Economic Theory (2007) 30: 385–394
DOI 10.1007/s00199-005-0062-6

EXPOSITA NOTE

M. Hossein Partovi · Michael R. Caputo

Existence of a universal comparative statics matrix for differentiable optimization problems

Received: 4 March 2005 / Accepted: 8 November 2005 / Published online: 15 December 2005

Abstract A universal existence theorem is established that yields exhaustive and constraint-free comparative statics information for a general, differentiable optimization problem in the preferred form of a semidefinite matrix. It subsumes all comparative statics formulations of differentiable optimization problems. Its relationship to comparative statics methods extant is established.

Keywords Comparative statics · Optimization problems · Rank inequalities

JEL Classification Numbers C600 · C610 · C650

1 Introduction

All differentiable comparative statics information stemming from the underlying optimization hypothesis originates from the semidefinite nature of the Hessian matrix of the Lagrangian function with respect to the decision variables at the optimum point. Given a differentiable constrained optimization problem, then, there arises the question of whether there exists a universal construction that yields the corresponding comparative statics information exhaustively, i.e., from which any

We are grateful to an anonymous referee for a particularly helpful comment on the contents of this paper.

M. Hossein Partovi
Department of Physics and Astronomy, California State University,
Sacramento, CA 95819–6041, USA
E-mail: hpartovi@csus.edu

Michael R. Caputo (✉)
Department of Economics, University of Central Florida,
P.O. Box 161400, Orlando, FL 32816–1400, USA
E-mail: mcaputo@bus.ucf.edu

other comparative statics matrix can be derived. The answer is affirmative, as we will show.

Specifically, we shall construct a semidefinite comparative statics matrix without any recourse to the geometry of the parameter space or to generalized compensated derivatives, in contrast to Partovi and Caputo (2006). The present construction, therefore, is almost entirely anchored in decision space, and achieves its results by relying on intrinsic, projective techniques to fully embody the semidefiniteness information contained in the constrained Hessian matrix of the underlying Lagrangian function. The result is a constraint-free *exhaustive* comparative statics matrix, that is to say, a semidefinite matrix which is linear in the partial derivatives of the decision functions with respect to the parameters. We shall refer to this matrix as the *universal* comparative statics matrix (hereafter the universal CSM). Because the universal CSM fully represents the underlying semidefiniteness information, it is expected to subsume any other CSM for the optimization problem which uses the same parameter set. This expectation is indeed born out for the existing methodologies as will be established in Section 3. Not surprisingly, the universal CSM is not as convenient to construct or apply, nor is it as intuitive to interpret, as is the CSM constructed by the method of generalized compensated derivatives utilized in Partovi and Caputo (2006). Nonetheless, from a theoretical point of view the existence theorem for the universal CSM established below is a more fundamental result than the comparative statics theorems of Silberberg (1974, Eq. (10)) , Hatta (1980, Theorems 6 and 7), and Partovi and Caputo (2006, Theorem 1), insofar as it provides a constructive proof for the existence of a constraint-free, exhaustive CSM for a generic differentiable optimization problem.

Because of the prominent role played by projection matrices in the proof of our existence theorem, we pause to briefly present some of their essential properties in the ensuing section. This strategy has the desirable effects of rendering the paper self-contained and increasing its accessibility. The section then culminates in a technical lemma that plays a crucial role in the proof of the aforementioned existence theorem.

2 Background material and a preliminary technical result

The class of constrained optimization problems under consideration is given by

$$\max_{\mathbf{x}\in \mathrm{D}} \left\{ f(\mathbf{x}, \mathbf{a}) \;\; \text{s.t.} \;\; g^k(\mathbf{x}, \mathbf{a}) = 0, \;\; k = 1, 2, \ldots, K \right\}, \tag{1}$$

where $f(\cdot) \in C^{(2)}$ and $g^k(\cdot) \in C^{(2)}$ on $\mathrm{D} \times \mathrm{P}^{\text{open}} \subset \Re^M \times \Re^N$ for $k = 1, 2, \ldots, K$, and $M, N > K$. We refer to D as the decision space and P^{open} as the parameter space in what follows. To avoid trivialities, we assume that the standard constraint qualification condition holds at the optimum point, i.e., the rank of the $K \times M$ Jacobian matrix $\partial \mathbf{g}(\mathbf{x}, \mathbf{a})/\partial \mathbf{x}$ is equal to K at the critical points of problem (1) for all $\mathbf{a} \in \mathrm{P}^{\text{open}}$. Furthermore, we assume that there exists a unique vector of $C^{(1)}$ decision functions $\mathbf{x}(\cdot)$, with values defined by

$$\mathbf{x}(\mathbf{a}) \stackrel{\text{def}}{=} \arg\max_{\mathbf{x}\in \mathrm{D}} \left\{ f(\mathbf{x}, \mathbf{a}) \;\; \text{s.t.} \;\; g^k(\mathbf{x}, \mathbf{a}) = 0, \;\; k = 1, 2, \ldots, K \right\}, \tag{2}$$

and that $\mathbf{x}(\mathbf{a}) \in \operatorname{int} \mathrm{D}$ for each value of $\mathbf{a} \in \mathrm{P}^{\mathrm{open}}$. We denote the corresponding value of the Lagrange multiplier vector by $\boldsymbol{\lambda}(\mathbf{a}) \in \Re^K$. Similarly, we assume that the set of parameter space normal vectors $\nabla^{\mathbf{a}} g^k(\mathbf{x}(\mathbf{a}), \mathbf{a})$, $k = 1, 2, \ldots, K$, is linearly independent, that is to say, the rank of the $K \times N$ Jacobian matrix $\partial \mathbf{g}(\mathbf{x}(\mathbf{a}), \mathbf{a})/\partial \mathbf{a}$ is equal to K for all $\mathbf{a} \in \mathrm{P}^{\mathrm{open}}$. By the implicit function theorem, this assumption implies, for all $\mathbf{x}(\mathbf{a}) \in \operatorname{int} \mathrm{D}$ and for each value of $\mathbf{a} \in \mathrm{P}^{\mathrm{open}}$, that the normal hyperplane in parameter space, $\mathrm{N}^{\mathrm{P}}(\mathbf{a})$, is of dimension K, i.e., $\dim[\mathrm{N}^{\mathrm{P}}(\mathbf{a})] = K$, and that its orthogonal complement $\mathrm{T}^{\mathrm{P}}(\mathbf{a})$, the tangent hyperplane, is of dimension $N - K$, i.e., $\dim[\mathrm{T}^{\mathrm{P}}(\mathbf{a})] = N - K$. We also define the value of the Lagrangian function $L(\cdot)$ as

$$L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{a}) \stackrel{\text{def}}{=} f(\mathbf{x}, \mathbf{a}) + \sum_{k=1}^{K} \lambda_k g^k(\mathbf{x}, \mathbf{a}). \tag{3}$$

We note in passing that the above set of assumptions is redundant. For example, the linear independence of the set of vectors $\nabla^{\mathbf{a}} g^k(\mathbf{x}(\mathbf{a}), \mathbf{a})$ implies the same for the set of vectors $\nabla^{\mathbf{x}} g^k(\mathbf{x}(\mathbf{a}), \mathbf{a})$, a fact that can be established by a contrapositive argument.

At this juncture it is convenient to introduce the compact notation $g^k_{,i}(\mathbf{x}, \mathbf{a}) \stackrel{\text{def}}{=} \frac{\partial}{\partial x_i} g^k(\mathbf{x}, \mathbf{a})$ and $g^k_{,\alpha}(\mathbf{x}, \mathbf{a}) \stackrel{\text{def}}{=} \frac{\partial}{\partial a_\alpha} g^k(\mathbf{x}, \mathbf{a})$. This notational convention implies that a subscript occurring to the right of a comma signifies partial differentiation. Moreover, Latin subscripts are used to denote differentiation with respect to decision variables, while Greek indices are used to denote differentiation with respect to parameters. Furthermore, the symbol "†" is used to denote transposition.

Let us now turn to a few basic properties of projection matrices. First, recall that the set of K normal vectors in decision space, namely the gradient vectors $\nabla^{\mathbf{x}} g^k(\mathbf{x}(\mathbf{a}), \mathbf{a})$, $k = 1, 2, \ldots, K$, is linearly independent by the aforementioned prototype constraint qualification assumption imposed on problem (1). Equivalently, by Theorem **4G** of Strang (1993), the $K \times K$ matrix defined by $\Gamma(\mathbf{a}) \stackrel{\text{def}}{=} \mathrm{G}(\mathbf{a})^{\dagger}\mathrm{G}(\mathbf{a})$, with elements $\Gamma_{kk'}(\mathbf{a}) \stackrel{\text{def}}{=} \sum_{i=1}^{M} g^k_{,i}(\mathbf{x}(\mathbf{a}), \mathbf{a})\, g^{k'}_{,i}(\mathbf{x}(\mathbf{a}), \mathbf{a})$, is symmetric and invertible, where $\mathrm{G}(\mathbf{a}) \stackrel{\text{def}}{=} [\nabla^{\mathbf{x}} g^1(\mathbf{x}(\mathbf{a}), \mathbf{a})\,|\nabla^{\mathbf{x}} g^2(\mathbf{x}(\mathbf{a}), \mathbf{a})\,|\cdots|\nabla^{\mathbf{x}} g^K(\mathbf{x}(\mathbf{a}), \mathbf{a})]$ is the $M \times K$ matrix whose columns are comprised of the linearly independent normal vectors in decision space. By the implicit function theorem this implies, for all $\mathbf{x}(\mathbf{a}) \in \operatorname{int} \mathrm{D}$ and for each value of $\mathbf{a} \in \mathrm{P}^{\mathrm{open}}$, that the normal hyperplane in decision space, $\mathrm{N}^{\mathrm{D}}(\mathbf{a})$, is of dimension K, i.e., $\dim[\mathrm{N}^{\mathrm{D}}(\mathbf{a})] = K$, and its orthogonal complement $\mathrm{T}^{\mathrm{D}}(\mathbf{a})$, the tangent hyperplane, is of dimension $M - K$, that is, $\dim[\mathrm{T}^{\mathrm{D}}(\mathbf{a})] = M - K$. Consequently, any decision space vector, say $\mathbf{w} \in \mathrm{D}$, possesses a unique orthogonal decomposition given by $\mathbf{w} = \mathbf{w}^n + \mathbf{w}^t$ corresponding to the above decomposition, where $\mathbf{w}^n \in \mathrm{N}^{\mathrm{D}}(\mathbf{a})$ and $\mathbf{w}^t \in \mathrm{T}^{\mathrm{D}}(\mathbf{a})$. Then by Theorem 2.8 of Härdle and Simar (2003), there exists a symmetric and idempotent $M \times M$ *projection matrix* $\mathrm{P}(\mathbf{a})$ defined by the formula $\mathrm{P}(\mathbf{a}) \stackrel{\text{def}}{=} \mathrm{G}(\mathbf{a})\Gamma(\mathbf{a})^{-1}\mathrm{G}(\mathbf{a})^{\dagger}$, with entries $\mathrm{P}_{ij}(\mathbf{a}) \stackrel{\text{def}}{=} \sum_{k=1}^{K}\sum_{k'=1}^{K} g^k_{,i}(\mathbf{x}(\mathbf{a}), \mathbf{a})\Gamma^{-1}_{kk'}(\mathbf{a}) g^{k'}_{,j}(\mathbf{x}(\mathbf{a}), \mathbf{a})$, such that $\mathbf{w}^n = \mathrm{P}(\mathbf{a})\mathbf{w} \in \mathrm{N}^{\mathrm{D}}(\mathbf{a})$, $\mathbf{w}^t = \mathrm{Q}(\mathbf{a})\mathbf{w} \in \mathrm{T}^{\mathrm{D}}(\mathbf{a})$, and $\mathbf{w}^{t\dagger}\mathbf{w}^n = 0$, where $\mathrm{Q}(\mathbf{a}) \stackrel{\text{def}}{=} [\mathrm{I}_M - \mathrm{P}(\mathbf{a})]$ and I_M is the identity matrix of order M. The geometry associated with projection matrices occurs precisely because $\mathrm{P}(\mathbf{a})$ projects onto the column space of the $M \times K$ matrix $\mathrm{G}(\mathbf{a})$, the columns of which are the linearly independent decision space normal vectors $\nabla^{\mathbf{x}} g^k(\mathbf{x}(\mathbf{a}), \mathbf{a})$, $k = 1, 2, \ldots, K$, which form a basis for $\mathrm{N}^{\mathrm{D}}(\mathbf{a})$, the normal hyperplane in decision space. Moreover, $\mathrm{Q}(\mathbf{a})$ projects onto the orthogonal complement of the column space of $\mathrm{G}(\mathbf{a})$, namely $\mathrm{T}^{\mathrm{D}}(\mathbf{a})$, the tangent hyperplane in decision space.

With these basic properties about projection matrices established, we now turn to the above mentioned technical result. As will be seen in due course, it plays an essential role in establishing the central result of this paper.

Lemma 1 *The symmetric* $M \times M$ *matrix* $\Lambda(\mathbf{a})$, *whose elements are given by* $\Lambda_{mm'}(\mathbf{a}) \overset{\text{def}}{=} \sum_{i=1}^{M} \sum_{j=1}^{M} \mathrm{Q}_{im}(\mathbf{a}) L_{,ij}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a})\, \mathrm{Q}_{jm'}(\mathbf{a})$, $m, m' = 1, 2, \ldots, M$, *is negative semidefinite.*

Proof First observe that the second-order necessary condition of the optimization problem defined by Eq. (1) is given by

$$\sum_{i=1}^{M} \sum_{j=1}^{M} v_i L_{,ij}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}) v_j \le 0$$
$$\forall\, \mathbf{v} \in \Re^M \ni \sum_{i=1}^{M} v_i g^k_{,i}(\mathbf{x}(\mathbf{a}), \mathbf{a}) = 0, k = 1, 2, \ldots, K. \tag{4}$$

For any vector $\mathbf{w} \in \Re^M$, recall that $\mathbf{w}^n = \mathrm{P}(\mathbf{a})\mathbf{w} \in \mathrm{N}^{\mathrm{D}}(\mathbf{a})$ and $\mathbf{w}^t = \mathrm{Q}(\mathbf{a})\mathbf{w} \in \mathrm{T}^{\mathrm{D}}(\mathbf{a})$, which implies that $\mathbf{w}^{t\dagger}\mathbf{w}^n = 0$, as is readily verified. Upon setting $v_i = \sum_{m=1}^{M} \mathrm{Q}_{im}(\mathbf{a}) w_m$, $i = 1, 2, \ldots, M$, and observing that $\mathrm{Q}(\mathbf{a})\mathrm{G}(\mathbf{a}) = \mathbf{0}_{M\times K}$, it follows from Eq. (4) that

$$\sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{m=1}^{M} \sum_{m'=1}^{M} w_m \mathrm{Q}_{im}(\mathbf{a}) L_{,ij}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a})\, \mathrm{Q}_{jm'}(\mathbf{a}) w_{m'} \le 0\, \forall\, \mathbf{w} \in \Re^M. \tag{5}$$

Since $L(\cdot) \in C^{(2)}$, the $M \times M$ matrix $\Lambda(\mathbf{a})$ is symmetric, where

$$\Lambda_{mm'}(\mathbf{a}) \overset{\text{def}}{=} \sum_{i=1}^{M} \sum_{j=1}^{M} \mathrm{Q}_{im}(\mathbf{a}) L_{,ij}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a})\, \mathrm{Q}_{jm'}(\mathbf{a}),$$
$$m, m' = 1, 2, \ldots, M. \tag{6}$$

Moreover, by Eq. (5) we may conclude that $\Lambda(\mathbf{a})$ is in fact negative semidefinite since the vector $\mathbf{w} \in \Re^M$ is arbitrary. □

This concludes the discussion of technical preliminaries.

3 The main theorem

The central result of the paper, to wit, the existence of a comprehensive and universal CSM, may now be stated.

Theorem 1 *The constrained optimization problem defined by Eq. (1) et seq. admits of an* $N \times N$ *constraint-free positive semidefinite comparative statics matrix* $\mathrm{U}(\mathbf{a})$, *the typical element of which is given by*

$$\mathrm{U}_{\mu\nu}(\mathbf{a}) \stackrel{\text{def}}{=} \sum_{i=1}^{M}\sum_{j=1}^{M} x_{i,\mu}(\mathbf{a})\mathrm{Q}_{ij}(\mathbf{a}) \Bigg[L_{,j\nu}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) - \sum_{l=1}^{M} L_{,jl}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) \sum_{k=1}^{K}\sum_{k'=1}^{K} g^{k}_{,l}\left(\mathbf{x}(\mathbf{a}), \mathbf{a}\right) \Gamma^{-1}_{kk'}(\mathbf{a}) g^{k'}_{,\nu}\left(\mathbf{x}(\mathbf{a}), \mathbf{a}\right) \Bigg],$$

$\mu, \nu = 1, 2, \ldots, N$. *Moreover, rank* $[\mathrm{U}(\mathbf{a})] \leq \min(M - K, N)$.

Proof Differentiate the identity $L_{,i}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) = f_{,i}\left(\mathbf{x}(\mathbf{a}), \mathbf{a}\right) + \sum_{k=1}^{K} \lambda_k(\mathbf{a}) g^k_{,i}\left(\mathbf{x}(\mathbf{a}), \mathbf{a}\right) = 0$, $i = 1, 2, \ldots, M$, with respect to a_μ, multiply the result by the element $\mathrm{Q}_{li}(\mathbf{a})$, sum over i, and use the fact that $\sum_{i=1}^{M} \mathrm{Q}_{li}(\mathbf{a}) g^k_{,i}\left(\mathbf{x}(\mathbf{a}), \mathbf{a}\right) = 0$ for $k = 1, 2, \ldots, K$, to get

$$\sum_{i=1}^{M} \left\{ \mathrm{Q}_{li}(\mathbf{a}) \left[L_{,i\mu}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) + \sum_{j=1}^{M} L_{,ij}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) x_{j,\mu}(\mathbf{a}) \right] \right\} = 0. \tag{7}$$

Next, premultiply Eq. (7) by $x_{l,\nu}(\mathbf{a})$ and sum over l. This operation yields

$$\sum_{i=1}^{M}\sum_{l=1}^{M} x_{l,\nu}(\mathbf{a}) \Bigg\{ \mathrm{Q}_{li}(\mathbf{a}) \Bigg[L_{,i\mu}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) + \sum_{j=1}^{M} L_{,ij}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}) \sum_{m=1}^{M} [\mathrm{P}_{jm}(\mathbf{a}) + \mathrm{Q}_{jm}(\mathbf{a})] x_{m,\mu}(\mathbf{a}) \Bigg] \Bigg\} = 0. \tag{8}$$

Observe that we have inserted the term $\delta_{jm} = \mathrm{P}_{jm}(\mathbf{a}) + \mathrm{Q}_{jm}(\mathbf{a})$ in Eq. (8), where δ_{jm} is the Kronecker delta. This insertion allows Eq. (8) to be rearranged as

$$\begin{aligned} \mathrm{U}_{\nu\mu}(\mathbf{a}) &\stackrel{\text{def}}{=} \sum_{i=1}^{M}\sum_{l=1}^{M} x_{l,\nu}(\mathbf{a}) \Bigg\{ \mathrm{Q}_{li}(\mathbf{a}) \Bigg[L_{,i\mu}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) + \sum_{j=1}^{M} L_{,ij}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) \sum_{m=1}^{M} \mathrm{P}_{jm}(\mathbf{a}) x_{m,\mu}(\mathbf{a}) \Bigg] \Bigg\} \\ &= -\sum_{l=1}^{M}\sum_{m=1}^{M}\sum_{i=1}^{M}\sum_{j=1}^{M} x_{l,\nu}(\mathbf{a}) \left\{ \mathrm{Q}_{il}(\mathbf{a}) L_{,ij}\left(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}\right) \mathrm{Q}_{jm}(\mathbf{a}) \right\} \\ &\quad \times x_{m,\mu}(\mathbf{a}), \quad \mu, \nu = 1, 2, \ldots, N, \end{aligned} \tag{9}$$

where the symmetry of the projection matrix Q($\mathbf{a}$) has been used on the right-hand side. By Lemma 1, the matrix U($\mathbf{a}$) is positive semidefinite. Next, use the definition of P($\mathbf{a}$) to arrive at

$$\sum_{m=1}^{M} \mathrm{P}_{jm}(\mathbf{a}) x_{m,\mu}(\mathbf{a}) = \sum_{m=1}^{M}\sum_{k=1}^{K}\sum_{k'=1}^{K} g^{k}_{,j}(\mathbf{x}(\mathbf{a}),\mathbf{a})\,\Gamma^{-1}_{kk'}(\mathbf{a}) g^{k'}_{,m}(\mathbf{x}(\mathbf{a}),\mathbf{a})\, x_{m,\mu}(\mathbf{a}). \tag{10}$$

Then use the identity $g^{k}_{,\mu}(\mathbf{x}(\mathbf{a}),\mathbf{a}) + \sum_{m=1}^{M} g^{k}_{,m}(\mathbf{x}(\mathbf{a}),\mathbf{a})\, x_{m,\mu}(\mathbf{a}) = 0$, which results from differentiating the identity $g^{k}(\mathbf{x}(\mathbf{a}),\mathbf{a})) = 0$ with respect to a_μ, to simplify Eq. (10), and substitute the result in $\mathrm{U}_{\mu\nu}(\mathbf{a})$. The result is

$$\mathrm{U}_{\mu\nu}(\mathbf{a}) \stackrel{\text{def}}{=} \sum_{i=1}^{M}\sum_{l=1}^{M} x_{l,\mu}(\mathbf{a}) \Bigg\{ \mathrm{Q}_{li}(\mathbf{a}) \Bigg[L_{,i\nu}(\mathbf{x}(\mathbf{a}),\boldsymbol{\lambda}(\mathbf{a}),\mathbf{a}) - \sum_{j=1}^{M} L_{,ij}(\mathbf{x}(\mathbf{a}),\boldsymbol{\lambda}(\mathbf{a}),\mathbf{a}) \sum_{k=1}^{K}\sum_{k'=1}^{K} g^{k}_{,j}(\mathbf{x}(\mathbf{a}),\mathbf{a})\Gamma^{-1}_{kk'}(\mathbf{a}) g^{k'}_{,\nu}(\mathbf{x}(\mathbf{a}),\mathbf{a}) \Bigg]\Bigg\},$$

$\mu, \nu = 1, 2, \ldots, N$, which establishes the semidefiniteness result once the dummy indices are changed as follows: $l \to i$, $i \to j$, and $j \to l$.

The upper bound to *rank* [U($\mathbf{a}$)] can be derived from the standard theorems that (a) the rank of an $N^R \times N^C$ matrix cannot exceed $\min(N^R, N^C)$, and (b) the rank of a product cannot exceed that of any of its factors. Since U($\mathbf{a}$) is an $N \times N$ matrix, its rank cannot exceed N. Moreover, because $\dim[\mathrm{T}^{\mathrm{D}}(\mathbf{a})] = M - K$, *rank* $[\mathrm{Q}(\mathbf{a})] = M - K$. Using the second formula of Eq. (9), and then recalling that $\partial\mathbf{x}(\mathbf{a})/\partial\mathbf{a}$ is an $M \times N$ matrix and that $\left[L_{,ij}(\mathbf{x}(\mathbf{a}),\boldsymbol{\lambda}(\mathbf{a}),\mathbf{a})\right]$ is an $M \times M$ matrix, the rank result follows from the above quoted rank theorems. □

It is worth noting that the rank of any CSM associated with the optimization problem (1) is limited by (a) the dimension of the *constrained* decision space, namely $M - K$, i.e., by $\dim[\mathrm{T}^{\mathrm{D}}(\mathbf{a})] = M - K$, and (b) the size of the parameter set, videlicet, N. Thus no CSM can be of higher rank than the smaller of these two integers. Note also that zeros in the spectrum of the $M \times M$ Hessian matrix $\left[L_{,ij}(\mathbf{x}(\mathbf{a}),\boldsymbol{\lambda}(\mathbf{a}),\mathbf{a})\right]$ or other "exceptional" circumstances may reduce the rank of a CSM below the value deduced under general conditions. Thus all that can be established in general is an upper bound to *rank* [U($\mathbf{a}$)]. To see this, observe that the $M \times N$ matrix $\partial\mathbf{x}(\mathbf{a})/\partial\mathbf{a}$ that appears in the second formula of Eq. (9) can have an arbitrarily small rank, including zero, implying the same for U($\mathbf{a}$). For example, if the optimization problem is given by $\max_{\mathbf{x}}[F^1(\mathbf{x}) + F^2(\mathbf{a})]$, then the decision functions do not depend on the parameters, thereby causing $\partial\mathbf{x}(\mathbf{a})/\partial\mathbf{a}$ to vanish identically. This implies the vanishing of U($\mathbf{a}$) and its rank.

We conclude our discussion of the universal CSM U($\mathbf{a}$) by establishing its relation to the CSM $\Omega(\mathbf{a})$ in Theorem 1 of Partovi and Caputo (2006). This relationship creates the link from U($\mathbf{a}$) to the comparative statics theorems of Silberberg (1974, Eq. (10)) and Hatta (1980, Theorems 6 and 7), for Partovi and Caputo (2006) have already established the connection between the latter two authors' results and $\Omega(\mathbf{a})$; see also Caputo (1999) in this regard.

To begin, let $\mathbf{t}^{\alpha} \in \mathrm{T}^{\mathrm{P}}(\mathbf{a}) \subset \Re^{N}$, $\alpha = 1, 2, \ldots, N-K$, be the set of so-called isovectors used in the construction of $\Omega(\mathbf{a})$ in Theorem 1 of Partovi and Caputo (2006). We intend to show that the relationship between $\Omega(\mathbf{a})$ and $\mathrm{U}(\mathbf{a})$ is given by $\Omega_{\alpha\beta}(\mathbf{a}) = \sum_{\mu=1}^{N}\sum_{\nu=1}^{N} t_{\mu}^{\alpha}\mathrm{U}_{\mu\nu}(\mathbf{a})t_{\nu}^{\beta}$. To establish the veracity of this claim, first observe that by Lemma 1 of Partovi and Caputo (2006), $x_{i;\alpha}(\mathbf{a}) \stackrel{\text{def}}{=} \sum_{\mu=1}^{N} t_{\mu}^{\alpha} x_{i,\mu}(\mathbf{a}) \in \mathrm{T}^{\mathrm{D}}(\mathbf{a})$ for $\alpha = 1, 2, \ldots, N-K$ and $i = 1, 2, \ldots, M$, thereby implying that $\sum_{i=1}^{M} x_{i;\alpha}(\mathbf{a})\mathrm{P}_{ij}(\mathbf{a}) = \sum_{i=1}^{M} x_{i;\alpha}(\mathbf{a})\mathrm{P}_{ji}(\mathbf{a}) = 0$ for $\alpha = 1, 2, \ldots, N-K$ and $j = 1, 2, \ldots, M$, since $\mathrm{P}(\mathbf{a})$ is the $M \times M$ projection matrix onto $\mathrm{N}^{\mathrm{D}}(\mathbf{a})$. This is a crucial property in the present derivation, and one that clearly shows how the application of the generalized compensated derivatives of Partovi and Caputo (2006) obviates the use of projection matrices. Next, using Eq. (9), transpose the indices μ and ν to get $\mathrm{U}_{\mu\nu}(\mathbf{a})$, substitute $\mathrm{U}_{\mu\nu}(\mathbf{a})$ into $\sum_{\mu=1}^{N}\sum_{\nu=1}^{N} t_{\mu}^{\alpha}\mathrm{U}_{\mu\nu}(\mathbf{a})t_{\nu}^{\beta}$, change the dummy indices of summation according to: $l \to i, i \to j$, and $j \to l$, and then employ the definitions $x_{i;\alpha}(\mathbf{a}) \stackrel{\text{def}}{=} \sum_{\mu=1}^{N} t_{\mu}^{\alpha} x_{i,\mu}(\mathbf{a})$ and $L_{,j;\beta}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}) \stackrel{\text{def}}{=} \sum_{\nu=1}^{N} t_{\nu}^{\beta} L_{,j\nu}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a})$. This sequence of operations yields the expression

$$\begin{aligned}\sum_{\mu=1}^{N}\sum_{\nu=1}^{N} t_{\mu}^{\alpha}\mathrm{U}_{\mu\nu}(\mathbf{a})t_{\nu}^{\beta} &= \sum_{i=1}^{M}\sum_{j=1}^{M} x_{i;\alpha}(\mathbf{a})\left[\delta_{ij} - \mathrm{P}_{ij}(\mathbf{a})\right] L_{,j;\beta}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}) \\ &+ \sum_{i=1}^{M}\sum_{j=1}^{M} x_{i;\alpha}(\mathbf{a})\mathrm{Q}_{ij}(\mathbf{a}) \\ &\times \left[\sum_{l=1}^{M} L_{,jl}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}) \sum_{m=1}^{M} \mathrm{P}_{lm}(\mathbf{a}) x_{m;\beta}(\mathbf{a})\right], \end{aligned} \tag{11}$$

for $\alpha, \beta = 1, 2, \ldots, N-K$. Using the aforementioned orthogonality property, that is to say, $\sum_{i=1}^{M} x_{i;\alpha}(\mathbf{a})\mathrm{P}_{ij}(\mathbf{a}) = \sum_{i=1}^{M} x_{i;\alpha}(\mathbf{a})\mathrm{P}_{ji}(\mathbf{a}) = 0, \alpha = 1, 2, \ldots, N-K$, and then taking account of the symmetry of $\mathrm{U}(\mathbf{a})$ or $\Omega(\mathbf{a})$, we find that Eq. (11) reduces to

$$\begin{aligned}\sum_{\mu=1}^{N}\sum_{\nu=1}^{N} t_{\mu}^{\alpha}\mathrm{U}_{\mu\nu}(\mathbf{a})t_{\nu}^{\beta} &= \sum_{i=1}^{M} x_{i;\beta}(\mathbf{a}) L_{,i;\alpha}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}), \mathbf{a}) \\ &= \Omega_{\alpha\beta}(\mathbf{a}), \alpha, \beta = 1, 2, \ldots, N-K,\end{aligned}$$

which is the intended result.

The relationship between $\Omega(\mathbf{a})$ and $\mathrm{U}(\mathbf{a})$ established above provides further insight into their rank properties. Recalling the theorems pertaining to the rank of a matrix quoted in the proof of Theorem 1, we conclude that the rank of $\Omega(\mathbf{a})$ is no larger than the smaller of the ranks of the $N \times N$ matrix $\mathrm{U}(\mathbf{a})$ and the $N \times (N-K)$ matrix $\mathrm{T} \stackrel{\text{def}}{=} \left[\mathbf{t}^1 \middle| \mathbf{t}^2 \middle| \cdots \middle| \mathbf{t}^{N-K}\right]$. In light of the fact that $rank(\mathrm{T}) = N-K$, we arrive at the result of Theorem 4 of Partovi and Caputo (2006), scilicet, that $rank\ [\Omega(\mathbf{a})] \leqslant \min(M-K, N, N-K) = \min(M-K, N-K)$, since $N-K \leq N$. This derivation also shows that when $N-K < N \leqslant M-K$, the rank of $\mathrm{U}(\mathbf{a})$ will in general be larger than that of $\Omega(\mathbf{a})$ because $rank\ [\mathrm{T}] = N-K$. The latter is in turn caused by the fact that the process of compensation employed by Partovi and Caputo (2006) prohibits differentiation with respect to the normal directions in parameter space, while the present construction via projection matrices does not.

A model for which the rank of $\mathrm{U}(\mathbf{a})$ will typically be larger than that of $\Omega(\mathbf{a})$ is the classical nonlinear general equilibrium model given by

$$\max_{\substack{x_{ij} \\ i=1,2,\ldots,I;\, j=1,2,\ldots,J}} \left\{ \sum_{j=1}^{J} p_j f^j(x_{1j}, x_{2j}, \ldots, x_{Ij}) \text{ s.t.} \sum_{j=1}^{J} x_{ij} = x_i,\ \ i = 1, 2, \ldots, I \right\},$$

where x_{ij} is the amount of factor i used in the production of good j, $f^j(\cdot)$ is the production function of good j, p_j is the market price of good j, and x_i is the fixed amount of factor i. Observe that in this model we have $M = IJ$, $N = I + J$, and $K = I$. Therefore, if $2I + J \leqslant IJ$, then $N - K < N \leqslant M - K$, and the rank of $\mathrm{U}(\mathbf{a})$ will in general be larger than that of $\Omega(\mathbf{a})$, as is straightforward to confirm.

With the main theorem now established, we turn to a brief application of it.

4 An application

Our objective here is to illustrate the use of Theorem 1, as well as its relation to Theorem 1 of Partovi and Caputo (2006), in the context of a familiar model, namely the utility maximization problem. We will find that the universal CSM derived for the utility maximization problem bears little resemblance to the Slutsky form and defies any attempt at intuitive understanding. Upon applying the reduction procedure established in Section 3 to this CSM, on the other hand, we recover the familiar Slutsky form. What this exercise underscores is that (a) information-equivalent CSM's can have different appearances, ranks, and intuitive significance, and (b) a judicious choice of form can significantly enhance the utility of a CSM. It is this latter feature that distinguishes the income-compensated derivative as the key ingredient in the Slutsky construction. More generally, the generalized compensated derivatives of Partovi and Caputo (2006) serve the purpose of guiding one's choice of the form in which the CSM emerges.

The problem under consideration is the archetype utility maximization problem, scilicet

$$\max_{\mathbf{x} \in \Re^{N-1}_{++}} \left\{ U(\mathbf{x}) \ \text{ s.t. } \ m - \mathbf{p}^{\dagger}\mathbf{x} = 0 \right\},$$

where $\mathbf{p} \in \Re^{N-1}_{++}$ is the price vector, $m > 0$ is the consumer's income, $\mathbf{a} \overset{\text{def}}{=} (\mathbf{p}, m) \in \Re^{N}_{++}$ is the parameter vector, $\mathbf{x}(\mathbf{a})$ is the value of the Marshallian demand function, and $\lambda(\mathbf{a})$ is the value of the marginal utility of income function. Observe that there is but one constraint function $g(\cdot)$, given by $g(\mathbf{x}, \mathbf{a}) \overset{\text{def}}{=} m - \mathbf{p}^{\dagger}\mathbf{x}$, and that $M = N - 1$. After a lengthy computation, we find that according to Theorem 1, that the universal CSM takes the following unfamiliar form:

$$\mathrm{U}_{\mu\nu}(\mathbf{a}) = -\sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \frac{\partial x_i(\mathbf{a})}{\partial p_\mu}\left[\delta_{ij} - \frac{p_i p_j}{\Gamma(\mathbf{a})}\right] \times \left[\lambda(\mathbf{a})\delta_{j\nu} + \sum_{l=1}^{N-1} U_{,jl}\left(\mathbf{x}(\mathbf{a})\right)\frac{p_l x_\nu}{\Gamma(\mathbf{a})}\right], \mu, \nu = 1, 2, \ldots, N-1, \tag{12}$$

$$\mathrm{U}_{\mu\nu}(\mathbf{a}) = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \frac{\partial x_i(\mathbf{a})}{\partial p_\mu}\left[\delta_{ij} - \frac{p_i p_j}{\Gamma(\mathbf{a})}\right] \times \left[\sum_{l=1}^{N-1} U_{,jl}\left(\mathbf{x}(\mathbf{a})\right)\frac{p_l}{\Gamma(\mathbf{a})}\right], \mu = 1, 2, \ldots, N-1; \nu = N, \tag{13}$$

$$\mathrm{U}_{\mu\nu}(\mathbf{a}) = -\sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \frac{\partial x_i(\mathbf{a})}{\partial m}\left[\delta_{ij} - \frac{p_i p_j}{\Gamma(\mathbf{a})}\right] \times \left[\lambda(\mathbf{a})\delta_{j\nu} + \sum_{l=1}^{N-1} U_{,jl}\left(\mathbf{x}(\mathbf{a})\right)\frac{p_l x_\nu}{\Gamma(\mathbf{a})}\right], \mu = N; \nu = 1, 2, \ldots, N-1, \tag{14}$$

$$\mathrm{U}_{\mu\nu}(\mathbf{a}) = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \frac{\partial x_i(\mathbf{a})}{\partial m}\left[\delta_{ij} - \frac{p_i p_j}{\Gamma(\mathbf{a})}\right]\left[\sum_{l=1}^{N-1} U_{,jl}(\mathbf{x}(\mathbf{a}))\frac{p_l}{\Gamma(\mathbf{a})}\right], \mu, \nu = N, \tag{15}$$

where $\Gamma(\mathbf{a}) \overset{\text{def}}{=} \sum_{n=1}^{N-1} p_n^2$. As is plainly obvious, this CSM does not resemble the Slutsky matrix, although it is a fully valid positive semidefinite CSM for the archetype utility maximization problem. Notice also that it is (a) redundant, in the sense that it is a rank $N-2$ matrix of order N due to the homogeneity of degree zero of the Marshallian demand functions in prices and income, and (b) unobservable, due to the appearance of the second-order partial derivatives of the utility function in its expression.

In order to transform Eqs. (12)–(15) into their familiar Slutsky form, we follow the procedure used in section 3 to demonstrate the relationship between $\mathrm{U}(\mathbf{a})$ and $\Omega(\mathbf{a})$. Because $\dim[\mathrm{T}^{\mathrm{P}}(\mathbf{a})] = N-1$ we must therefore determine $N-1$ vectors that form a basis for $\mathrm{T}^{\mathrm{P}}(\mathbf{a})$. A natural and simple set of basis vectors is given by $\mathbf{t}^\alpha = (0_1, 0_2, \ldots, 0_{\alpha-1}, 1_\alpha, 0_{\alpha+1}, \ldots, 0_{N-1}, x_\alpha(\mathbf{a}))$, $\alpha = 1, 2, \ldots, N-1$, where the subscript on a numerical element of the vector indicates its position within the vector. That $\mathbf{t}^\alpha \in \mathrm{T}^{\mathrm{P}}(\mathbf{a})$ for $\alpha = 1, 2, \ldots, N-1$ can be verified by showing that $\nabla^{\mathbf{a}} g(\mathbf{x}(\mathbf{a}), \mathbf{a}) \cdot \mathbf{t}^\alpha = 0$ for $\alpha = 1, 2, \ldots, N-1$, while the fact that the only solution to the homogeneous system of linear equations $\sum_{\alpha=1}^{N-1} c_\alpha \mathbf{t}^\alpha = \mathbf{0}_N \in \Re^N$ is the null vector $\mathbf{c} = \mathbf{0}_{N-1} \in \Re^{N-1}$ confirms that the vectors $\mathbf{t}^\alpha$, $\alpha = 1, 2, \ldots, N-1$, form a basis for $\mathrm{T}^{\mathrm{P}}(\mathbf{a})$. The classical Slutsky matrix can now be derived from the formula $\Omega_{\alpha\beta}(\mathbf{a}) = \sum_{\mu=1}^{N}\sum_{\nu=1}^{N} t_\mu^\alpha \mathrm{U}_{\mu\nu}(\mathbf{a}) t_\nu^\beta$. Note that the Slutsky matrix $\Omega(\mathbf{a})$ is of order $N-1$ while its rank is $N-2$, so it too is redundant. On the other hand, the Slutsky form has the important empirical advantage of being observable, in contrast to U(**a**), as remarked above.

5 Conclusion

We have constructed a universal CSM for an arbitrary differentiable optimization problem, as summarized in Theorem 1. For theoretical purposes, this theorem is a fundamental result, as it defines a framework in which any other differentiable comparative statics method must be subsumed. Not surprisingly, the universal construction method is not as practically convenient, or intuitively appealing, as the method of generalized compensated derivatives of Partovi and Caputo (2006), a point we emphasized in section 4. On the other hand, recalling that the universal construction is applicable to any differentiable system governed by an extremum principle, a category which includes numerous physical and mathematical problems in quite diverse fields of inquiry, one realizes the vast scope and reach of the theorem.

References

Caputo, M.R.: The relationship between two dual methods of comparative statics. J Econ Theory **84**, 243–250 (1999)

Härdle, W., Simar, S.: Applied multivariate statistical analysis. Berlin Heidelberg New York: Springer 2003

Hatta, T.: Structure of the correspondence principle at an extremum point. Rev Econ Stud **47**, 987–997 (1980)

Partovi, M.H., Caputo, M.R.: A complete theory of comparative statics for differentiable optimization problems. Metroeconomica **57**, (2006, forthcoming)

Silberberg, E.: A revision of comparative statics methodology in economics, or, how to do comparative statics on the back of an envelope. J Econ Theory **7**, 159–172 (1974)

Strang, G.: Introduction to linear algebra. Wellesley: Wellesley-Cambridge Press 1993