be derived under more general conditions using ordinal methods.[3] In general, one should think of differential and ordinal methods as complements, whose strengths can be combined to yield sharp and general results when analyzing the monotonicity of solutions to an optimization or equilibrium problem. Second, in many problems the solution is generically unique, so that the question of which selections to consider naturally disappears when $X(t)$ is a singleton that is differentiable in t. If there are multiple solutions that are locally unique, GMA considers monotonicity of each corresponding selection, and thus necessarily also the most interesting ones (e.g., stable selections).[4] Third, in equilibrium analysis, one may be interested precisely in equilibrium evolutions that are smooth with respect to parameters. Indeed, ruling out solutions that suddenly jump after an infinitesimal parameter move may be a desirable assumption when analyzing equilibrium monotonicity.[5] Fourth, GMA easily accommodates the introduction of additional constraints to discriminate between solutions. For example, nonnegativity of solutions may be imposed as an additional inequality to the information set $R(t)$. Fifth, monotonicity can be weak or strict, or even stronger. While ordinal methods yield weak monotonicity, the differential nature of GMA makes is possible to establish strict monotonicity results.[6] As noted earlier, GMA can be used in the form of 'quantitative monotonicity analysis' to obtain explicit bounds for the response rate of model solutions to changes in parameters.[7]

GMA extends the boundaries of MCS conceptually by providing a systematic method for investigating the interplay between model structure and available information (including voluntary restrictions) about model solutions. The approach is constructive, as it endogenously generates the largest cones of directions in the parameter space that are compatible with monotonicity.[8] Parameter directions found this way are likely to have an intuitive

[3]For example, such exploration was used in Quah and Strulovici (2008).

[4]The transversality theorem provides simple conditions under which differentiability of solutions obtains generically, as a consequence of the invertibility of $D_x f$ (Guillemin and Pollack 1974, p. 68).

[5]Similar assumptions are often made in optimization problems, cf. Persico (2000, Assumption A1).

[6]Edlin and Shannon (1998) focus on strict monotonicity of solutions. Their analysis is based on differentiability and first-order conditions.

[7]See Section 3.5 and the application in Section 4.4.

[8]Jensen (2007) also considers general cones for comparative statics. However, in his analysis the

interpretation (for example, they can express a notion of "proximity," as in Section 4.2), which can be difficult to determine otherwise.

The paper builds on Strulovici and Weber (2008), providing more general foundations and results for monotonicity analysis, examining new issues such as quantitative monotonicity analysis and robustness, and proposing economic applications that illustrate the versatility of the method. Section 2 introduces the concepts of the paper. Section 3 contains core theoretical results. Section 4 provides applications of GMA to constrained optimization, non-supermodular games, aggregation theory, quantitative monotonicity analysis, and monotone comparative dynamics. Section 5 concludes.

2 Concepts and Notation

The GMA approach seeks to establish the monotonicity of a 'criterion' in certain parameter movements. The (monotonicity) criterion is a function of a variable that solves a 'primitive equation.' The approach can accommodate additional solution requirements which are not encoded in the primitive equation, but are available in the form of an 'information structure.' An information structure describes a subset of solutions to the primitive equation. The method proceeds to compute 'pseudo-gradients' to estimate the criterion change in the different parameter directions. These estimates are then used to obtain 'monotonicity directions' in the parameter space, along which a criterion change is guaranteed to be nonnegative.

Primitive Equation. Let $X(t)$ denote the set of solutions $x(t)$ to the *primitive equation*

$$f(x,t) = 0, \tag{1}$$

where the (*decision*) *variable* x lies in a smooth manifold $\mathcal{X} \subseteq \mathbb{R}^n$, the *parameter* t lies in a convex subset $\mathcal{T}$ of $\mathbb{R}^m$, and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ (with $n \geq 1$) is a continuously differentiable (smooth) function.[9]

Monotonicity Criterion. Instead of considering only the monotonicity of solutions $x(t)$ to the primitive equation, we introduce a *monotonicity criterion* $\hat{x} = \phi(x)$, where ϕ is a

direction cones are treated as exogenous primitives.

[9]All of our results can easily be extended to allow for infinite-dimensional parameters, although, for simplicity, we focus on the case with finite-dimensional parameter vectors. Some of the results also generalize to infinite-dimensional variables, as in Section 4.5.

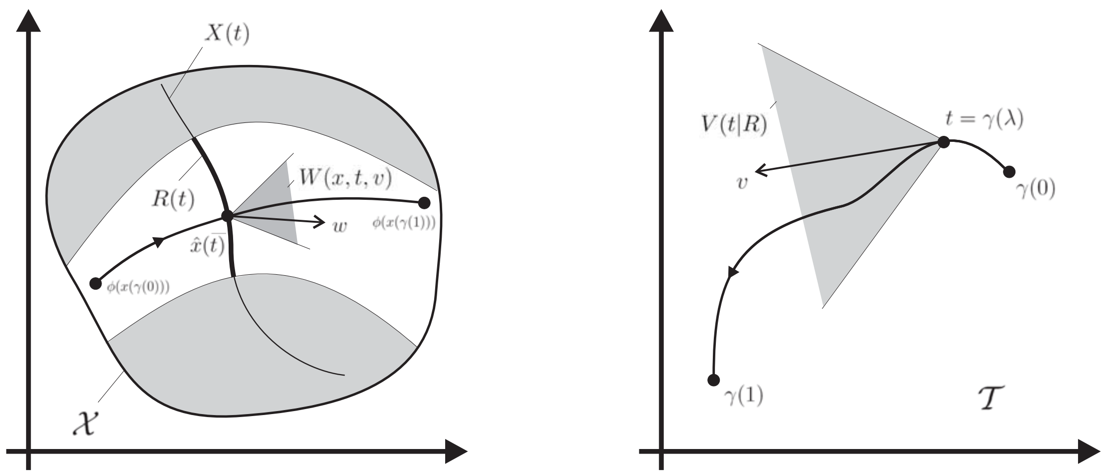

Figure 1: GMA Primitives (for $\hat{x} = \phi(x) \equiv x$).

d-dimensional differentiable function of x (the notation $\hat{x}$ will sometimes be used instead of $\phi(x)$ for simplicity). For example, monotonicity of the first component of $x(t)$ can be examined by setting $\phi(x) = x_1$. The criterion can also be used to examine the monotonicity of aggregates of the original variables. If, for example, $x = (x_1, \ldots, x_n)$ represents consumption decisions of n individuals, then the comparative statics of aggregate consumption can be examined by choosing a criterion of the form $\phi(x) = x_1 + \cdots + x_n$. To avoid obscuring the exposition of the GMA method, we often de-emphasize the presence of the criterion ϕ by simply writing $\hat{x}$ instead of $\phi(x)$. Analyzing the monotonicity of the original variables amounts to setting $\phi(x) = x$ or, equivalently, $\hat{x} = x$.

Information Structure. For any $t \in \mathcal{T}$, let $R(t) \subset X(t)$ denote the subset of solutions to (1) which satisfy a number of additional constraints, given a parameter value of t. Such additional constraints may include nonnegativity, second-order conditions, or stability requirements, depending on whether (1) derives from an optimization or an equilibrium problem.[10] The collection $R = \{R(t)\}_{t \in \mathcal{T}}$ defines a parameterized *information structure* of the problem. The crucial point is that, when searching for monotonicity directions, one may be able to exploit not only the fact that any element $x(t)$ of $R(t)$ solves the primitive equation (by definition), but also that it satisfies the additional requirements implied by the inclusion $x(t) \in R(t)$. A map $x : t \mapsto x(t) \in R(t)$ is a *selection* of R. We say that R is a *refinement* of R', denoted by $R \subseteq R'$, if $R(t) \subseteq R'(t)$ for all t.

[10] GMA (particularly Theorem 3 and Corollary 1) can be extended to conditions that cannot be summarized by a subset of $\mathcal{X}$, for example conditions involving derivatives of $x(t)$. Such conditions can be useful to narrow down the set of pseudo-gradients, as in the example of Section 4.1.

Symbol	*Dimension*	*Interpretation*
$x \in \mathcal{X}$	n	Variable
$\hat{x} = \phi(x) \in \phi(\mathcal{X})$	$d \in \{1, \ldots, n\}$	Monotonicity criterion
$t \in \mathcal{T}$	m	Parameter
$v \in \mathbb{R}^m$	m	Parameter direction
$w(x,t) \in \mathbb{R}^n$	n	Pseudo-gradient of f at (x,t) in direction v, satisfies (2)
$\phi : \mathcal{X} \to \mathbb{R}^d$	d	Criterion function
$\gamma : [0,1] \to \mathcal{T}$	m	Path in the parameter space
$R(t)$	n	Information set at t (=subset of $X(t)$)
$X(t)$	n	Set of solutions to (1) at t
$V(t\|R)$	m	GMA-cone at t given R, defined in (4)
$W(x,t,v)$	n	Set of pseudo-gradients at (x,t) in direction v
$\bar{W}(t,v\|R)$	n	Set of pseudo-gradients at t in direction v ($= \cup_{x \in R(t)} W(x,t,v)$)
$\mathcal{X} \subset \mathbb{R}^n$	n	Set of feasible decisions (=manifold)
$\mathcal{T} \subset \mathbb{R}^m$	m	Parameter set (and the underlying vector space)

Table 1: Summary of Notation.

Parameter Paths. A continuously differentiable map $\gamma : [0,1] \to \mathcal{T}$ is called a *path* of the parameter space. A parameter $t \in \mathcal{T}$ is *on* γ if $t = \gamma(\lambda)$ for some $\lambda \in [0,1]$. A path γ *starts at* $t \in \mathcal{T}$ if $\gamma(0) = t$. It starts at t *in the direction* $v \in \mathbb{R}^m$ if $\dot{\gamma}(0_+) = d\gamma/d\lambda|_{\lambda=0_+} = v$. A path γ that starts at t in the direction v will be referred to as a (t,v)*-path.* A selection x of R is γ*-differentiable* if $\lambda \mapsto x(\gamma(\lambda))$ is differentiable on $(0,1)$, and is right-differentiable at 0.

Pseudo-Gradients. A *pseudo-gradient* of f at (x,t) in direction v is a vector w in $\mathbb{R}^n$ such that

$$D_x f(x,t)\, w + D_v f(x,t) = 0 \tag{2}$$

for some $x \in R(t)$. The set of pseudo-gradients is denoted by $W(x,t,v)$. This is the set of potential gradients assumed by any smooth selection of R, as the parameter moves from t in direction v. Since any solution to (1) at t which may be of interest is localized to the information set $R(t)$, the set of all possible pseudo-gradients for a parameter movement from t in direction v is

$$\bar{W}(t,v|R) = \{w \in W(x,t,v) : x \in R(t)\}. \tag{3}$$

The set $\bar{W}(t,v|R)$ contains all vectors of $\mathbb{R}^n$ which *may* be *actual* gradients of some differential solution to (1) as the parameter moves from t in direction v.

Monotonicity Path. γ is a *monotonicity path* if $\hat{x} = \phi(x)$ is nondecreasing along γ for any γ-differentiable selection x.

Monotonicity Directions. Given an information structure R and a parameter $t \in \mathcal{T}$, let

$$V(t|R) = \{v \in \mathbb{R}^m : D\phi(x)w \geq 0,\ x \in R(t),\ w \in W(x,t,v)\} \tag{4}$$

denote the set of parameter directions in which the monotonicity criterion is nondecreasing.[11] As is shown in the next section, $V(t|R)$ forms a convex cone, containing the initial directions of all monotonicity paths starting at t. It will henceforth be referred to as the *GMA-cone* at t.

Table 1 summarizes the notation, and Figure 1 depicts the main GMA primitives.

3 Theoretical Results

In this section, we first show that the set of pseudo-gradients contains all actual gradients of smooth selections of R (Proposition 1). We then establish that the set $V(t|R)$ of monotonicity directions is a convex cone (Proposition 2), that this cone contains all directions guaranteeing monotonicity of ϕ evaluated at the solutions (Theorem 1, the "necessity" part), and that the cone widens as R becomes more informative (Proposition 2). We then establish that trajectories following directions in the GMA-cone are monotonicity paths for ϕ (Theorem 3, the "sufficiency" part), and provide a simple characterization of monotonicity when $D_x f$ is invertible (Corollary 1). After summarizing the different steps of GMA, we show how information contained in R may be folded into the primitive equation (1), which establishes a form of equivalence between various initial descriptions of any given problem. The analysis is then modified to consider several other important concepts of monotonicity: strict monotonicity and quantitative monotonicity (Propositions 4 and 5). The latter concept, which to the best of our knowledge is new, can be used to derive "robust" comparative statics, as illustrated by Section 4.4 in the context of productivity estimation. Another theoretical aspect of the method, concerning the inclusion of constraints, is treated as an illustrative application in Section 4.1.

[11] While the GMA-cone clearly depends on the criterion function ϕ, we choose to ignore this dependence in the notation.

3.1 Construction and Properties of Monotonicity Paths

PROPOSITION 1 (PSEUDO-GRADIENTS) *If γ is a (t,v)-path and x a γ-differentiable selection of R, then*

$$D_v x(t) = \left.\frac{d}{d\lambda}\right|_{\lambda=0_+} x(\gamma(\lambda)) \in W(x(t),t,v).$$

Proof. Right-differentiating both sides of $f(x(\gamma(\lambda)),\gamma(\lambda)) = 0$ at $\lambda = 0_+$ yields

$$D_x f(x(t),t) \left.\frac{dx(\gamma(\lambda))}{d\lambda}\right|_{\lambda=0_+} + D_v f(x(t),t) = 0,$$

which satisfies (2) with $x(t) \in R(t)$. ■

The next result characterizes the shape of the set of all directions v in which the monotonicity criterion increases.

PROPOSITION 2 (MONOTONICITY DIRECTIONS) *For any t, $V(t|R)$ is a convex cone.*

Proof. Let $t \in \mathcal{T}$. We first show that $V(t|R)$ is a cone. It clearly contains the origin. For any t, $v \in V(t|R)$, $\alpha > 0$, and $x \in R(t)$, we have $W(x,t,\alpha v) = \alpha W(x,t,v)$, since $D_{\alpha v} f = \alpha D_v f$. Therefore, $D\phi(x)w \geq 0$ for $w \in W(x,t,\alpha v)$ if and only if $D\phi(x)w \geq 0$ for $w \in W(x,t,v)$. To establish the convexity of $V(t|R)$, consider any $v, v' \in V(t|R)$ and $x \in R(t)$. Since $D_{v+v'} f(x,t) = D_v f(x,t) + D_{v'} f(x,t)$, it follows that $W(x,t,v+v') = W(x,t,v) + W(x,t,v')$, and hence that $D\phi(x)w \geq 0$ for all $w \in W(x,t,v+v')$. Since the last inequality holds for all $x \in R(t)$, it follows that $v + v' \in V(t|R)$. ■

We now show that $V(\cdot|R)$, when considered on its entire domain $\mathcal{T}$, generates *all* monotonicity paths along which the criterion is monotonic. Since GMA is concerned with smooth solutions, we assume (possibly by restricting the information structure) that for all t, each element of $R(t)$ can be reached by a smooth selection of R. Precisely, R is *smooth* if for any path γ, parameter t on γ, and vector $x \in R(t)$, there exists a γ-differentiable selection y of R such that $y(t) = x$. Smooth information structures rule out isolated points in the graph of $X(t)$, which are irrelevant for the GMA concept of monotonicity.

THEOREM 1 (NECESSITY) *Suppose that R is smooth and that $v \notin V(t|R)$. Then there exists a (t,v)-path γ and a γ-differentiable selection x such that some component of the criterion $\hat{x} = \phi(x)$ is decreasing in a right-neighborhood of $\lambda = 0$.*

Proof. By definition, $v \notin V(t|R)$ implies that there exist $x \in R(t)$, $w \in W(x,t,v)$, and $i \in \{1,\ldots,d\}$ such that the i-th component of $D\phi(x)w$ is negative. Let γ be any path such that $\gamma(0) = t$ and $\dot{\gamma}(0_+) = v$. By smoothness of R, there exists a γ-differentiable selection x of R such that $x(0) = x$. By construction, the i-th component of $D\phi(x(0))w$ is negative for $w \in W(x,t,\dot{\gamma}(0_+))$, implying that $\phi_i(x(\gamma(\lambda)))$ is decreasing in a right-neighborhood of $\lambda = 0$. ∎

Thus, the set $V(t|R)$ is in effect the *largest* cone of directions v (i.e., the "GMA-cone") at t, for which the criterion $\hat{x}(t)$ is guaranteed to be monotonic at the beginning of any (t,v)-path.

The mapping $V(\cdot|R)$ can be viewed as a set-valued vector field: trajectories generated by $V(\cdot|R)$ are parameter paths with the property that at any t on such a path, the direction of the path belongs to $V(t|R)$. As the notation indicates, the shape of the cone $V(t|R)$ depends on the point t in the parameter space and on the available information structure R. Standard methods in comparative statics consider only the case where the GMA-cone is fixed to the positive orthant, i.e., where $V(t|R) = \mathbb{R}^m_+$. Trajectories generated by such a cone require that all components of the parameter path be nondecreasing. Here, by contrast, the GMA-cone $V(t|R)$ is determined endogenously by the primitives and the available additional information about solutions to the problem (1). As the information structure becomes finer, the GMA-cone cannot become smaller.

Theorem 2 (Information-Structure Refinements) *If $R \subseteq R'$, then $V(t|R') \subseteq V(t|R)$ for all $t \in \mathcal{T}$.*

Proof. For a given $t \in \mathcal{T}$ fix an arbitrary $v \in V(t|R')$. Then for any $x \in R(t) \subset R'(t)$ and any $w \in W(x,t,v)$ we have $D\phi(x)w \geq 0$, which implies by (4) that $v \in V(t|R)$. ∎

The finer the available information structure about the solution set, the larger the set of monotonicity directions in the parameter space. This establishes a partial order over information structures R, which is similar to Blackwell's (1951) order for the comparison of information sources.[12]

[12] An information source A is at least as informative as an information source B in the sense of Blackwell, if for *any* decision, subject to some minor technical requirements, the (outcome-contingent) payoffs attainable using information source B can also be attained using information source A. Analogously, an information set R is at least as useful for generating MCS as the information set R' if $R \subseteq R'$, since then $W(t|R') \subseteq W(t|R)$, i.e., the set of all monotonicity directions generated by R contains the one generated by R'. If R' is obtained from R by the removal of constraints, then the partial order is valid

The next result is the backbone of the GMA method: it states that the criterion $\hat{x}$ is nondecreasing along any direction of the cone V.

Theorem 3 (Sufficiency) *If γ is a path such that*

$$\dot{\gamma}(\lambda) \in V(\gamma(\lambda)|R) \tag{5}$$

for all $\lambda \in [0,1]$, then it is a monotonicity path.

Proof. Fix any γ-differentiable selection x of R. For any $\lambda \in [0,1)$, let $t = \gamma(\lambda)$ and $v = \dot{\gamma}(\lambda)$ (or the right-derivative of γ if $\lambda = 0$). By Proposition 1, $D_v x(t) \in W(x(t), t, v)$. By construction of γ, this implies that $D\phi(x(t))D_v x(t) \geq 0$. Now, for $\lambda < \lambda'$, the fundamental theorem of calculus implies that

$$\hat{x}(\gamma(\lambda')) - \hat{x}(\gamma(\lambda)) = \phi(x(\gamma(\lambda'))) - \phi(x(\gamma(\lambda))) = \int_{\lambda}^{\lambda'} D\phi(x(\gamma(\ell)))D_{\dot{\gamma}(\ell)}x(\gamma(\ell))d\ell \geq 0,$$

hence that the criterion $\hat{x}(\gamma(\lambda))$ is nondecreasing in λ. ■

The differential inclusion (5) describes the dynamics of monotonicity paths (interpreting λ as time).[13] When $D_x f$ is invertible and $\mathcal{T}$ is finite-dimensional, the pseudo-gradient is unique and can be directly computed, which simplifies the statement of Theorem 3.

Corollary 1 *Suppose that $v : \mathcal{T} \to \mathbb{R}^m$ is a smooth vector field such that*

$$-D\phi(x)\left[\begin{array}{cc} (D_x f)^{-1}(x,t) & D_t f(x,t) \end{array}\right] v(t) \geq 0 \tag{6}$$

for all $t \in \mathcal{T}$ and $x \in R(t)$. Then, trajectories generated by v are monotonicity paths with respect to ϕ.

Proof. With $D_x f$ invertible, the only vector w satisfying (2) is $-[D_x f]^{-1} D_v f$, so that the condition (4) for all $x \in R(t)$ and $w \in W(x,t,v)$ reduces to (6). ■

3.2 Method Summary

The GMA approach can be summarized as follows.

for *any* problem of the form (1) in which $\mathcal{X}$ and $\mathcal{T}$ are fixed.

[13] Differential inclusions are a standard tool in the description of dynamic systems (see, e.g., Smirnov (2002)).

1. *Step 1 (Pseudo-Gradients)* For all relevant t and v, use (3) to compute the set of pseudo-gradients. This set describes the possible changes of any smooth selection of R along any (t, v)-path, or – in other words – the solution sensitivity to any (smooth) change of parameters at t in direction v.

2. *Step 2 (GMA-Cone)* For each t, solve (4) to determine the convex cone $V(t|R)$ of monotonicity directions at t.

   Alternatively, one may wish to impose a cone of monotonicity directions and derive conditions on the primitives under which parameter changes of interest are monotonicity paths. For example, this is what standard comparative statics does, when V is identically equal to the positive orthant. This approach is further explored in Section 3.3.

3. *Step 3 (Monotonicity Paths)* When the GMA-cone is not reduced to zero, selecting one monotonicity direction for each t generates a vector field. It is typically possible to generate a smooth vector field (at least for a large subset of $\mathcal{T}$). Each of its trajectories is then a monotonicity path. By Theorem 3 this can be accomplished by solving the differential inclusion (5). If, however, $V(t|R) = \{0\}$, one needs to gather more information about the solutions to (1) to narrow down the set of pseudo-gradients which by Proposition 2 can only widen the GMA-cone. Alternatively, it may be useful to choose a less demanding criterion function ϕ, for example by reducing its dimension d (which may be achieved by dropping or aggregating variables).

4. *Step 4 (Optional - Reparametrization)* If monotonicity paths cover $\mathcal{T}$, they may naturally be used to provide a new parametrization of the initial problem, under which monotonicity obtains with respect to the first of these new parameters. More precisely, if one can find a (Lipschitz-continuous) function $\sigma : \mathcal{T} \times [0, 1] \to \mathbb{R}^m$, such that

   $$\gamma(0) = t,\ \dot{\gamma}(\lambda) = \sigma(\gamma(\lambda), \lambda)\ \forall\, \lambda \in (0, 1) \quad \Rightarrow \quad \gamma_t(\lambda) = \gamma(\lambda)\ \forall\, \lambda \in (0, 1)$$

   for all $t \in \mathcal{T}$, then, by the rectifiability theorem for direction fields (see, e.g., Arnold and Il'yashenko (1988)), it is possible to (at least locally) transform coordinates such that monotonicity is guaranteed. The function σ can be obtained by pasting together the differential equations that describe γ_t by selecting a continuous σ such that

   $$\sigma(t, \lambda) = \dot{\gamma}_t(\lambda) \in V(\gamma_t(\lambda)|R)$$

for all $t \in \mathcal{S}$. One method is to use a $(k-1)$-dimensional plane $\mathcal{P}$ that is transverse (i.e., never collinear) to the (Lipschitz-continuous) vector field described by $v(t) = \sigma(\gamma_t(0), 0)$ for all $t \in \mathcal{S}$. Then, any $t \in \mathcal{S}$ corresponds to a unique point on the plane (described by the $m-1$ contra-variant coordinates $s_2, \ldots, s_m$) and the time it takes to get to or from the plane $\mathcal{P}$ to t (described by the co-variant coordinate s_1), depending on the direction of the flow. This defines a new parametrization $t = \varphi(s)$ in $s = (s_1, \ldots, s_m)$ on $\mathcal{S}$ such that the criterion $\hat{x}(\varphi(s_1, s_2, \ldots, s_n))$ is nondecreasing in s_1.[14]

3.3 Imposed Monotonicity Paths

The method described thus far discloses all paths in the parameter space along which the criterion $\hat{x} = \phi(x)$ is nondecreasing for all smooth solutions x. In some problems, one may wish to reverse the question, and find conditions under which some particular directions of the parameter space are guaranteed to yield monotonicity of $\phi(x)$. For example, usual comparative statics consider the case in which V contains the positive orthant of the parameter space. Such conditions can be tested by checking that extreme rays of the corresponding GMA-cone, which are unit vectors of the parameter space, belong to V. Imposing that extreme rays of the positive orthant belong to V generates inequalities that are equivalent to the differential characterization of supermodularity as applied to objective functions of optimization problems.

More generally, suppose that one wishes to ensure that a particular cone V_0 of parameter directions guarantees the monotonicity of smooth solutions (or some criterion thereof), and let Δ denote a set of vertices (directions in $\mathcal{T}$) generating the cone V_0. Then the following result obtains, the proof of which is immediate and therefore omitted.

Proposition 3 *If $\Delta \subset V(t|R)$, then the criterion $\hat{x} = \phi(x)$ is nondecreasing along any trajectory generated by V_0, i.e., any path γ with $\dot{\gamma} \in V_0$ is a monotonicity path.*

Section 4.3 illustrates this result in the context of parameter aggregation.

[14] Further details on global reparametrizations, using additional tools in the theory of ordinary differential equations, have been developed elsewhere (Strulovici and Weber 2008). As the applications in Section 4 demonstrate, it is often enough to restrict attention to the first three steps of GMA to obtain a fairly complete picture of the monotonicity properties of the solutions to a given problem of the form (1).

3.4 Folding Information into the Primitive Equation

The analysis so far has maintained a clear distinction between the primitive equation (1) and the additional conditions imposed by the information structure R. In fact, the additional information encapsulated in R can often be folded into the primitive equation. To see this, consider an example. Suppose that

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

where the smooth constraint function $h(x,t)$ takes values in $\mathbb{R}^k$ (with $k \geq 1$). This situation arises when (1) represents the first-order necessary optimality condition of a maximization problem. The corresponding necessary second-order optimality condition (at any interior solution x, not on the boundary of $\mathcal{X}$) is that $f_x(x,t) \leq 0$. Define $\mathcal{Z} = \mathcal{X} \times \mathbb{R}^k$, $z = (x,y)$ with $y \in \mathbb{R}^k$. The information from $R(t)$ can be folded into the initial equation if f is extended to

$$\hat{f}(z,t) = \begin{bmatrix} f(x,t) \\ g(x,t) \\ h(x,t) + y^2 \end{bmatrix},$$

where $y^2 = ((y_1)^2, \ldots, (y_k)^2) \in \mathbb{R}^k_+$ represents a vector whose elements are the nonnegative slacks associated with the components of the inequality constraint.

To further illustrate this technique, suppose that x is one-dimensional, and that $R(t) = \{x : f(x,t) = 0 \text{ and } f_x(x,t) \leq 0\} \subset \mathbb{R}$. Consider any smooth selection $x : t \mapsto x(t)$ of R. If $f_x(x(t),t) < 0$, then $x'(t) = -f_t(x(t),t)/f_x(x(t),t)$ by the implicit function theorem. However, this theorem cannot be applied if $f_x(x(t),t) = 0$. Suppose now that the second-order condition is incorporated into f. The resulting equation is

$$\hat{f}(x,y,t) = \begin{bmatrix} f(x,t) \\ f_x(x,t) + y^2 \end{bmatrix} = 0.$$

Differentiating these equations yields[15] relations: $f_x x' + f_t = 0$ and $f_{xx} x' + f_{xt} + 2yy' = 0$. For $y(t) \neq 0$, f_x is invertible and the implicit function theorem can be used. If $y(t_0) = 0$ for some parameter t_0, the second relation yields $x'(t_0) = -f_{xt}(x(t_0),t_0)/f_{xx}(x(t_0),t_0)$, which describes the sensitivity of the selection $x(t)$ at t_0, despite the singularity of the equation at that particular parameter and solution (as long as $f_{xx}(x(t_0),t_0) \neq 0$). Thus, folding

[15] Smoothness of $y(t)$ obtains by the relation $f_x(x(t),t) + y^2(t) = 0$, as well as the smoothness of f_x and x in t.

the second-order condition into f implies sharp predictions for the pseudo-gradient, even at singularity points.[16]

3.5 Quantitative Monotonicity Analysis

Another innovation of this paper is to examine at some degree of generality the rate at which solutions are monotonic. We believe that this question, for which ordinal methods are ill-suited, may arise in many important instances, as illustrated by the application in Section 4.4. In contrast, this question is a natural extension of our method. *Quantitative monotonicity analysis* investigates the magnitude of the change in the solution to the primitive equation (1) as a consequence of parameter variations. We start by extending our previous result to the strict monotonicity of solutions to (1), considering directions in

$$\hat{V}(t|R) = \{v \in \mathbb{R}^m : D\phi(x)w > 0,\ x \in R(t),\ w \in W(x,t,v)\},$$

which for all $t \in \mathcal{T}$ is a subset of the GMA-cone $V(t|R)$.

Proposition 4 *If γ is a differentiable path such that $\dot{\gamma}(\lambda) \in \hat{V}(\gamma(\lambda)|R)$ for all $\lambda \in [0,1]$, then $\phi(x)$ is (strictly) increasing along γ.*

Proof. Consider any γ-differentiable selection x of R and $\lambda \in [0,1)$, and let $t = \gamma(\lambda)$ and $v = \dot{\gamma}(\lambda) \in \hat{V}(t|R)$. From Proposition 1, $D_v x(t) \in W(x(t), t, v)$. By assumption, $D\phi(x(t))D_v x(t) > 0$. For $0 \leq \lambda < \lambda' \leq 1$, it is

$$\phi(x(\lambda')) - \phi(x(\lambda)) = \int_{\lambda}^{\lambda'} D\phi(x(\gamma(\ell)))D_{\dot{\gamma}(\ell)}x(\gamma(\ell))d\ell.$$

The last integral is positive, which shows the result. ∎

If one can find a positive lower bound for the sensitivity of the criterion, the result can be strengthened as follows.

Proposition 5 (Quantitative Monotonicity Analysis) *Let γ be a differentiable path. If there exists a vector $\rho = (\rho_1, \ldots, \rho_d)$ with strictly positive components such that*

$$w \in \bar{W}(\gamma(\lambda), \dot{\gamma}(\lambda)|R),\ x \in R(\gamma(\lambda)) \quad \Rightarrow \quad D\phi(x(\gamma(\lambda)))\, w \geq \rho$$

[16] As long as $R(t)$ describes a closed subset of the solution $X(t)$, the above method works. Indeed, a theorem by Whitney (Postnikov 1987, p. 20) states that *any* closed set $\mathcal{Z}$ at t can be written in the form $\hat{f}(z,t) = 0$ with an appropriate function $\hat{f}$ that is smooth in z. The smoothness of $\hat{f}$ in (z,t) can also be guaranteed by Whitney's theorem, as long as the graph of the set-valued mapping, which assigns the relevant constraint set $\mathcal{Z}(t)$ to each parameter value t, is closed.

for all $\lambda \in (0, 1)$*, then*

$$0 \leq \lambda < \lambda' \leq 1 \quad \Rightarrow \quad \hat{x}(\gamma(\lambda')) - \hat{x}(\gamma(\lambda)) \geq \rho(\lambda' - \lambda).$$

The proof is essentially identical to the proof of Proposition 4 and is therefore omitted.[17] By a similar argument, one can derive upper bounds on the rate of monotonicity. Proposition 5 can be used, for example, to show that x increases faster (or slower) than a parameter (taking $\rho = 1$). The value of ρ provides valuable information about the rate of increase (or decrease), giving rise to quantitative statements. A simpler result than Proposition 5 has been used when $\phi(x) = x_i$ for some $i \in \{1, \ldots, n\}$, t is finite-dimensional, and the path γ amounts to increasing a single component t_j of the parameter vector (Samuelson 1947). Indeed, suppose that $D_x f$ is invertible that $i = j = 1$. Then Corollary 1 with $v = (1, 0 \ldots, 0)^T$ and $D\phi = (1, 0, \ldots 0)$ yields the condition $m(x,t)^T a(x,t) \geq \rho$, where $m(x,t)$ is the first row of the matrix $-[D_x f]^{-1}(x,t)$ and $a(x,t)$ is the vector $(df_1/dt_1, df_2/dt_1, \ldots, df_n/dt_1)$. An application of quantitative monotonicity analysis to robust inference is provided in Section 4.4.

4 Applications

In Section 4.1 we show how constraints can be incorporated into the framework of Section 3 and give two illustrations of GMA with constraints. The first recovers Chipman's (1977) normal-good theorem for supermodular, strongly concave utility functions (Section 4.1). The second derives a condition on the marginal rate of substitution of a utility function in order for a good to be a Giffen good. Section 4.2 is a key application of the paper. It considers a non-supermodular game in which each player has two decisions (investment and location) and a one-dimensional type. Using GMA, we obtain conditions for investment decisions to be nondecreasing in parameter directions that increase a certain 'proximity' between players. This application also illustrates the use of criterion functions (to select a subset of the variables) and of the GMA-cone (to describe parameter directions in which the players' proximity increases). The notion of proximity is derived endogenously. The subsequent application, presented in Section 4.3, illustrates another aspect of GMA, demonstrating the consequences of imposing a large cone of monotonicity directions. Specifically, it is shown that requiring monotonicity with respect to any additive increasing function of the parameters implies that the solution depends only

[17] Proposition 5 can easily be extended to the case in which ρ is a function of the parameter.

on that function. The result is proved for the case where the aggregate is a parameter sum. A simple transformation of each parameter component generalizes this proof to any additive, increasing function of the parameters. This result, which goes beyond pure comparative-statics analysis, provides a good illustration of the use of GMA-cones. In Section 4.4, quantitative monotonicity analysis is used as a tool for robust inference about unobservable parameters in the context of productivity estimation. The last application, discussed in Section 4.5, concerns comparative dynamics. More than all of the preceding applications, it shows the importance of pseudo-gradients for monotonicity analysis. Proof is provided for a general theorem for comparative dynamics, which implies a simple result about the implications of local properties of the law of motion on the monotonicity of solutions (related to Huggett (2003)). It is then shown how *global* properties of the dynamic equation also can be exploited by the theorem, in a result that can be interpreted as a stylized model of the influence of positive global cycles on the growth of individual firms.

4.1 Constrained Optimization

Constrained optimization problems often prevent the use of ordinal methods, by violating crucial lattice assumptions.[18] GMA can accommodate problem constraints in a natural manner, as long as they are described in terms of smooth functions.[19] The general argument is described first, then applied both to prove Chipman's normal-good theorem for supermodular, strongly concave utility functions, and to derive a condition under which a good is a Giffen good. Consider the problem

$$\begin{aligned} &\max\nolimits_{y\in\mathbb{R}^n} G(y,t), \\ &\text{s.t.}\;\; g(y,t)=0, \end{aligned}$$

where $G : \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}^l$ and $g : \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}^k$ are twice differentiable. A necessary condition for optimality (Bertsekas 1995, p. 255) is the existence of a k-dimensional vector ν such that at the optimum

$$F(y,\nu,t) = D_yG(y,t) + \nu \cdot D_yg(y,t) = 0.$$

[18] A notable exception is Quah's (2007) ordinal method, which provides conditions on the transformation of constraints to guarantee monotonicity for the entire vector of decision variables.

[19] As pointed out in Footnote 16 at the end of Section 3.4, provided that the problem constraints confine the variable to a closed set, the assumption of a constraint representation in terms of smooth functions is not restrictive, at least from a theoretical point of view.

Together with the k equations $g(y,t)=0$, this determines a system of $n=l+k$ equations in n variables $x=(y,\nu)$,

$$f(x,t)=\begin{bmatrix} D_yG(y,t)+\nu\cdot D_yg(y,t) \\ g(y,t) \end{bmatrix}=0,$$

corresponding to the primitive equation (1) of our analysis. The approach can easily be extended to allow for both equality and inequality constraints (cf. Section 3.4).

Example: Normal Goods. As an illustration, consider the budget-constrained optimal-consumption problem

$$U(t)=\max_{y:\, g(y,t)=0} G(y),$$

where $g(y,t)=p\cdot y-t$, $p\gg 0$ denotes the price vector for a commodity bundle y, and $t>0$ is the available budget. Our goal is to show that if G is increasing, strongly concave, and supermodular, then any smooth optimizer selection $y(t)$ is nondecreasing in t. The constraint $g(y,t)=0$ does not define a lattice if there are three or more goods.[20] With G concave and differentiable, the first-order condition and budget constraint can be written in the form

$$f(x,t)=\begin{bmatrix} D_yG(y)-\lambda p \\ g(y,t) \end{bmatrix}=0,$$

where $x=(y,\lambda)$ and λ is the nonnegative Lagrange multiplier associated with the budget constraint.

In this problem, the natural criterion function is $\phi(x)=\phi(y,\lambda)=y$, since only monotonicity in demand matters. The information set $R(t)$ consists of two constraints. The first constraint is the equation itself, $f(x,t)=0$. The second constraint stems from the observation that the value function U is concave in t as[21] the maximum of a concave function subject to the convex constraint $p\cdot y\le t$. Since $\lambda=U'(t)$, concavity of U implies that $\lambda'(t)\le 0$. The constraints $f(x,t)=0$ and $\lambda'(t)\le 0$ define the information structure of the problem. Finally, normal-good monotonicity means that any wealth increase should increase demand in all goods. This implies that the GMA-cone must be $V=[0,\infty)$ (since t lies in $\mathbb{R}$, a nontrivial GMA-cone is necessarily one of the rays $[0,\infty)$ or $(-\infty,0]$).

[20] With two goods, adopting the new order (reversing the sign of the quantity of the second good) is enough to recover the lattice structure and show the result.

[21] See for example Luenberger (1969, p. 216).

For any x and t, the pseudo-gradients $w(x,t)$ must solve, by definition,

$$\begin{bmatrix} D_y f w_y(x,t) - p w_\lambda(x,t) \\ p \cdot w_y(x,t) - 1 \end{bmatrix} = 0,$$

where w_y and w_λ are the components of w corresponding to y and λ. Strong concavity of G implies that $D_y f$ is invertible. Since G is concave and supermodular, the diagonal elements of $D_y f$ are nonpositive, while its off-diagonal elements are nonnegative. This implies (see, e.g., Samuelson 1947) that all elements of $[D_y f]^{-1}$ are nonpositive. Therefore, $w_y(x,t) = w_\lambda(x,t)[D_y f]^{-1} p$ has nonnegative components. This proves the following result.

Proposition 6 (Normal Goods) *If G is strongly concave, supermodular, and twice differentiable, then $y_i(t)$ is nondecreasing in t for all $i \in \{1, \ldots, n\}$.*

The result is identical to Chipman (1977). The proof there is very similar to ours, but does not proceed from a general approach to comparative statics: it uses properties of the indirect utility function instead of treating Lagrange multipliers as part of an enlarged system of equations. This leads to one subtle difference: while we exploit the fact that the Lagrange multiplier has a nonpositive derivative to prove good normality, Chipman first shows that demand either increases for all goods or decreases for all goods, and concludes by observing that, the budget having expanded, all demand has necessarily increased for some good.[22]

Example: Giffen Goods. Consider the budget-constrained optimal-consumption problem

$$\max_{(x,y)\in B(p,q)} u(x,y) \tag{7}$$

where u is increasing and concave, p, q are positive prices, and consumption (x,y) is constrained by the budget set

$$B(p,q) = \{(x,y) \geq 0 : px + qy \leq 1\}.$$

We use GMA to derive conditions under which the second good, say, behaves as a Giffen good, i.e., such that the demand $y(p,q)$ for the second good increases in q. We focus on interior solutions. From our earlier analysis, the problem can be rewritten in the form

$$\begin{bmatrix} u_x(x,y) - \lambda p \\ u_y(x,y) - \lambda q \\ px + qy - 1 \end{bmatrix} = 0, \tag{8}$$

[22]See also Quah (2007) for a treatment of comparative statics of this and other constrained optimization problems without differentiability assumptions.

where λ is the Lagrange multiplier associated with the budget constraint.

Applying Step 1 of the GMA-method (cf. Section 3.2) yields a description of the set of pseudo-gradients as a solution to

$$\begin{bmatrix} u_{xx}(x,y) & u_{xy}(x,y) & -p \\ u_{xy}(x,y) & u_{yy}(x,y) & -q \\ p & q & 0 \end{bmatrix} \begin{bmatrix} w_x \\ w_y \\ w_\lambda \end{bmatrix} + \begin{bmatrix} -\lambda & 0 \\ 0 & -\lambda \\ x & y \end{bmatrix} \begin{bmatrix} v_p \\ v_q \end{bmatrix} = 0 \tag{9}$$

in terms of $w = (w_x, w_y, w_\lambda)^T$ for any given parameter vector (p, q), candidate solution (x, y), and (nonzero) parameter direction (v_p, v_q). Using a symbolic solver (and the equalities $u_x/p = u_y/q = \lambda$), we find that the unique solution for w_y is

$$w_y = \frac{1}{q} \frac{\left(\frac{u_x}{u_y}\right)_x (xv_p + yv_q) + \frac{u_x}{u_y}\left(v_p - \frac{u_x}{u_y} v_q\right)}{\frac{u_x}{u_y}\left(\frac{u_x}{u_y}\right)_y - \left(\frac{u_x}{u_y}\right)_x} = \frac{1}{q} \frac{(x\mu_x + \mu)v_p + (y\mu_x - \mu^2)v_q}{\mu\mu_y - \mu_x}, \tag{10}$$

where $\mu(x, y) = u_x(x, y)/u_y(x, y) > 0$ is the consumer's marginal rate of substitution between the two goods. To characterize situations where the second good is a Giffen good, we use the alternative Step 2 (cf. Section 3.3) of the method by imposing the parameter change $v_p = 0$ and $v_q = 1$, and asking that w_y be nonnegative. Concavity[23] of u can be used to show that the denominator in the rightmost expression of (10) is positive. This yields the following result.

Proposition 7 (Giffen Goods) *$y(p, q)$ is (locally) increasing in q if and only if $\mu^2 < y\mu_x$ evaluated at $(x(p, q), y(p, q))$.*

This result illustrates how constraints can be dealt with to obtain conditions on the primitive of the problem (here, the marginal rate of substitution between goods) that yield monotonicity. Intuitively the condition states that the second good is Giffen provided that the marginal rate of substitution of the first good with respect to the second is small compared to the elasticity of that rate of substitution with respect to x: if the price of the second good increases, an increase in x would reduce the marginal value of the first good compared to the second by so much that it is better to increase consumption of the second good instead. Sørenson (2007) provides simple examples for Giffen goods, which rely on a Leontief-type kink in the demand curve, whereas the condition in Proposition 7 works with smooth primitives.

[23] More precisely, we assume that the Hessian of u is negative definite.

4.2 Equilibrium of a Non-Supermodular Game

Supermodular games have been widely studied with comparative-statics techniques, yielding interesting results about the existence and monotonicity of equilibria (Vives 1990; Milgrom and Roberts 1990; Echenique 2002). Far less is known about monotonicity properties of non-supermodular games. Such games provide a good illustration of how GMA can be used to explore a complex comparative-statics problem and get some insights about its structure. We consider a game with two players in which each player chooses two actions: a nonnegative quantity q_i and a location z_i on the real line. Players have a type $t_i \in \mathbb{R}$ corresponding to some preference in the location space. When choosing q_i, player $i \in \{1, 2\}$ obtains the payoff

$$\Pi^i(q_i, z_i, z_j, t_i) = q_i R^i(z_i, z_j, t_i) - C^i(q_i),$$

where C^i is a strictly convex cost function, R^i is player i's revenue function, and j denotes the index of the other player. For example, if two firms engage in a joint venture, q_i may represent firm i's investment, z_i its final product positioning, t_i its initial preference or position, R^i is i's return and C^i its opportunity cost from investing q_i in the joint venture rather than in other projects. We assume that R^i is supermodular in (z_i, z_j) and (z_i, t_i) and strictly concave in z_i. Note, however, that Π^i is not in general supermodular in (q_i, z_i), so that standard results from the theory of supermodular games do not apply. In fact, it is easy to verify that q_i cannot be monotonic in t_i or t_j over the entire parameter space. Therefore, monotonicity results are possible only if one looks for other parameter moves.

Our goal is therefore to find parameter moves that jointly increase players' equilibrium investments (q_1^*, q_2^*). The above supermodularity assumptions ensure that an increase in player types increases their location decision. Still, such an increase does not necessarily increase player i's revenue R^i or his action q_i.

Equilibrium conditions, based on the first order conditions for the two actions of each of the two players, form a system of four equations:

$$f(x, t) = \begin{bmatrix} R^i - (C^i)'' \\ R^i_1 \\ R^j - (C^j)'' \\ R^j_1 \end{bmatrix} (x, t) = 0,$$

where $x = (q_1, z_1, q_2, z_2)$ and $t = (t_1, t_2)$, and R^i_k denotes R^i's partial derivative with respect to its k-th variable. Hessian and cross-partial derivative matrices of f are

$$D_x f = \begin{bmatrix} -(C^1)'' & 0 & 0 & R_2^1 \\ 0 & R_{11}^1 & 0 & R_{12}^1 \\ 0 & R_2^2 & -(C^2)'' & 0 \\ 0 & R_{12}^2 & 0 & R_{11}^2 \end{bmatrix}, \quad D_t f = \begin{bmatrix} R_3^1 & 0 \\ R_{13}^1 & 0 \\ 0 & R_3^2 \\ 0 & R_{13}^2 \end{bmatrix}.$$

Since we are interested in variations of q_1^* and q_2^*, we use the criterion

$$\phi(q_1, z_1, q_2, z_2) = (q_1, q_2).$$

From Corollary 1, monotonicity directions $v = (v_1, v_2)$ are given by the condition

$$-D\phi(x)[(D_x f)^{-1} D_t f](x,t)v(t) \geq 0.$$

After simplification, this yields $Av \geq 0$, where

$$A_{ii} = \kappa^i [R_3^i + \delta R_{13}^i R_2^i R_{12}^j],$$

and

$$A_{ij} = \kappa^i [\delta R_{13}^j R_2^i (-R_{11}^i)],$$

with $\kappa^i = -[(C^i)'']^{-1} > 0$ and $\delta = (R_{11}^1 R_{11}^2 - R_{12}^1 R_{12}^2)^{-1}$. We assume that $(-R_{11}^i) > R_{12}^i$ for $i \in \{1, 2\}$, ensuring that δ is positive. This condition means that the marginal return of a player's action is more sensitive to his own action than to the other player's. We also assume that if $R_1^i = R_1^j = 0$ and $t_i > t_j$, then R_2^i and R_3^j are nonnegative while R_2^j and R_3^i are nonpositive. This means that, in equilibrium (when $R_1^i = R_1^j = 0$), player returns decrease if one widens the gap between their preferences (conditions on R_3's), and player i's return increases if player j's action moves in the direction of i's preference (conditions on R_2's). Those assumptions will be satisfied, for example, if one can check independently that $t_i > t_j$ implies that $z_i \geq z_j$ in equilibrium (this is likely to hold if player returns are symmetric), and that on the domain $\{(z_i, z_j, t_i, t_j) : z_i \geq z_j \text{ and } t_i \geq t_j\}$, the above inequalities hold, as can easily be checked for any particular form of the return functions.

Under those assumptions, one would expect that, as player preferences get closer to each other, players will benefit more from their interaction and thus invest more. Indeed, the off-diagonal elements of the matrix A are positive for i and negative for j, which already implies that if $t_i > t_j$, an increase of t_i (higher type) causes player j to reduce investment, while an increase in t_j causes player i to increase investment. To obtain global comparative statics, however, a joint move of player types is required, which brings them closer in a particular direction. Thus, suppose, without loss of generality, that $t_1 > t_2$, and that t_1 is decreased by R^2_{13}, while t_2 is increased by R^1_{13}. The effect on (q_1^*, q_2^*) is proportional to

$$R^2_{13}(-R^1_3) + \gamma((-R^1_{11}) - R^2_{12}), R^1_{13}R^2_3 + \gamma((-R^2_{11}) - R^1_{12}),$$

where $\gamma = R^1_{13}R^2_{13} > 0$. As a result, investments increase in equilibrium as long as a player's action has more impact on his marginal return than on the other player's, that is if $(-R^i_{11}) > R^j_{12}$. The slope R^1_{13}/R^2_{13} can be interpreted as the ratio of players' action-type complementarity. If player 1 exhibits more action-type complementarity than player 2, increasing investments will be guaranteed only if 2's preference is increased by a larger amount than that by which 1's preference is decreased. That ratio is constant if players incur a quadratic cost by moving from t_i to z_i: $R^i = -\mu_i(z_i - t_i)^2 + \psi^{i,1}(z_i, z_j) + \psi^{i,2}(z_j, t_i)$ for $i = 1, 2$. With symmetric complementarity ($\mu_1 = \mu_2$), investments increase as player types get closer to each other, i.e., as $|t_1 - t_2|$ decreases.[24]

4.3 Aggregation

GMA can be used to examine the comparative statics of a given criterion in terms of aggregates of parameters, such as an arithmetic mean or, equivalently, a sum. As an example, we use the GMA-method to establish, under fairly general conditions, that a solution to (1) is monotonic with respect to the sum $t_1 + \cdots + t_m$ of the parameters only if it is *independent* of the individual parameters conditional on that sum.[25] The proof of this result illustrates the use of imposed monotonicity paths (cf. Section 3.3). Given a criterion ϕ and an information structure R, by Corollary 1, the condition

$$(x, v) \in R(t) \times V(t|R) \quad \Rightarrow \quad -D\phi(x)[(D_x f)^{-1} D_t f](x, t)v \geq 0,$$

implies that trajectories generated by $V(t|R)$ are monotonicity paths. By Theorem 1, the last condition is also necessary to guarantee monotonicity.

Let $t = (t_1, \ldots, t_m)$ be a parameter vector, and consider the sum of its components, $s_m = t_1 + \cdots + t_m$. Solution monotonicity in s_m is stronger than solution monotonicity in any given component of t: the former requires not only that the solution is nondecreasing

[24] For this symmetric quadratic case, there is an n-player equivalent to the result: types get "closer" if the vector $t = (t_1, \ldots, t_n)$ gets closer (for the usual Euclidean distance) to the first bisector of the parameter space, i.e., the line whose direction is given by the vector $(1, \ldots, 1)$.

[25] By a simple nonlinear re-scaling of the parameter vector, this result can be extended to any additive function that is strictly increasing in the parameters.

in each component of t, but also that it (weakly) increases when a parameter t_i is increased by more than the amount by which another parameter t_j is decreased, keeping other parameters fixed. For simplicity, it is assumed that $X(t)$ is nonempty for all t, and that $f_x < 0$ everywhere. These assumptions are naturally satisfied when (1) is obtained as the first-order necessary optimality condition for the maximization of a strongly concave objective function on a convex set. The solution set $X(t) \equiv R(t)$ contains only a single element, denoted by $x(t)$.

First, consider the simple case when $m = 2$. The first three steps of the GMA-method in Section 3.2 yield that the GMA-cone with respect to monotonicity in s_2 is

$$V(t|R) = \{(\tau_1, \tau_2) \in \mathbb{R}^2 : \tau_1 + \tau_2 \geq 0\}.$$

This cone itself is generated by $\Delta = \{(-1, 1), (1, -1), (1, 1)\}$. Hence, f must satisfy

$$(x, v) \in R(t) \times \Delta \quad \Rightarrow \quad -f_x^{-1} \begin{bmatrix} f_{t_1}(x,t) & f_{t_2}(x,t) \end{bmatrix} v \geq 0. \tag{11}$$

By substituting the different $v \in \Delta$, the last relation can be rewritten in the form

$$x \in R(t) \quad \Rightarrow \quad \min \{f_{t_1} + f_{t_2}, f_{t_1} - f_{t_2}, -f_{t_1} + f_{t_2}\} (x, t) \geq 0,$$

which implies that $f_{t_1}(x, t) = f_{t_2}(x, t) \geq 0$ for all $x \in R(t)$. Let $s = (s_1, s_2) = (t_1, t_2) = t$ define a smooth change of parameters by $t = \varphi(s) = (s_1, s_2 - s_1)$, and let

$$F(x, s) = f(x, \varphi(s)).$$

Then s_1 is a contra-variant coordinate, as $F_{s_1}(x, s) = f_{t_1}(x, \varphi(s)) - f_{t_2}(x, \varphi(s))$ vanishes on $R(\varphi(s))$. Fix s and consider a differentiable selection $p \mapsto x(s, p)$. Since $F_x(x(\varphi(s)), s) = f(x(\varphi(s)), \varphi(s)) < 0$ by assumption, the solution in the new coordinates, $x(\varphi(s))$, must be independent of s_1. It can therefore depend only on the parameter aggregate s_2. Note that the information structure in this example is given by $R(t) \equiv \{x(t)\}$: the mere knowledge that $x(t)$ is a solution to (1), along with the stipulated requirement that x be nondecreasing in s_2, is enough to imply that that $x(\varphi(s))$ depends only on s_2. This conclusion generalizes to more than two parameter dimensions.

Theorem 4 (Monotonicity and Dependence on Aggregates) *Assume that the solution to (1) is a singleton, i.e., $X(t) \equiv \{x(t)\}$ on $\mathcal{T}$. If*

$$s'_m = t'_1 + ... + t'_m > t_1 + \cdots + t_m = s_m \quad \Rightarrow \quad x(t') \geq x(t)$$

for all $t' = (t'_1, \ldots, t'_m), t = (t_1, \ldots, t_m) \in \mathcal{T}$, then there exists a function $y : \mathbb{R} \to \mathcal{X}$ such that $x(t) = y(s_m)$ on $\mathcal{T}$, i.e., the solution depends <u>*only*</u> *on the aggregate s_m.*

Proof. The result has been already established for $m = 2$. For $m > 2$ consider the following induction step. Assume that the result is true for $k \geq 2$ parameters, and examine the case of $k + 1$ parameters, so that $t = (t_1, \ldots, t_{k+1})$. For a fixed t_{k+1}, the solution $x(t)$ is nondecreasing in s_k and thus depends only on s_k, by induction hypothesis. Thus, x depends only on s_k and t_{k+1}. Moreover, x is monotonic in the sum $s = s_k + t_{k+1}$. Applying the argument for the two-dimensional result to the parameter vector (s_k, t_{k+1}) then establishes that there is a function $y : \mathbb{R} \to \mathcal{X}$ such that $x(t) = y(s_k + t_{k+1}) = y(s_{k+1})$, which concludes our proof. ■

Theorem 4 implies that the only way a unique solution to a parameterized equation of the form (1) can be monotonic in a parameter average is if this average is a "sufficient statistic" for all the parameters involved in its computation. There have been numerous studies in aggregation and statistics concerning the question of which structures of problem (1) imply that solutions depend solely on a given parameter aggregate (see, e.g., the treatise by Blackorby and Shorrocks (1996)). The monotonicity analysis performed here considers the converse question and therefore yields necessary structural conditions: monotonicity in an additive aggregate implies that the aggregate must be a sufficient statistic for its components.[26]

In aggregation theory, a central question is to determine when some macroeconomic variable will be monotonic in a certain aggregate value of household characteristics. Combining Theorem 4 with previous aggregation results, the above analysis suggests that aggregate monotonicity can occur only if the macroeconomic variable is determined by an equation where the aggregate value is the only determinant parameter. Individual parameters may influence the function only through a multiplicative factor that does not affect the root of the equation.

[26]Another way to make the connection between monotonicity and dependence on additive aggregates is as follows. Suppose that $x(t)$ is monotonic in the sum $s_2 = t_1 + t_2$. Fix $t' = (t'_1, t'_2)$, and let $s'_2 = t'_1 + t'_2$, $a = \sup_{t: t_1 + t_2 < s'_2} \{x(t)\}$, and $b = \inf_{t: t_1 + t_2 > s'_2} \{x(t)\}$. By assumption, $a \leq x(t') \leq b$. Continuity of $x(t)$ in t, as a consequence of the maximum theorem (Berge 1963), then implies that $a = x(t') = b$. This is true for all $x(t)$ such that $t_1 + t_2 = s'_2$. Therefore, $x(t)$ must be constant on the line $t_1 + t_2 = s'_2$. This shows that x depends only on the parameters through their sum.

4.4 Robust Productivity Estimates

An important problem in applied industrial organization is the comparison of firms' private productivity forecasts given their observable investment decisions. Olley and Pakes (1996) consider a model in which firms base investment decisions on current capital stock (observed by outsiders) and productivity forecasts (unobserved by outsiders). They provide conditions under which investment decisions are monotonic in productivity forecasts. Clearly, if a firm's productivity forecast were the only parameter influencing its investment decision, one could simply rank the firms' productivity forecasts based the observed investment decisions. However, since capital stock materially affects investment decisions, such a comparison of productivity forecasts across firms is valid only for a fixed level of capital stock.

Quantitative monotonicity analysis can be used to show that it is possible to rank two firms' productivity estimates, as long as their capital stocks are close, and their investment decisions differ sufficiently. In that sense the monotonicity disclosed by GMA is *robust* with respect to perturbations, and thus allows for *robust inference*. To simplify the exposition, we first consider a static setting where investment decisions are based on a simple point estimate of a firm's productivity, before discussing the dynamic setting in which investment decisions are always based on the current forecast for the next realization of the underlying Markovian productivity process.

4.4.1 Static Setting

Suppose that the first-order condition of a firm's investment problem is

$$f(x, k, p) = 0,$$

where x is the investment decision, k is the firm's capital stock, and p is its productivity estimate. Assuming invertibility of f_x, the pseudo-gradient is (with $t = (k, p)$)

$$w(x, t) = -f_x^{-1} \left[\begin{array}{cc} f_k & f_p \end{array} \right] (x, t).$$

Since we are interested only in the monotonicity of any solution $x(k, p)$, the criterion function here is simply $\phi(x) \equiv x$. Since $f_x < 0$ at any solution (second-order condition, assuming strict concavity of the objective in the underlying maximization problem), the firm's investment decision increases if $f_k dk + f_p dp$ is positive. Suppose one has observed two firms' investment decisions, $x_1 = x(k_1, p_2)$ and $x_2 = x(k_2, p_2)$, and it is known that their capital stocks are close, i.e., such that $|k_2 - k_1| \le \kappa$ for some $\kappa > 0$. Then, if $f_p > \rho$ for some $\rho > 0$ and $|f_k| < \eta$ for some $\eta < 0$ on the entire domain of f, then $x_2 - x_1 > \theta > 0$ implies $p_2 > p_1 + (\theta - \kappa\eta)/\rho$. In particular, an observed investment difference between firm 2 and firm 1 of more than $\kappa\eta$ implies that firm 2's productivity forecast is *surely* larger than firm 1's productivity forecast, no matter what the firms' precise capital stocks are.

4.4.2 Dynamic Setting

In the dynamic setting considered by Olley and Pakes, productivity follows a stochastic Markov process. The only parameters affecting a firm's investment decision are capital stock, productivity level, and the productivity forecast.[27]

Letting $\tilde{p}$ (or $\tilde{p}|p$) denote the (random) next-period productivity level given that the current productivity level is p, a firm's value function $\Pi(k,p)$ solves the Bellman equation

$$\Pi(k,p) = \max_x \{\pi(k,p) - c(x,k) + \beta E[\Pi(x + (1-\delta)k, \tilde{p})|p]\},$$

where π is the current-period gross payoff, c is the adjustment-cost function, β is the per-period discount factor, and δ represents the per-period depreciation rate. Convexity of c in x and concavity of π imply that the optimal dynamic investment policy $x(k,p)$ satisfies the first-order necessary optimality condition

$$f(x,k,p) = \beta E[\Pi_k(x + (1-\delta)k, \tilde{p})|p] - c_x(x,k) = 0.$$

Applying the above analysis yields the following result.

Proposition 8 *If there exists $\eta > 0$ such that*

$$|\beta(1-\delta)E[\Pi_{kk}(x + (1-\delta)k, \tilde{p})|p] - c_{xk}(x,k)| < \eta$$

for all (x,k,p), then

$$x(k_2,p_2) > x(k_1,p_1) > \eta|k_2 - k_1| \quad \Rightarrow \quad p_2 > p_1,$$

where $x(k_i,p_i)$, for $i \in \{1,2\}$, denotes the optimal investment given a current capital stock of k_i and a current productivity estimate of p_i.

[27] In accordance with recent literature (cf. Olley and Pakes (1996)), we abstract from the firm's age in this example. For simplicity, we also rule out exit, which means that salvage values are sufficiently low.

Therefore, any upper bounds on the concavity of the firm's value function and on the cross-derivative of the cost function with respect to capital and productivity allow for a robust inter-firm comparison of private productivity estimates based on observed investments, as long as capital stocks in the industry are not too different.

4.5 Comparative Dynamics of Equilibrium Paths

Given a parameterized description of a time-varying process in a Euclidean space, *monotone comparative dynamics* obtain when a parameter increase results in the process taking on higher values at all points after the initial time. Such an approach to comparative dynamics, with applications to growth, has been studied by Brock and Mirman (1972), Mendelsohn and Sobel (1980), Becker (1983), and Amir et al. (1991).

In what follows, we take the view that comparative dynamics is equivalent to a comparative-statics problem in which the variable is the entire process, and the equation $f = 0$ describes its law of motion.

This section exploits the concept of pseudo-gradients to derive comparative dynamics. The variable x is a function of time, the evolution of which is described by some dynamic equation. The goal is to determine nontrivial changes of the dynamic equation that increase the value of x at each instant. Cast in the GMA framework, comparative dynamics imply that pseudo-gradients must satisfy a differential equation. This equation is first used to prove a simple result relating local properties of the dynamic equation to the monotonicity of x, and then to show how *global* properties of the dynamic equation can be incorporated into monotonicity analysis, providing a highly stylized interpretation of the impact of global cycles on comparative dynamics (Proposition 10).

For this application, $\mathcal{X}$ is infinite-dimensional. In keeping with standard notation in dynamic systems, the parameter is denoted by α instead of t, and t now denotes time. Suppose that the evolution of a process can be described in terms of the initial value problem

$$\dot{x}(t) = g(x, t, \alpha), \quad x(0) = x_0, \tag{12}$$

for all $t \in [0, T]$, where $x_0 \in \mathbb{R}^N$, $T \in \mathbb{R}_{++}$, and $\alpha \in \mathcal{A} = [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R}^m$ (with $\underline{\alpha} < \bar{\alpha}$) are given constants, and $g : \mathbb{R}^N \times [0, T] \times \mathbb{R}^m \to \mathbb{R}^N$ is a smooth function. If $x(t, \alpha)$ is any given solution to the initial value problem (12), then *monotone comparative dynamics*

(MCD) for this dynamic system obtains if

$$\hat{\alpha} \geq \alpha \quad \Rightarrow \quad \Big(x(t, \hat{\alpha}) \geq x(t, \alpha), \ \forall \, t \in [0, T] \Big) . \tag{13}$$

To treat this problem within our framework, we first split the time interval $[0, T]$ into n subintervals $\mathcal{I}_k = [t_{k-1}, t_k]$, $k \in \{1, \ldots, n\}$, where $t_k = kT/n$ and $n \geq 1$. Discretizing the initial value problem (12) accordingly yields

$$x_k - x_{k-1} = \left(\frac{T}{n} \right) g_k(x_k, \alpha), \quad k \in \{1, \ldots, n\}, \tag{12'}$$

where $x_k = x(t_k)$ and $g_k(\cdot, \alpha) = g(\cdot, t_k, \alpha)$. Interpreting $(x_1, \ldots, x_n)$ as an (nN)-dimensional variable and $(g_1, \ldots, g_n)$ as an (nN)-dimensional function, the discretized initial value problem (12') is of the form (1) with $f = (f_1, \ldots, f_n)$, and

$$f_k(x_1, \ldots, x_n, \alpha) = x_k - x_{k-1} - (T/n) g_k(x_k, \alpha).$$

The corresponding pseudo-gradient with respect to any direction $(\alpha, v) \in \mathcal{A} \times \mathbb{R}^m$ is

$$W_n(\alpha, v) = \left\{ w \in \mathbb{R}^N : w_{k+1} - \left(1 + \left(\frac{T}{n} \right) \frac{\partial g_k(x_k, \alpha)}{\partial x_k} \right) w_k = \left(\frac{T}{n} \right) \frac{\partial g_k(x_k, \alpha)}{\partial \alpha} \cdot v, \ 0 \leq k \leq n-1 \right\},$$

where $w_0 = 0$. By taking the limit as $n \to \infty$, we therefore obtain an N-dimensional pseudo-gradient with functional components

$$W_\infty(\alpha, v) = \left\{ w \in C^1 \left([0, T], \mathbb{R}^m \right) : \dot{w}(t) - \frac{\partial g(x, t, \alpha)}{\partial x} w = \frac{\partial g(x, t, \alpha)}{\partial \alpha} v, \ w(0) = 0, \ t \in [0, T] \right\}$$

for any $(\alpha, v) \in \mathcal{A} \times \mathbb{R}^m$. The set of directions w that constitutes the pseudo-gradient $W_\infty(\alpha, v)$ is defined by the linear initial value problem

$$\dot{w}(t) - \frac{\partial g(x(t, \alpha), t, \alpha)}{\partial x} w = \frac{\partial g(x(t, \alpha), t, \alpha)}{\partial \alpha} v, \quad w(0) = 0,$$

for all $t \in [0, T]$, any solution of which can be written in the form

$$w(t, \alpha, v) = F(t, \alpha) \left(\int_0^t F^{-1}(s, \alpha) \frac{g(x(s, \alpha), s, \alpha)}{\partial \alpha} \, ds \right) v,$$

where F is an $N \times N$ fundamental matrix[28] satisfying the matrix differential equation

$$\dot{F}(t) = \frac{\partial g(x(t, \alpha), t, \alpha)}{\partial x} F(t)$$

for all $t \in [0, T]$ and $F(0) = 1$. The positive cone $V \subset \mathbb{R}^m$ that admits MCD in the sense of (13) corresponds to the set of directions v such that $w(t, \alpha, v) \geq 0$.

[28]For $N = 1$, $F(t, \alpha) = \exp \left[\int_0^t \frac{\partial g(x(s, \alpha), s, \alpha)}{\partial x} \, ds \right]$, $t \in [0, 1]$. When $N > 1$, the fundamental matrix cannot be given in closed form.

Theorem 5 *If γ is a differentiable (α, v)-path such that*

$$F(t,\gamma(\lambda))\left(\int_0^t F^{-1}(s,\gamma(\lambda))\frac{g(x(s,\gamma(\lambda)),s,\gamma(\lambda))}{\partial\alpha}\,ds\right)\dot{\gamma}(\lambda)\geq 0, \tag{14}$$

for all $\lambda \in [0,1]$ and all $t \in [0,T]$, then

$$x(t,\gamma(\hat{\lambda})) \geq x(t,\gamma(\lambda))$$

for all $\hat{\lambda} \geq \lambda \in [0,1]$ and all $t \in [0,T]$.

As a first application of Theorem 5, suppose that α is one-dimensional, and that one wishes to find conditions under which $x(t,\alpha)$ is nondecreasing in α. Since F is positive, inequality (14) reduces to

$$\int_0^u \exp\left(-\int_0^t \frac{\partial g}{\partial x}(x(s,\alpha),s,\alpha)ds\right)\frac{\partial g}{\partial\alpha}(x(t,\alpha),t,\alpha)dt \geq 0$$

for $u \in [0,T]$. This inequality is always true if $\partial g/\partial\alpha$ is nonnegative.

Proposition 9 *If $m = 1$ and $\partial g/\partial\alpha$ is nonnegative, then $x(t,\alpha)$ is nondecreasing in α.*

This proposition can also be proved by dynamic analysis,[29] but is much simpler to prove using Theorem 5. Huggett (2003) proves a similar result in discrete time, allowing for Markov uncertainty.

As another application, suppose that x represents a firm's capital, the growth of which at time t depends on some investment return $h(x,t)$ and some exogenous factor $k(t)$. The goal is to assess the impact of $k(t)$ on the growth of x: with $\dot{x}(t) = h(x,t) + \alpha k(t)$, we wish to derive conditions on h and k so that an increase in the weight α of the external shock k increases x. If k changes sign (e.g., exhibiting cyclical behavior), $\partial g/\partial\alpha$ is not always positive, so that Proposition 9 does not apply. From Theorem 5, monotonicity obtains if

$$\int_0^u \exp\left(-\int_0^t h_x(x(s,\alpha),s)ds\right)k(t)dt \geq 0$$

for all $u \in [0,T]$. In order to obtain monotonicity, there must be some sense in which k has an overall positive effect on x, despite sometimes having locally negative effects. Thus,

[29] Let $\alpha < \beta$ and $y(t) = x(t,\beta) - x(t,\alpha)$. Then $y(0) = 0$ and $y'(t) = g(x(t,\beta),t,\beta) - g(x(t,\alpha),t,\alpha)$, so $y'(t) \geq 0$ whenever $y(t) = 0$. We wish to show that $y(t) \geq 0$ for all t. The only problem is if $y(t) = 0$ and $y'(t) = 0$ for some t. Then consider the smallest $s > t$ such that $y'(s) \neq 0$. By construction, $y(t') = 0$ on $[t,s]$, so $y'(s) \geq 0$, implying that $y'(s) > 0$. Thus, y can never become negative.

we assume that for all t, $K(t) = \int_0^t k(s)ds \geq 0$. We now use the following lemma, whose proof is omitted.[30]

Lemma 1 *If ψ is such that $\int_0^t \psi(s)ds \geq 0$ for $t \in [0,T]$ and v is a decreasing function, then $\int_0^t v(s)\psi(s)ds \geq v(t)\int_0^t \psi(s)ds$.*

Applying the lemma to $\psi(t) = k(t)$ and $v(t) = \exp\left(-\int_0^t h_x(x(s,\alpha),s)ds\right)$, which is decreasing in t if h_x is nonnegative, we have proved the following result.[31]

Proposition 10 *Suppose that $\dot{x}(t) = h(x,t) + \alpha g(t)$ and $x(0) = x_0$, with h nondecreasing in x and $\int_0^t g \geq 0$ for all t.[32] Then $x(t,\alpha)$ is nondecreasing in α for all t.*

This result can be interpreted (within a highly stylized, deterministic context) as follows: if a firm's growth is subject to exogenous global cycles, each of which has an *overall* positive effect, an increased impact of these cycles on the firm results in the firm possessing higher capital at all times.

5 Conclusion

Analyzing the monotonicity of solutions to optimization or equilibrium problems is often a difficult task. Despite the recent advances in MCS, many problems in economic theory and other fields do not fit the mold of standard techniques, which are based on supermodularity on lattices or on other ordinal concepts, and one is led to either focus on very simple models or make strong assumptions in order to obtain intuitive monotonicity results. GMA extends the previous literature on comparative-statics analysis in several directions. First, it provides a new way of *endogenously* generating parameter moves that yield monotonicity. This may result, for example, in partitions of the parameter space into regions where simple but region-dependent parameter moves do indeed yield monotonicity. It may also result in the emergence of some particular aggregate of the parameters, such as the notion of parameter closeness derived in Section 4.2. This construction is

[30] The lemma is nontrivial, since ψ has no sign requirement. Quah and Strulovici (2008) prove a different version of the lemma, which is equivalent to this one by a change of variables.

[31] To our knowledge, this result cannot be proved using existing theorems of comparative dynamics (since g can take negative values). It can be derived along the lines of Footnote 29.

[32] As usual, we assume that h and g are smooth, to guarantee existence and uniqueness of $x(\cdot,\alpha)$ for all α.

achieved thanks to the introduction of a new concept, that of a *pseudo-gradient*, which summarizes all available information (provided in the form of an information set) for the determination of monotonicity properties. Second, GMA naturally permits consideration of the monotonicity of *functions* of a solution instead of the monotonicity of the solution itself as in the standard analysis. This flexibility proves important when analyzing the monotonicity properties of complex multivariate models. Third, GMA may be used to derive conditions under which monotonicity obtains with respect to imposed moves of the parameters, for example corresponding to parameter aggregates. Fourth, the method introduced here is easily extended to derive bounds on the sensitivity of variables (or functions thereof) to parameters. Such quantitative results largely escape the reach of standard ordinal methods. Fifth, GMA introduces several concepts and insights regarding the nature of comparative statics, for example in dealing with constraints or exploiting additional information about the problem structure. Owing to its simplicity and systematic nature, GMA is a natural and potentially crucial first step to analyze monotonicity in complex problems, and is in many ways complementary to ordinal methods.

References

Amir, R., Mirman, L., Perkins, W. (1991) "One-Sector Nonclassical Optimal Growth: Optimality Conditions and Comparative Dynamics," *International Economic Review*, Vol. 32, No. 3, pp. 625–644.

Arnold, V.I., Il'yashenko, Yu.S. (1988) "Ordinary Differential Equations," in: Arnold, V.I., Anosov, D.V. (Eds.) *Dynamical Systems, Vol. I*, Encyclopaedia of Mathematical Sciences, Springer, New York, NY.

Athey, S. (2002) "Monotone Comparative Statics Under Uncertainty," *Quarterly Journal of Economics*, Vol. 117, No. 1, pp. 187–223.

Becker, R. (1983) "Comparative Dynamics in the One-Sector Optimal Growth Model," *Journal of Economic Dynamics and Control*, Vol. 6, No. 1, pp. 99–107.

Berge, C. (1963) *Topological Spaces*, Oliver and Boyd, London, UK. Reprinted by Dover Publications, Mineola, NY, in 1997.

Bertsekas, D.P. (1995) *Nonlinear Programming*, Athena Scientific, Belmont, MA.

Blackwell, D. (1951) “Comparison of Experiments,” Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, pp. 93–102.

Blackorby, C., Shorrocks, A.F. (1996) *Separability and Aggregation: The Collected Works of W. M. Gorman, Vol. I*, Oxford University Press, Oxford, UK.

Brock, W., Mirman, L. (1972) “Optimal Economic Growth and Uncertainty: The Discounted Case,” *Journal of Economic Theory*, Vol. 4, No. 3, pp. 497–513.

Chipman, J. (1977) “An empirical implication of Auspitz-Lieben-Edgeworth-Pareto complementarity,” *Journal of Economic Theory*, Vol. 14, No. 1, pp. 228–231.

Echenique, F. (2002) “Comparative Statics by Adaptive Dynamics and the Correspondance Principle,” *Econometrica*, Vol. 70, No. 2, pp. 833–844.

Edlin, A., Shannon, C. (1998) “Strict Monotonicity in Comparative Statics,” *Journal of Economic Theory*, Vol. 81, No. 1, pp. 201–219.

Guillemin, V., Pollack, A. (1974) *Differential Topology*, Prentice-Hall, Englewood Cliffs, NJ.

Hicks, J.R. (1939) *Value and Capital*, Clarendon Press, Oxford, UK.

Hugget, M. (2003) “When are Comparative Dynamics Monotone?,” *Review of Economic Dynamics*, Vol. 6, No. 1, pp. 1–11.

Jensen, M.K. (2007) “Monotone Comparative Statics in Ordered Vector Spaces,” *The B.E. Journal of Theoretical Economics*, Vol. 7, No. 1 (Topics), Article 35.

Luenberger, D.G. (1969) *Optimization by Vector Space Methods*, Wiley, New York, NY.

Mendelsohn, R., Sobel, M. (1981) “Capital Accumulation and the Optimization of Renewable Resources,” *Journal of Economic Theory*, Vol. 23, No. 2, pp. 243–260.

Milgrom, P., Roberts, J. (1990) “Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities,” *Econometrica*, Vol. 58, No. 6, pp. 1255–1277.

Milgrom, P., Shannon, C. (1994) “Monotone Comparative Statics,” *Econometrica*, Vol. 62, No. 1, pp. 157–180.

Olley, G.S., Pakes, A. (1996) “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, Vol. 64, No. 6, pp. 1263–1297.

Persico, N. (2000) "Information Acquisition in Auctions," *Econometrica*, Vol. 68, No. 1, pp. 135–148.

Postnikov, M.M. (1987) *Smooth Manifolds*, Lectures in Geometry, Semester III, Mir Publishers, Moscow, Russia.

Quah, J.K.-H. (2007) "The Comparative Statics of Constrained Optimization Problems," *Econometrica*, Vol. 75, No. 2, pp. 401–431.

Quah, J.K.-H., Strulovici, B.H. (2008) "Comparative Statics, Informativeness, and the Interval-Dominance Order," Working Paper, Oxford University.

Samuelson, P. (1941) "The Stability of Equilibrium: Comparative Statics and Dynamics," *Econometrica*, Vol. 9, No. 2, pp. 97–120.

Samuelson, P. (1947) *Foundations of Economic Analysis*, Harvard University Press, Cambridge, MA.

Smirnov, G.V. (2002) *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Providence, RI.

Sørenson, P.N. (2007) "Simple Utility Functions with Giffen Demand," *Economic Theory*, Vol. 31, pp. 367–370.

Strulovici, B., Weber, T.A. (2008) "Monotone Comparative Statics: Geometric Approach," *Journal of Optimization Theory and Applications*, Vol. 137, No. 3, pp. 641–673.

Topkis, D. (1968) *Ordered Optimal Solutions*, Doctoral Dissertation, Stanford University, Stanford, CA.

Topkis, D. (1998) *Supermodularity and Complementarity*, Princeton University Press, Princeton, NJ.

Veinott, A.F. (1989) *Lattice Programming*, Unpublished notes for lectures initially delivered at Johns Hopkins University.

Vives, X. (1990) "Nash Equilibrium with Strategic Complementarities," *Journal of Mathematical Economics*, Vol. 19, No. 3, pp. 305–321.