Exact and Useful Optimization Methods for Microeconomics

Balder, Erik

Mathematical Institute, University of Utrecht

21 April 2008

Exact and Useful Optimization Methods for Microeconomics

Erik J. Balder

Abstract This paper points out that the treatment of utility maximization in current textbooks on microeconomic theory is deficient in at least three respects: breadth of coverage, completeness-cum-coherence of solution methods and mathematical correctness. Improvements are suggested in the form of a Kuhn-Tucker type theorem that has been customized for microeconomics. To ensure uniqueness of the optimal solution *stringent quasiconcavity*, an apparently new adaptation of the notion of strict quasiconcavity, is introduced. It improves upon an earlier notion formulated by Aliprantis, Brown and Burkinshaw. The role of the domain of differentiability of the utility function is emphasized. This is not only to repair a widespread error in the microeconomic literature but also to point out that this domain can be chosen sensibly in order to include the maximization of certain nondifferentiable utility functions, such as Leontiev utility functions. To underscore the usefulness of the optimality conditions obtained here, five quite different instances of utility maximization are completely solved by a single coherent method.

1 Introduction

In the currently popular textbooks [8, 10, 11, 13] on microeconomic theory and mathematical economics, all of which profess to be rigorous and precise, the treatment of the fundamental subject of utility maximization would seem to show considerable shortcomings in the following three respects: breadth of coverage, completeness-cum-coherence of solution methods and mathematical correctness. Similar shortcomings also show up in related textbooks with a less formal orientation [9, 14]. These shortcomings are briefly reviewed and explained in Remark 2. As a consequence, the reader will search in vain in the standard literature on microeconomic theory and mathematical economics for a method to derive the Marshallian

Mathematical Institute, University of Utrecht, P.O. Box 80010, 3508TA Utrecht, The Netherlands, e-mail: E.J.Balder@uu.nl

demand function that meets the following criteria: it must be coherent, complete and correct (i.e., based on generally accepted principles of correct mathematical reasoning starting from, say, the Kuhn-Tucker and Weierstrass' theorems), and it must be operationally useful by being applicable to at least the following five basic instances of standard utility functions, where $\ell \in \mathbb{N}$ stands for the number of commodities: (i) Cobb-Douglas utility function, $\ell > 2$, (ii) CES utility function (more precisely, this instance will be separated below into three different cases (iia), (iib) and (iic)), $\ell > 2$, (iii) linear utility function with positive coefficients, $\ell \geq 2$, (iv) the utility function $u(x_1, x_2) = x_1^2(x_2 + 1)$ or any similar one leading to partial corner point solutions, and (v) Leontiev utility function, $\ell \geq 2$. The actual state of affairs in the standard literature is even more perplexing than stated above: the problems signaled in Remark 2 already affect the classical instances (i)-(ii) and already for $\ell = 2$. The appendix offers a more detailed account of this.

To motivate the above five instances, I first observe that there are very good reasons for including instances with more than two commodities in the preceding list. Indeed, although not many of the aforementioned references point this out, the standard utility maximization problem in microeconomics (see (1)) can immediately be reduced to an optimization problem over an interval if $\ell = 2$. This follows from the fact that its optimal solution must be budget-balanced; cf. Theorem 1(a). Now an interval optimization problem is quite elementary. One can solve it very directly by means of sign diagrams of the derivative, possibly supported by the use of computer algebra packages. This explains why I take $\ell > 2$ in instances (i)-(ii): it is a minimal check of operational usefulness. Next, instances (iii)-(iv) require the solution method to be able to detect multiple solutions and corner point solutions. The latter form a standard subject in intermediate microeconomics courses [4]. Remarkably, even though at least three of the above-mentioned textbooks indicate the possibility of corner point or multiple solutions in the form of illustrations (see figure 3.D.4 in [10], figure 1.10 in [8] and figure 1.17 of [13]), none of them manages to treat this subject in an analytically satisfactory and complete manner. Finally, instance (v) has a nondifferentiable utility function, which presents a quite unusual operational challenge for the solution method. I observe that in [8, 10, 11, 13] this Leontiev instance is not included in the mainstream approach.

In section 3 I shall develop an efficient solution method that meets all of the above criteria; it is based on optimality results for general utility maximization that are presented in section 2. This development started in my Utrecht lecture notes, because I was dissatisfied with what the standard microeconomics literature had to offer. Essentially, the optimality results take the form of a Kuhn-Tucker theorem, Theorem 1, that has been customized for microeconomics. To a high degree such customization depends on a special property of utility functions that is commonly found in microeconomics: they are strictly increasing. I observe beforehand that the classical Cobb-Douglas and CES instances (i)-(ii) require that the model can deal with utility functions that are defined on the nonnegative orthant $\mathbb{R}_+^\ell$ only, possibly with nondifferentiability on the boundary $\mathbb{R}_+^\ell \setminus \mathbb{R}_{++}^\ell$. Together with the applicability of my method to the nondifferentiable instance (v), this indicates that careful consideration of the differentiability domain of the utility function plays an important role

in this paper (cf. Remark 2). Parts (a) to (c) of the customized Kuhn-Tucker theorem come close to what is done on [8, p. 23 ff.], but, as my applications to instances (i), (iia)-(iib) and (iv) will show in particular, it is part (d) of Theorem 1, absent in [8], which makes a considerable difference. This part uses an apparently new adaptation of the notion of strict quasiconcavity, called *stringent quasiconcavity*, which improves upon a related earlier notion by Aliprantis, Brown and Burkinshaw [2] (see Remark 3). It allows the method, in its handling of sufficient conditions for optimality and uniqueness, to go beyond interior optimal solutions and makes it possible to derive corner point solutions in a rigorous, coherent and efficient way, for instance in instance (iv), whose utility function is *not* strictly quasiconcave on $\mathbb{R}_+^2$.

In sum, this paper improves on the usual literature on utility maximization in microeconomics by presenting a custom-made Kuhn-Tucker theorem that exploits the usual strict monotonicity of utility functions in microeconomics via the new notion of stringent quasiconcavity and that pays careful attention to the utility function's domain of differentiability.

2 Customized optimality results for microeconomics

Let $u : \mathbb{R}_+^\ell \to \mathbb{R}$ be a continuous function, the *utility function*. Let Ω be an open set that is contained in $\mathbb{R}_+^\ell$ (whence in the strictly positive orthant $\mathbb{R}_{++}^\ell$); the function u is supposed to be differentiable on Ω [6]. In section 3 I shall choose for Ω the strictly positive orthant $\mathbb{R}_{++}^\ell$ so as to treat instances (i)-(iv), but to deal with the Leontiev instance (v) I am going to choose Ω differently. Throughout I suppose that u is *strictly increasing* on $\mathbb{R}_+^\ell$; that is to say, for every x and x' in $\mathbb{R}_+^\ell$ the following must hold: if $x_i > x_i'$ for every $i = 1, 2, \ldots, n$, then $u(x) > u(x')$. For $p \in \mathbb{R}_{++}^\ell$ (*price vector*) and $y \in \mathbb{R}_+$ (*income*) the consumer's *utility maximization problem* is as follows:

$$\text{maximize } u(x) \text{ over all } x \in \mathbb{R}_+^\ell \text{ such that } p \cdot x \leq y \tag{1}$$

and it is desired to determine *all* global optima (if any). The problem is well-defined because $0 \in B$. Here $B := \{x \in \mathbb{R}_+^\ell : p \cdot x \leq y\}$ stands for the feasible set of this problem, which is called the *budget set*. Of course, for $y = 0$ the problem is trivial, because then $B = \{0\}$. I shall also use $B_0 := \{x \in \mathbb{R}_+^\ell : p \cdot x = y\}$ to denote the so-called *budget plane*. A vector x in $\mathbb{R}_+^\ell$ is said to be *budget-balanced* if $p \cdot x = y$, that is to say, if it belongs to the budget plane B_0. Recall that $u : \mathbb{R}_+^\ell \to \mathbb{R}$ is defined to be *quasiconcave* on $\mathbb{R}_+^\ell$ if $u(tx + (1-t)x') \geq \min(u(x), u(x'))$ for every pair $x, x' \in \mathbb{R}_+^\ell$ and every $t \in [0, 1]$. I shall say that the function $u : \mathbb{R}_+^\ell \to \mathbb{R}$ is *stringently quasiconcave* if it is both quasiconcave on $\mathbb{R}_+^\ell$ and if it has the following property, which I shall call *property* (S): for every pair $x, x' \in \mathbb{R}_+^\ell$ with $x \neq x'$ and $u(x) = u(x') > u(0)$ one has $u(\frac{1}{2}x + \frac{1}{2}x') > u(x) = u(x')$. It is worth noting that in the main text of this paper the stringent quasiconcavity notion is *always* considered on the full domain $\mathbb{R}_+^\ell$ of the utility function u (but see Remark 1(ii)). Observe that stringent quasiconcavity is a

modification of the classical notion of strict quasiconcavity which excludes certain points that are sub-optimal (if $y > 0$) and only works with points x and x' at the same utility level. Clearly, a sufficient condition for $u : \mathbb{R}_+^\ell \to \mathbb{R}$ to be stringently quasiconcave is that it is quasiconcave on $\mathbb{R}_+^\ell$ and strictly quasiconcave when restricted to the set $C := \{x \in \mathbb{R}_+^\ell : u(x) > u(0)\}$ (i.e., if $u(tx + (1-t)x') > \min(u(x), u(x'))$ for every pair $x, x' \in C$, $x \neq x'$, and every $t \in (0, 1)$). A related sufficient condition for stringent quasiconcavity is given below in Proposition 1; it shows the utility function to be stringently quasiconcave in some of the previously mentioned instances. Because it turns out to work so very well for utility maximization, it is surprising that the stringent quasiconcavity notion was neither introduced before in economics nor implicitly exploited in another guise. Yet this appears to be the case.

Theorem 1. (*a*) *The consumer's utility maximization problem (1) has an optimal solution. Moreover, every optimal solution is budget-balanced.*

(*b*) *Suppose that (1) has an optimal solution x^* which is such that $x^* \in \Omega$. Then there exists $\lambda \geq 0$ such that*

$$\nabla u(x^*) = \lambda p. \tag{2}$$

(*c*) *If $x^* \in \Omega$ is budget-balanced and such that (2) holds for some $\lambda > 0$, then x^* is an optimal solution of (1), provided that u is quasiconcave on $\mathbb{R}_+^\ell$.*

(*d*) *If u has property* (*S*), *then (1) has a unique optimal solution. In particular, if u is stringently quasiconcave, then any budget-balanced $x^* \in \Omega$ for which (2) holds for some $\lambda > 0$, is the unique optimal solution of (1).*

This result has the familiar makeup of results in optimization theory: existence, followed by necessary conditions for optimality that are sharpened into sufficient conditions and even a uniqueness condition. Simple examples show that the above formulation is sharp; for instance, taking $\ell = 1$, $u(x) := (x - 1)^3$, $y = 1$ and $p = 1$ shows that the possibility $\lambda = 0$ cannot be excluded in part (*b*), etc.

Lemma 1. *Suppose that u is quasiconcave on $\mathbb{R}_+^\ell$. Then for every $x \in \Omega$ and $x' \in \mathbb{R}_+^\ell$*

$$u(x) \leq u(x') \text{ implies } \nabla u(x) \cdot (x' - x) \geq 0.$$

Proof. For $t \in [0, 1]$ let $\phi(t) := u((1 - t)x + tx')$. Then $u(x) \leq u(x')$ implies $\phi(t) \geq u(x) = \phi(0)$. So $\phi(t)$ attains a minimum over $[0, 1]$ for $t = 0$. Also ϕ, the composition of u and a linear mapping, is differentiable from the right at 0, because u is differentiable at $x \in \Omega$. It follows that $\phi'(0) \geq 0$, as a first order necessary condition for optimality. The desired inequality then follows by the chain rule. □

Proof. Now Theorem 1 can be proved.

(*a*) Existence of an optimal solution x^* follows by the Weierstrass theorem, because u is continuous and B is a nonempty compact set. For $y = 0$ the identity $B = B_0 = \{0\}$ causes budget balancedness to hold trivially. For $y > 0$, $x^* \notin B_0$ would imply $p \cdot x^* < y$. So setting $\tilde{x}_i := x_i^* + t$ for $i = 1, \ldots, \ell$ would result in a contradiction for $t > 0$ sufficiently small, because then $\tilde{x} \in B$ and $u(\tilde{x}) > u(x^*)$ by the fact that u increases strictly.

(b) The hypothesis $x^* \in \Omega$ implies that x^* is also an optimal solution of the auxiliary optimization problem

$$\text{maximize } u(x) \text{ over all } x \in \Omega \text{ with } p \cdot x \leq y,$$

which has only one inequality constraint. At this stage it is traditional to invoke a version of the Kuhn-Tucker theorem and preferably – to avoid incomplete or inaccurate arguments – a version that allows open domains of definition for its functions: e.g., Theorems 1.D.3-1.D.4 in [13] can be used. However, the following extremely simple argument can be used instead, based on budget-balancedness in Theorem 1(a) (incidentally, a different but similarly simple argument, based on polar considerations, can also be given for the consumer's expenditure minimization problem to prove a result that is very similar to Theorem 1). By part (a) the bundle x^* is also an optimal solution of the auxiliary optimization problem

$$\text{maximize } u(x) \text{ over all } x \in \Omega \text{ with } p \cdot x = y.$$

Here the constraint can be expressed as $x_\ell = (y - \sum_{i=1}^m p_i x_i)/p_\ell$, where $m := \ell - 1$. Hence, the previous optimization problem can be rephrased in the variables $x_1, \ldots, x_m$ as follows

$$\text{maximize } \tilde{u}(x_1, \ldots, x_m) := u(x_1, \ldots, x_m, \frac{y - \sum_{i=1}^m p_i x_i}{p_\ell}) \text{ over all } (x_1, \ldots, x_m) \in \tilde{\Omega}.$$

Here $\tilde{\Omega}$ is the set of all $(x_1, \ldots, x_m) \in \mathbb{R}^m$ such that $(x_1, \ldots, x_m, (y - \sum_{i=1}^m p_i x_i)/p_\ell) \in \Omega$. By continuity of linear maps it is clear that $\tilde{\Omega}$ is open. Because $\tilde{u}$ attains its maximum over the open set $\tilde{\Omega}$ at $\tilde{x}^* := (x_1^*, \ldots, x_m^*)$, it follows from standard calculus results and the chain rule that $0 = \partial \tilde{u}(\tilde{x}^*)/\partial x_i = \partial u(x^*)/\partial x_i - p_i p_\ell^{-1} \partial u(x^*)/\partial x_\ell$ for $i = 1, \ldots, \ell - 1$. So setting $\lambda := p_\ell^{-1} \partial u(x^*)/\partial x_\ell$ gives (2).

(c) Suppose that $x^* \in B_0 \cap \Omega$ were not optimal. Then there would be $\hat{x} \in B$ such that $u(\hat{x}) > u(x^*)$. So $u((1-t)\hat{x}) > u(x^*)$ would hold for $t > 0$ small enough, by continuity of u. By Lemma 1 the quasiconcavity hypothesis for u implies

$$\nabla u(x^*) \cdot (x - x^*) \geq 0 \text{ for every } x \in \mathbb{R}_+^\ell \text{ with } u(x) \geq u(x^*).$$

By $\nabla u(x^*) = \lambda p$ and $\lambda > 0$, it then follows that $p \cdot ((1-t)\hat{x} - x^*) \geq 0$, whence $p \cdot \hat{x} > (1-t)p \cdot \hat{x} \geq y$ by the given budget-balancedness of x^* (observe that $p \cdot \hat{x} > 0$ by $\hat{x} \neq 0$). This contradicts $\hat{x} \in B$.

(d) First, if $y = 0$ then $B = \{0\}$, so $x^* = 0$ is the unique optimal solution. Next, if $y > 0$, then $u(x^*) > u(0)$ must hold for any optimal solution x^*, because $(t, t, \ldots, t)$ belongs to B for $t > 0$ small enough and because u is strictly monotone. Now suppose that x^* and x^{**} were two different optimal solutions of (1). Then $u(x^*) = u(x^{**}) = $ optimal value of (1) and $u(x^*) > u(0)$ by the previous argument. Define $\hat{x} := \frac{1}{2}x^* + \frac{1}{2}x^{**}$; then $\hat{x} \in B$ and property (S) gives $u(\hat{x}) > \frac{1}{2}u(x^*) + \frac{1}{2}u(x^{**}) = u(x^*)$. This contradicts the optimality of x^*. So the optimal solution is unique. The final part

of the statement is an immediate consequence of combining part (c) with uniqueness. □

Remark 1. (i) Without monotonicity in Theorem 1, the budget-balancedness of the optimal solution can obviously not be maintained, but a fair part of Theorem 1 continues to hold when u is nondecreasing and the details are as follows. In part (a) the existence of at least one budget-balanced optimal solution is still guaranteed and part (b) continues to hold as stated. Part (c) of Theorem 1 remains meaningful by linking the additional possibility $x^* \in B\backslash B_0$ with the multiplier $\lambda = 0$. In that case the optimality of x^* can be guaranteed if u, next to being nondecreasing, is concave on $\mathbb{R}_+^\ell$ (the example $\ell = 1$, $x^* = 1$, $u(x) := (x-1)^3$, $y = 2$ and $p = 1$ shows that mere quasiconcavity is insufficient in this situation). Finally, part (d) is without significance: to have property (S), a nondecreasing u must be strictly increasing, a situation that is already covered by Theorem 1(d) itself.

(ii) Some utility functions have $\mathbb{R}_{++}^\ell$ as their natural domain of definition. Let $u :$ $\mathbb{R}_{++}^\ell \to \mathbb{R}$ be such a function and suppose that u is continuous and strictly increasing on $\mathbb{R}_{++}^\ell$, as well as differentiable on some open set $\Omega \subset \mathbb{R}_{++}^\ell$. To avoid trivialities, the new situation requires $y > 0$. Then parts (b)-(d) of Theorems 1 continue to hold, naturally with quasiconcavity on $\mathbb{R}_+^\ell$ replaced by quasiconcavity on $\mathbb{R}_{++}^\ell$ and with property (S) redefined as follows: for every pair $x, x' \in \mathbb{R}_{++}^\ell$ with $x \neq x'$ and $u(x) = u(x')$ one has $u(\frac{1}{2}x + \frac{1}{2}x') > u(x) = u(x')$. An obvious sufficient condition for these two properties to hold is that u be strictly quasiconcave. As shown by the example $\ell = 2$, $u(x_1, x_2) := x_1 + x_2$, part (a) of Theorem 1 needs adjustment: an optimal solution is guaranteed to exist under the extra condition that the set $C_v := \{x \in \mathbb{R}_{++}^\ell : u(x) \geq v\}$ is closed for every $v \in u(\mathbb{R}_{++}^\ell)$. Namely, given p and $y > 0$, fix any $\bar{x} \in \mathbb{R}_{++}^\ell$ with $p \cdot \bar{x} \leq y$ and set $\bar{v} := u(\bar{x})$. Then

$$\sup_{x \in \mathbb{R}_{++}^\ell, p \cdot x \leq y} u(x) = \sup_{x \in C_{\bar{v}}, p \cdot x \leq y} u(x)$$

and on the right side a continuous function is maximized over a nonempty compact set.

Remark 2. A different application of the Kuhn-Tucker theorem is obtained if, instead of working with the above open set $\Omega \subset \mathbb{R}_+^\ell$, one uses a model where the utility function u is defined and differentiable on an open set Ω' (for instance, this could be $\mathbb{R}^\ell$ itself) that *contains* $\mathbb{R}_+^\ell$. In that case optimality of x^* for (1) can be expressed equivalently as optimality of x^* for the following optimization problem:

maximize $u(x)$ over all $x \in \Omega'$ with $p \cdot x \leq y$ and $-x_i \leq 0$, $i = 1, \dots, \ell$.   (3)

Precisely such an application was chosen in [9, 10, 11, 13], but in doing so one forms a model that no longer applies to all Cobb-Douglas or CES utility functions, because for general parameter values such functions are not differentiable at points in $\mathbb{R}_+^\ell \backslash \mathbb{R}_{++}^\ell$! In [9, p. 131 ff.], [10, p. 50 ff.], [11, Theorem 22.1, Example 22.1] and [14, section 2.2] (the latter reference discusses this for profit maximization) this has led to imprecise or incorrect formulations of their necessary first order optimality

conditions and, in the case of [9, 10, 11], to an incorrect application to the Cobb-Douglas instance (i).[1] The appendix to this paper provides more details to support this claim. Apart from this shortfall, the derivation of these optimality conditions is standard and can be found in [10, 11]: by applying the Kuhn-Tucker theorem [13] to (3), which has $\ell + 1$ inequality constraints, one now obtains as a first-order necessary condition for optimality

$$\frac{\partial u(x^*)}{\partial x_i} \leq \lambda p_i, \text{ with equality if } x_i^* > 0, \, i = 1, \ldots, \ell$$

instead of (2) and this holds for every optimal x^* in Ω', a set that now includes the boundary of $\mathbb{R}_+^\ell$. For utility functions that fit into this model, *but only for those*, this formulation gives meaningful and somewhat sharper results. These remarks do not affect [8]; however, the optimality results in that reference presuppose that the optimal solution belongs to $\mathbb{R}_{++}^\ell$, as is exemplified by the application to the Cobb-Douglas instance in exercise 1.20 of [8]. That is not enough to allow a treatment of the corner point solution instances (iii)-(iv). It reflects a common shortcoming of the above references: except for [14, p. 57], none would seem to treat corner point solutions analytically.

From part (d) of Theorem 1 it is evident that stringent quasiconcavity can help to solve the optimization problem (1), but it should be kept in mind that this is very much due to the hypothesis that the utility function u is strictly increasing: property (S) ignores bundles at the lowest utility level, which is $u(0)$. For operational use I shall state a simple sufficient condition for stringent quasiconcavity. To prepare for it, I observe that the properties of u in Theorem 1 cause the range of u to be an interval, namely $[u(0), u_\infty)$, where $u_\infty := \sup_{x \in \mathbb{R}_+^\ell} u(x)$ (this notation allows for the possibility that u_∞ equals $+\infty$). To see this, define $\psi(t) := u(t, t, \ldots, t)$ for $t \geq 0$; this is a continuous, strictly increasing function on $\mathbb{R}_+$. Clearly, $u_\infty = \sup_{t \geq 0} \psi(t)$ and the supremum cannot be attained. Because for any $x \in \mathbb{R}_+^\ell$ the value $u(x)$ lies between $\psi(0)$ and $\psi(t)$ for some $t > 0$ sufficiently large (this follows by strict monotonicity and continuity of u), the intermediate value theorem can be invoked to finish the argument.

Proposition 1. *If $\{x \in \mathbb{R}_+^\ell : u(x) > u(0)\}$ is convex and if there exists a strictly increasing function $h : (u(0), u_\infty) \to \mathbb{R}$ such that the composition mapping $x \mapsto h(u(x))$ is strictly quasiconcave on $\{x \in \mathbb{R}_+^\ell : u(x) > u(0)\}$, then u is stringently quasiconcave.*

Proof. By a previous remark, it is enough to prove that u is quasiconcave on $\mathbb{R}_+^\ell$ and strictly quasiconcave on $\{x \in \mathbb{R}_+^\ell : u(x) > u(0)\}$. Let $\alpha \in \mathbb{R}$. If $\alpha \leq u(0)$, then $\{u \geq \alpha\} := \{x \in \mathbb{R}_+^\ell : u(x) \geq \alpha\} = \mathbb{R}_+^\ell$. If $\alpha > u(0)$, then for any $x, x' \in \{u \geq \alpha\} \subset \{u > u(0)\}$, $x \neq x'$, and for any $t \in (0, 1)$ the given strict quasiconcavity property

---

[1] Notwithstanding its general Kuhn-Tucker Theorem 1.D.3, reference [13] considers utility maximization only for utility functions defined on all of $\mathbb{R}^\ell$ (see its pp. 134-135); on pp. 223-224 this has resulted in an ad hoc solution of instance (i).

implies $h(u(tx + (1-t)x')) > \min(h(u(x)), h(u(x'))) \geq h(\alpha)$. So $u(tx + (1-t)x') > \alpha$ by strict monotonicity of h. For the set $\{u \geq \alpha\}$ I conclude that it equals $\mathbb{R}^{\ell}_+$ for every $\alpha \leq u(0)$ and that it is strictly convex for every $\alpha > u(0)$. Hence, u is certainly quasiconcave on $\mathbb{R}^{\ell}_+$ and the desired strict quasiconcavity of u on $\{u > u(0)\}$ also follows from the previous conclusion. $\qquad\square$

Example 1 (*i*) A utility function $u : \mathbb{R}^{\ell}_+ \to \mathbb{R}_+$ of Cobb-Douglas type is given by $u(x) := \Pi^{\ell}_{i=1} x_i^{\alpha_i}$; here all $\alpha_i > 0$. This function is not strictly quasiconcave on $\mathbb{R}^{\ell}_+$ (observe that $u(x_1, 0, 0, \ldots, 0) = 0$). However, it is stringently quasiconcave and I use Proposition 1 to show this. Observe first that the set $\{x \in \mathbb{R}^{\ell}_+ : u(x) > u(0)\} = \mathbb{R}^{\ell}_{++}$ is convex. On $(u(0), u_\infty) = (0, +\infty)$ I choose $h(t) := \log(t)$; this is a strictly increasing function, which gives $h(u(x)) = \sum^{\ell}_{i=1} \alpha_i \log(x_i)$ on $\mathbb{R}^{\ell}_{++}$. Because each function $x_i \mapsto \log(x_i)$ is strictly concave on $\mathbb{R}_{++}$, it easily follows that the function $x \mapsto h(u(x))$ is strictly concave on $\mathbb{R}^{\ell}_{++}$. Therefore, it follows from Proposition 1 that u is stringently quasiconcave on $\mathbb{R}^{\ell}_+$.

(*ii*) A utility function of CES type is given by $u(x) := (\sum^{\ell}_{i=1} x_i^{\rho})^{1/\rho}$ with either (*iia*) $0 < \rho < 1$, (*iib*) $\rho < 0$ or (*iic*) $\rho > 1$. In case (*iic*) u is not stringently quasiconcave, as is easy to see. In case (*iia*) the function u is defined on $\mathbb{R}^{\ell}_+$, as usual, but in case (*iib*) its domain of definition has to be $\mathbb{R}^{\ell}_{++}$, because of the negative exponent ρ. See also Remark 1(*ii*) below. I observe that u is strictly increasing in both cases (*iia*) and (*iib*) (and in case (*iic*) as well).

(*iia*) I will show that in case (*iia*) u is stringently quasiconcave and even strictly quasiconcave on $\mathbb{R}^{\ell}_+$. First, I observe that the set $\{x \in \mathbb{R}^{\ell}_+ : u(x) > u(0)\} = \mathbb{R}^{\ell}_+ \setminus \{0\}$ is convex (because this only leaves out the origin, the argument can immediately be adapted to imply that u is even strictly quasiconcave on $\mathbb{R}^{\ell}_+$). On $(u(0), u_\infty) = (0, +\infty)$ I choose $h(t) := t^{\rho}$, a strictly increasing function; then $x \mapsto h(u(x)) = \sum^{\ell}_{i=1} x_i^{\rho}$ is easily seen to be strictly concave on $\mathbb{R}^{\ell}_+$, because each function $x_i \mapsto x_i^{\rho}$ is strictly concave on $\mathbb{R}_+$. So the conditions in Proposition 1 certainly hold; hence, u is stringently quasiconcave on $\mathbb{R}^{\ell}_+$.

(*iib*) As observed above, the domain of definition is $\mathbb{R}^{\ell}_{++}$ in this case. In particular, $u(0)$ is not well-defined, so stringent quasiconcavity loses its meaning (but see Remark 1(*ii*) below for an adaptation). However, by analogy to Proposition 1 I can still show that u is strictly quasiconcave on $\mathbb{R}^{\ell}_{++}$. Indeed, in this case $h(t) := -t^{\rho}$ is strictly increasing on $(0, +\infty)$ and $x \mapsto h(u(x)) = -\sum^{\ell}_{i=1} x_i^{\rho}$ is strictly concave on $\mathbb{R}^{\ell}_{++}$, because each $x_i \mapsto -x_i^{\rho}$ is strictly concave on $\mathbb{R}_+$ by $\rho < 0$. It now follows easily that u is strictly quasiconcave on $\mathbb{R}^{\ell}_{++}$.

(*iii*) Let $u : \mathbb{R}^2_+ \to \mathbb{R}$ be given by $u(x_1, x_2) = x_1^2(x_2 + 1)$, as in my instance (*iv*). This function is not strictly quasiconcave on $\mathbb{R}^2_+$ (note that $u(0, x_2) = 0$ for all $x_2 \geq 0$). However, u is stringently quasiconcave. To prove this, I observe first that the set $\{x \in \mathbb{R}^2_+ : u(x) > u(0)\} = \{(x_1, x_2) : x_1 > 0, x_2 \geq 0\}$ is convex. On $(u(0), u_\infty) = (0, +\infty)$ I choose $h(t) := \log(t)$, a strictly increasing function; then $x \mapsto h(u(x)) = 2\log(x_1) + \log(x_2 + 1)$ is evidently strictly concave on $\{(x_1, x_2) : x_1 > 0, x_2 \geq 0\}$ (repeat the reasoning in instance (*i*)). Hence, u is stringently quasiconcave by Proposition 1.

The above example will be used for applications of Theorem 1 to the instances (i), (iia), (iib) and (iv). It is well-known that more can be said, although this is not of importance for my applications. For instance, for $\sum_i \alpha_i \leq 1$ the Cobb-Douglas utility function in part (i) of Example 1 is concave on $\mathbb{R}_+^\ell$ and in part (iia) the same is true for the CES utility function. Results from [5] or [12] can be invoked to derive such additional properties: see Remark 4 in [5, p. 123] or see Theorem 2.5.3 and its Examples 5, 6 in [12].

Remark 3. In [2] Aliprantis, Brown and Burkinshaw consider utility maximization for a pure exchange consumer, but their analysis extends effortlessly to that for an ordinary consumer. For $\ell = 2$ they correctly and completely solve instances (i), (iia) and (iv) in a coherent way, based on using existence and necessary first order optimality conditions, similar to parts (a) and (b) of Theorem 1, aided by considerations involving strict quasiconcavity of u on $\mathbb{R}_{++}^2$: see the solutions to problems 1.2.1, 1.3.2 and 1.3.4 in [1, pp. 25-26, pp. 34-35]. Instance (iii) is not treated in [2] and instance (v) is solved in an ad hoc fashion; the former would not seem to be out of reach of the general method presented in [2], but the latter would seem to be. Thus, in the area of utility maximization the book [2], although not devoted to general microeconomic theory as such, managed to reach further than the references mentioned above and it did so flawlessly.

For purposes going beyond mere computations, Definition 1.3.4 of [2] defines a *neoclassical preference* to be a continuous preference relation on $\mathbb{R}_+^\ell$ that has certain monotonicity properties. In terms of the representing utility function of such a preference relation, which exists by [10, Proposition 3.C.1], this definition comes down to the following two possibilities: either (NC_1) u is strongly increasing and strictly quasiconcave on $\mathbb{R}_+^\ell$ or (NC_2) strongly increasing and strictly quasiconcave on $\mathbb{R}_{++}^\ell$ and $u(x) > u(x')$ for every $x \in \mathbb{R}_{++}^\ell$ and $x' \in \mathbb{R}_+^\ell \setminus \mathbb{R}_{++}^\ell$. Here I use standard terminology from [8, 10] – the one used in [2] is somewhat different. Observe that CES utility functions satisfy (NC_1) but not (NC_2) and that Cobb-Douglas utility functions satisfy (NC_2) but not (NC_1); thus, the above definition is inherently two-pronged. I observe that if possibility (NC_2) obtains, then one has actually the following special property, which seems not to have been stated explicitly in [2]: everywhere on the boundary of $\mathbb{R}_+^\ell$ the function u is equal to $u(0)$ (to see this, it suffices to compare any point on the boundary with $(t, t, \ldots, t)$ for sufficiently small $t > 0$ and use continuity and strict monotonicity of u to finish the argument). I claim that both possibilities (NC_1) and (NC_2), and therefore the definition of a neoclassical preference relation as a whole, are subsumed by the more general notion of a strictly increasing utility function that is stringently quasiconcave (incidentally, this continues to hold if in (NC_1) and (NC_2) "strongly increasing" is replaced by the less demanding "strictly increasing"). For possibility (NC_1) this is immediately obvious. As for (NC_2), by continuity of u and strict quasiconcavity of u on $\mathbb{R}_{++}^\ell$ it easily follows that u is quasiconcave on $\mathbb{R}_+^\ell$, which is the closure of $\mathbb{R}_{++}^\ell$. To show property (S), recall from the discussion of (NC_2) in the preceding lines that if $x, x' \in \mathbb{R}_+^\ell$, $x \neq x'$, have $u(x) = u(x') > u(0)$ then this implies $x, x' \in \mathbb{R}_{++}^\ell$. Therefore, the fact that possibility (NC_2) results in strict quasiconcavity of u on $\mathbb{R}_{++}^\ell$ implies $u(\frac{1}{2}x + \frac{1}{2}x') > u(x) = u(x')$. As already said, this shows that utility functions that

correspond to a neoclassical preference are stringently quasiconcave. The converse is not true, even when the neoclassical preference definition is adapted so as to encompass strictly increasing utility functions: for instance, in Example 1(iii) it was demonstrated that in instance (iv) the utility function u is stringently quasiconcave, but not strictly quasiconcave on $\mathbb{R}_+^2$; hence, for this u possibility (NC_1) is out of the question. On the other hand, (NC_2) is also impossible because of $u(1,0) \not< u(\frac{1}{2}, \frac{1}{2})$.

3 Testing for operational usefulness

Here I shall discuss the use of Theorem 1 as a means to meet my criterion of operational usefulness for utility maximization; similar applications to expenditure minimization are also possible. So my task is to derive complete solutions for the Marshallian demand function, using Theorem 1 in a coherent way, for each of the following utility functions:

(i) $u(x) = \Pi_{i=1}^{\ell} x_i^{\alpha_i}$ with all $\alpha_i > 0$,

(ii) $u(x) = (\sum_{i=1}^{\ell} a_i x_i^{\rho})^{1/\rho}$ with all $a_i > 0$ and (a) $\rho \in (0,1)$, (b) $\rho < 0$ or (c) $\rho > 1$,

(iii) $u(x) = \sum_{i=1}^{\ell} a_i x_i$, with all $a_i > 0$,

(iv) $u(x_1, x_2) = x_1^2 (x_2 + 1)$,

(v) $u(x) = \min_{1 \leq i \leq \ell} b_i x_i$, with all $b_i > 0$.

These five functions are continuous and strictly increasing. In instance (v) I suppose $\ell \geq 2$, so as to avoid overlap with instance (iii). My solution method, which is based on familiar reasoning in optimization theory [3], goes as follows. The starting point is that Theorem 1(a) guarantees that an optimal solution of (1) exists and is budget-balanced. It is useful to introduce the following term: an *optimality candidate* is a vector $x^* \in B_0 \cap \Omega$ that satisfies the first order necessary optimality condition (2). Then it follows from parts (a)-(b) of Theorem 1 that the optimal solution of (1) must be an optimality candidate, provided that it belongs to Ω. Subsequently, if u happens to be quasiconcave then part (c) applies and all optimality candidates (if there are any) are indeed optimal solutions. If in addition u has property (S), then the solution is of course complete (but only then, for I wish to determine *all* globally optimal solutions of (1) and in the previous steps the so-called *remainder set $B_0 \backslash \Omega$* has not been inspected). In sum, if the utility function is stringently quasiconcave, then an optimality candidate, when found, is immediately known to be the unique optimal solution of (1). If there does not exist an optimality candidate or if u is not quasiconcave or fails to possess property (S), a careful look at the values that u attains on the remainder set $B_0 \backslash \Omega$ is needed, and these values should be compared with the maximum value of all the optimality candidates already found (if any). For $\ell = 2$ the latter is easy, but for $\ell > 2$ it can be somewhat of a challenge: see my solution of instances (iic), (iii) and (v).

It must be mentioned that the alternative optimality conditions that I mentioned in Remark 2 can also be used, but only for instances such as (iii) and (iv), where the

utility functions are differentiable on all of $\mathbb{R}^\ell$. This gives $\ell+1$ multipliers and it is well-known from nonlinear programming [3] that one should then work with the set of active indices of $x^* \in B_0$ (i.e., those indices i for which $x_i^* = 0$). A good example of this is presented in section 4.3 of [14] for a linear technology (the derivation there is nonrigorous because it is exclusively based on using first-order necessary conditions). Another observation is that ad hoc methods to solve specific problems abound. For example, it is well-known that instance (i) can be solved by trivially eliminating the boundary of $\mathbb{R}_+^\ell$, after which one can apply the logarithmic transformation. However, such a transformation is already contained in Example $1(i)$, where it is part of a fairly systematic solution method.

Solution of instance (i) To solve problem (1) for the utility function $u(x) := \Pi_{i=1}^\ell x_i^{\alpha_i}$ on $\mathbb{R}_+^\ell$ with all $\alpha_i > 0$, I choose $\Omega = \mathbb{R}_{++}^\ell$. From Example $1(i)$ I already know that u is stringently quasiconcave. So if I can find $x^* \in B_0 \cap \Omega$ that satisfies (2), then it must be the unique optimal solution by Theorem $1(d)$. In search of such x^*, I combine (2) with $p \cdot x^* = y$ and verify concretely that the x^* found belongs to Ω. This is a simple algebraic task (incidentally, note that the possibility $\lambda = 0$ in (2) leads to $x^* \notin \Omega$). It yields $x_i^* = \alpha_i y/(\alpha p_i)$, $i = 1, \ldots, \ell$, where $\alpha := \sum_{j=1}^\ell \alpha_j$. This outcome happens to be strictly positive, so the unique optimal solution has been found.

Solution of instance (iia) To solve problem (1) for the utility function $u(x) := (\sum_{i=1}^\ell x_i^\rho)^{1/\rho}$ on $\mathbb{R}_+^\ell$ with $\rho \in (0,1)$, I choose $\Omega = \mathbb{R}_{++}^\ell$. From Example $1(iia)$ I know that u is stringently quasiconcave.

  Step 1: special case $a_i = 1$ for all i. If I can find $x^* \in B_0 \cap \Omega$ that satisfies (2), then it must be the unique optimal solution by Theorem $1(d)$. To find such x^*, I combine (2) with $p \cdot x^* = y$ and then check that their solution belongs to Ω (the possibility $\lambda = 0$ in (2) can be excluded, because it leads to nonsensical expressions). This is a simple algebraic task, which gives the strictly positive expression $x_i^* = p_i^{r-1} y/(p_1^r + \cdots + p_\ell^r)$, $i = 1, \ldots, \ell$, where $r := \rho/(\rho - 1)$. So the unique optimal solution has been determined.

  Step 2: general case. I absorb the coefficients into the variables by using the variables $\xi_i := a_i^{1/\rho} x_i$. Then the utility maximization problem for $u(x) := (\sum_{i=1}^\ell a_i x_i^\rho)^{1/\rho} = (\sum_{i=1}^\ell \xi_i^\rho)^{1/\rho}$ can be rewritten as the one already solved in step 1, provided that the prices are adapted by setting $q_i := a_i^{-1/\rho} p_i$. So by step 1 the unique optimal solution is $\xi_i^* = q_i^{r-1} y/(q_1^r + \cdots + q_\ell^r)$, i.e., $x_i^* = a_i^{1-r} p_i^{r-1} y/(a_1^{1-r} p_1^r + \cdots + a_\ell^{1-r} p_\ell^r)$.

Solution of instance (iib) To solve problem (1) for the utility function $u(x) := (\sum_{i=1}^\ell x_i^\rho)^{1/\rho}$ on $\mathbb{R}_{++}^\ell$ with $\rho < 0$, I choose $\Omega = \mathbb{R}_{++}^\ell$ and appeal to Remark $1(ii)$. Because of Example $1(iib)$, the solution method is almost exactly the same as the one presented above for case (iia): again it follows that any $x^* \in B_0 \cap \Omega$ that satisfies (2) must be the unique optimal solution and again this leads to the same expression for the optimal solution x_i^* as in case (iia).

Solution of instance (iic) To solve problem (1) for the utility function $u(x) := (\sum_{i=1}^\ell x_i^\rho)^{1/\rho}$ on $\mathbb{R}_+^\ell$ with $\rho > 1$, I choose $\Omega = \mathbb{R}_{++}^\ell$. In contrast to case (iia), the function u is not stringently quasiconcave and this explains why I follow a com-

pletely different line of attack. It concentrates on the remainder set $B_0 \cap (\mathbb{R}_+^\ell \setminus \mathbb{R}_{++}^\ell)$ and uses parts (a) and (b) of Theorem 1.

Step 1: special case $a_i = 1$ for all i. By Theorem $1(a)$ an optimal solution of (1) for $u(x) := (\sum_{i=1}^\ell x_i^\rho)^{1/\rho}$ exists and it belongs to B_0. I claim that any such optimal solution must have precisely $\ell - 1$ coordinates equal to 0. This can be proved by means of a contradiction argument. For suppose that an optimal solution x^* of (1) had precisely $k > 1$ nonzero coordinates; without loss of generality I can suppose that these are the first k coordinates $x_1^*, \ldots, x_k^*$ (or else I could relabel the coordinates). The given optimality of x^* implies a fortiori that $z^* := (x_1^*, \ldots, x_k^*)$, the vector composed of the first k coordinates of x_k^*, is the optimal solution of the following problem: maximize $\tilde{u}(z) := (\sum_{i=1}^k z_i^\rho)^{1/\rho}$ over all $z \in \mathbb{R}_+^k$ such that $\tilde{p} \cdot z \leq y$. Here $\tilde{p} := (p_1, \ldots, p_k)$. Observe that z^* belongs to $\mathbb{R}_{++}^k$ by definition of k. So I can apply Theorem $1(b)$ to the aforementioned maximization problem, which is a version of (1) with dimension k. From the theorem it follows that there exists $\lambda \geq 0$ such that $\nabla \tilde{u}(z^*) = \lambda \tilde{p}$. Writing this out, the same algebra as in case (iia) gives $z_i^* = p_i^{r-1} y / (p_1^r + \cdots + p_k^r)$, $i = 1, \ldots, k$. Because $z_i^* = x_i^*$ for all $i \leq k$ and $x_i^* = 0$ for all $i > k$, it would follow with a little algebra that $x^* = (x_1^*, \ldots, x_k^*, 0, \ldots, 0)$ satisfies $u(x^*) = y / (\sum_{i=1}^k p_i^r)^{1/r}$. But the present case has $r := \rho/(\rho - 1) > 0$, so $u(y/p_1) > u(x^*)$ would hold by $k \geq 2$ and this contradiction proves my claim. It follows easily that the optimal solution, known to exist, must be one of the corner points $x^i := (0, \ldots, 0, y/p_i, 0, \ldots, 0)$, $i = 1, \ldots, \ell$, of B_0. It is now easy to conclude that any corner point x^j whose index j is such that it corresponds to the lowest price, i.e., $p_j = \min_{1 \leq i \leq \ell} p_i$, is optimal (so multiple optimal solutions can occur when two or more commodities share the lowest price).

Step 2: general case. The same absorption trick as used in case (iia), i.e., setting $\xi_i := a_i^{1/\rho} x_i$ and $q_i := a_i^{-1/\rho} p_i$, gives precisely the same optimal solution(s) as obtained in step 1, i.e., the corner point(s) of B_0 which correspond to the lowest value(s) $a_i^{-1/\rho} p_i$.

Solution of instance (iii) To solve problem (1) for the linear utility function $u(x) := \sum_{i=1}^\ell a_i x_i$, with all $a_i > 0$, I choose $\Omega = \mathbb{R}_{++}^\ell$. The function u is clearly not stringently quasiconcave, but it is so simple that a solution, essentially based on parts (a) and (b) of Theorem 1, can be provided in each of the following two cases. Case 1: $a := (a_1, \ldots, a_\ell)$ is a scalar multiple of p, say $a = \mu p$ for some $\mu \in \mathbb{R}$, and then $\mu > 0$ of course. Case 2: a is not a scalar multiple of p.

Case 1: By Theorem $1(b)$, for any optimal $x^* \in \Omega := \mathbb{R}_{++}^\ell$ there exists $\lambda \geq 0$ such that $a = \lambda p$. In the present case this was already true (take $\lambda = \mu > 0$), so every $x^* \in B_0 \cap \Omega$ is an optimality candidate. Next, u is quasiconcave, so every optimality candidate is also an optimal solution. Again this offers no news, because all optimality candidates x^* satisfy $a \cdot x^* = \mu p \cdot x^* = \mu y$ and because $a \cdot x = \mu p \cdot x \leq \mu y$ for all $x \in B$. It remains to inspect the remainder set $B_0 \setminus \Omega$. Here that is the set of all x^* in B_0 with at least one coordinate equal to zero. The value $u(x^*)$ of any $x^* \in B_0 \setminus \Omega$ is $a \cdot x^* = \mu p \cdot x^* = \mu y$, which is the same value as found before. Hence, I conclude that in case 1 the set of all optimal solutions is B_0, that is to say the union of $B_0 \cap \Omega$ and $B_0 \setminus \Omega$.

Case 2: This time the necessary condition $a = \lambda p$ is incompatible with the present situation. So there are no optimality candidates at all, which means that the optimal solution, known to exist by Theorem 1(a), must belong to the remainder set $B_0 \backslash \Omega$. Now for every x in B I have $a \cdot x = \sum_i a_i p_i^{-1} p_i x_i \leq \alpha p \cdot x \leq \alpha y$, where $\alpha := \max_i a_i / p_i$. Let I be the set of those indices i for which $a_i / p_i = \alpha$. Then I is nonempty and it is a strict subset of $\{1, \ldots, \ell\}$ (or else I would find $a = \alpha p$, which cannot be true in the present case). Any corner point $x^i := (0, \ldots, 0, y/p_i, 0, \ldots, 0)$, $i \in I$, belongs to B_0 and achieves $a \cdot x^i = \alpha y$. So I conclude $\sup_{x \in B} a \cdot x = \alpha y$. I claim that the set of optimal solutions is the intersection $B_0 \cap \cap_{i \notin I} \{x \in \mathbb{R}_+^\ell : x_i = 0\}$, which is a face of B_0. As for one inclusion, let x^* belong to this intersection. Then $a \cdot x^* = \sum_{i \in I} a_i x_i^* = \sum_{i \in I} a_i p_i^{-1} p_i x_i^* = \alpha p \cdot x^* = \alpha y$, so x^* is optimal. Conversely, let x^* be optimal. Then Theorem 1(a) implies $x^* \in B_0$ and $\alpha y = a \cdot x^*$. Hence, $\alpha y = \alpha \sum_{i \in I} p_i x_i^* + \sum_{i \notin I} a_i x_i^*$. By $p \cdot x^* = y$ this implies $\sum_{i \notin I}(a_i - \alpha p_i) x_i^* = 0$, so it follows that $x_i^* = 0$ for every $i \notin I$ (note that $a_i - \alpha p_i < 0$ for each $i \notin I$). This proves the desired characterization of the set of optimal solutions in case 2.

Solution of instance (iv) To solve problem (1) for the utility function $u(x_1, x_2) := x_1^2(x_2 + 1)$ on $\mathbb{R}_+^2$, I choose $\Omega = \mathbb{R}_{++}^2$. By Example 1$(iii)$ u is stringently quasiconcave. So if I can find $x^* \in B_0 \cap \Omega$ which satisfies (2), then it must be the unique optimal solution by Theorem 1(d). I solve $p \cdot x^* = y$ and (2), the latter amounting to $x_2^* = p_1(2p_2)^{-1} x_1^* - 1$, and then I pick the solution, if any, that belongs to Ω. From the former two equations I also obtain that $\lambda = 0$ if and only if $(x_1^*, x_2^*) = (0, y/p_2)$, a vector which is not in Ω. So I can proceed with $\lambda > 0$. I solve the two equations and this gives $x_1^* = \frac{2}{3}(y + p_2)/p_1$, $x_2^* = \frac{1}{3}(y - 2p_2)/p_2$. Because $x^* \in \Omega$ is needed for x^* to be an optimality candidate, I distinguish between the following two cases (note that $y > 2p_2$ is equivalent to $x^* \in \Omega$):

Case 1: $y > 2p_2$. Then $x^* = (\frac{2}{3}(y + p_2)/p_1, \frac{1}{3}(y - 2p_2)/p_2)$ belongs to Ω. Hence, it is an optimality candidate and by the previous argument it must also be the unique optimal optimal solution.

Case 2: $y \leq 2p_2$. In this case there is not any optimality candidate. Therefore, I know that the best u-value over the remainder set $B_0 \backslash \Omega$ is the optimal value. The set $B_0 \backslash \Omega$ contains only two vectors, the corner points $(y/p_1, 0)$ and $(0, y/p_2)$. Of these, the former one gives the highest u-value. Combining cases 1 and 2, I conclude that the optimal solution is given by

$$(x_1^*, x_2^*) = \begin{cases} (\frac{2}{3}(y + p_2)/p_1, \frac{1}{3}(y - 2p_2)/p_2) & \text{if } y > 2p_2, \\ (y/p_1, 0) & \text{if } y \leq 2p_2. \end{cases}$$

This is only partly a corner solution; graphical illustrations can help greatly to understand why this is so.

Solution of instance (v) To solve problem (1) for the utility function $u(x) = \min_{1 \leq i \leq \ell} b_i x_i$ on $\mathbb{R}_+^\ell$ with all $b_i > 0$, I choose Ω to be the open set of all $x \in \mathbb{R}_{++}^\ell$ such that $b_i x_i \neq b_j x_j$ for any $i \neq j$. Then u is locally of the form $u(x) = b_k x_k$ near each point x^* of Ω (the index k being unique to x^*), so there are no optimality candidates at all, because in such x^* condition (2) amounts to $b_k e_k = \lambda p$, which does not have

a solution (recall that $\ell \geq 2$). Here e_k again denotes the k-th unit vector. It follows that the optimal solution, known to exist, must belong to the remainder set $B_0 \backslash \Omega$. For any x in $B_0 \backslash \Omega$ either of the following two cases can occur: case 1: $x \in \mathbb{R}_{++}^{\ell}$ and there is a tie in the form of an equality $b_i x_i = b_j x_j$ for some $i \neq j$ or case 2: some coordinate of x is equal to zero. In case 2 $u(x)$ equals zero, which means the certain sub-optimality of x. So I concentrate on those vectors in $B_0 \cap \mathbb{R}_{++}^{\ell}$ that have ties. For any x in B I have $x_i \geq u(x)/b_i$ for all i, giving $y \geq \sum_i p_i x_i \geq u(x) \sum_i p_i/b_i$, i.e., $u(x) \leq y/\beta$, with $\beta := \sum_i p_i/b_i > 0$. In fact, the same reasoning shows that $u(x) < y/\beta$ whenever the set of all indices i for which $b_i x_i > u(x)$ is nonempty. So $u(x^*) = y/\beta$ for $x^* \in B$, which amounts to x^* being optimal in view of the previous lines, requires $b_i x_i^* = b_j x_j^* = u(x^*)$ for all i and j. Clearly, there is precisely one such x^* and it is given by $x_i^* = b_i^{-1} y/\beta$, $i = 1, \ldots, n$; so this is the optimal solution.

Appendix

This appendix details the findings that I summarized in Remark 2. I discuss what I found in the textbooks [8, 10, 11, 13], all of which make explicit claims about being rigorous and precise, and also in the two less formally oriented textbooks [9, 14] which were added in a second round. Reference [7] is also reviewed briefly, at the request of one of the editors. The intermediate textbook [4], which I cite for other reasons, is no part of these comparisons.

Example 1. (Mas-Colell-Whinston-Green [10]) In *Microeconomic Theory* [10, p. 50] the utility function u is defined on the consumption set $\mathbb{R}_+^{\ell}$. On p. 53, which states first-order necessary conditions for optimality in problem (1), u is required to be continuously differentiable in the following way: *"If $u(\cdot)$ is continuously differentiable, an optimal consumption bundle ... can be characterized ..."*. I quote this in full, because what is meant by differentiability of functions on closed sets can be a very confusing matter (for instance, some authors will call a function differentiable on $\mathbb{R}_+^{\ell}$ if and only if it is differentiable in every interior point of $\mathbb{R}_+^{\ell}$, i.e., if it is differentiable on $\mathbb{R}_{++}^{\ell}$; see [13, p. 81]). In Example 3.D.1O [10, p. 56] the first-order necessary conditions are applied to the Cobb-Douglas instance $u(x_1, x_2) = k x_1^{\alpha} x_2^{1-\alpha}$, with $0 < \alpha < 1$. This function does not have partial derivatives in points with either $x_1 = 0$ or $x_2 = 0$. In view of this application, it cannot be that the necessary conditions on p. 53 should contain a partial derivative $\frac{\partial u(x^*)}{\partial x_i}$ for any x^* on the boundary of $\mathbb{R}_+^{\ell}$. Yet this is so: the first-order condition (3.D.1) on p. 53 contains the following explicit clause *" with equality if $x_i^* > 0$ "*, which would make no sense if boundary

points x^* were excluded a priori. In fact, on p. 54 one finds: *"Figure 3.D.4(b) depicts the first-order condition ... when the consumer's optimal bundle x^* lies on the boundary of the consumption set."* The trouble is that the proof on p. 53 of [10] relies on the Kuhn-Tucker Theorem M.K.2 from p. 959. But the latter result is formulated for functions whose domain is all of $\mathbb{R}^\ell$. See Remark 2 for further explanation. To those readers who might object that, after all, in Example 3.D.1 boundary points are automatically sub-optimal and need not really be subjected to (3.D.1), the response must be that, mathematically/logically speaking, such an ad hoc readjustment is beside the point: necessary conditions for optimality should apply to *all* points for which they have been formulated. Here the result and the function were simply not made for each other. There is also another misplaced application of first order necessary conditions in [10]. Example 3.D.1 on p. 56 treats the Cobb-Douglas family in a well-known ad hoc fashion by the logarithmic transformation. However, the transformed utility function is once more *hors concours* for the first-order condition (3.D.1), which is formulated for a *real-valued* function that is defined on all of $\mathbb{R}^\ell_+$. Writing *"Since* $\log 0 = -\infty$ ..." on [10, p. 55] cannot alter this fact.

Example 2. (Simon-Blume [11]) Example 22.1 in *Mathematics for Economists* [11] treats utility maximization for the Cobb-Douglas instance with $\ell = 2$. In Example 22.1 one finds the following: *"We use Theorems 22.1 and 22.2 to compute the demand function ... for the Cobb-Douglas utility function $U(x_1, x_2) = x_1^a x_2^b$."* Here $a, b > 0$ and $a + b = 1$. However, Theorems 22.1 and 22.2 have the same deficiency as observed in Example 1: they are formulated for a utility function $U : \mathbb{R}^\ell_+ \to \mathbb{R}$ that is stated to be $\mathscr{C}^1$ (i.e., continuously differentiable). Just as on [10, p. 53], the condition *"... and equality holds if $x_i^* > 0$"* appears in a formula that uses partial derivatives of the utility function at the point $(x_1^*, \ldots, x_\ell^*)$. So, just as in [10] there is a mismatch: these results require derivatives to be well-defined on all of $\mathbb{R}^\ell_+$, a requirement that the above Cobb-Douglas utility function cannot meet. Compared to [10], the book by Simon and Blume offers a longer discussion of (continuous) differentiability. In fact, on p. 371 its definition is formally stated, but only for a function defined on all of $\mathbb{R}^\ell$.

Example 3. (Takayama [13]) The approach taken in *Mathematical Economics* [13] suffers from similar shortcomings. The utility maximization problem is stated on [13, p. 135] for a utility function that appears to be defined on all of $\mathbb{R}^\ell$. The first-order necessary conditions for optimality are obtained directly from the Kuhn-Tucker type result on p. 134, which works with $\mathbb{R}^\ell$ for the domain of the functions. Upon closer inspection, it is surprising that a far better Kuhn-Tucker type result is available in [13, Theorem 1.D.3]. This result allows open domains of definition for the functions involved in the optimization problem, but the aforementioned application on pp. 134-135 does not benefit from this. The subsequent treatment of the Cobb-Douglas instance in [13, pp. 223-224] (for $\ell = 2$) is presented correctly, but with no apparent connections to the material on pp. 134-135 (see also footnote 32 on p. 231 of [13]).

Example 4. (Jehle-Reny [8]) As already indicated in section 2, the first-order optimality conditions on p. 23 ff. in *Advanced Microeconomic Theory* [8] are virtually

the same as parts (a)–(c) of Theorem 1. Nonetheless, there are serious complications with the way [8, Chapter A2] presents the underpinning notions for its optimality conditions. This is because [8] fails to distinguish (total) differentiability from partial differentiability: only a formal definition of partial differentiability is presented, namely in Definition A2.1. Consequently, Example 4 of [6, p. 69] forms a counterexample against the intuitive definition of the directional derivative on [8, p. 463], which is exclusively based on partial differentiability. This can also help to explain why no chain rule for (totally) differentiable functions has been presented in [8]; yet this rule is applied. Moreover, a formula for the total differential can be found on pp. 487-488, but but neither the adjective "total" nor the term "differential" would seem to have been defined in [8]. Last but not least, as already mentioned in section 2 the treatment of (1) in [8] is exclusively occupied with optimal solutions which already belong to $\mathbb{R}^{\ell}_{++}$. Therefore, it is unable to treat situations with corner point solutions, such as instances (iii)-(iv). Hence its Exercise 1.20 about deriving the Marshallian demand for the Cobb-Douglas instance seems slightly incomplete.

Example 5. (Luenberger [9]) On p. 130 of *Microeconomic Theory* [9] the standard choice of $\mathbb{R}^{\ell}_{+}$ is made for the commodity space. The first-order conditions on p. 131 and p. 133 are correctly stated for interior points of $\mathbb{R}^{\ell}_{+}$. However, at the bottom of p. 133, still in the same section and with no mention of any change of the model, the following is stated: *"We assumed above that the solution occurred at a point* **x** *with* **x** $> \mathbf{0}$ ... *If the solution occurs at a point with* $x_1 > 0$ *and* $x_2 = 0$, *the first-order condition is that* $MRS_{21} > p_1/p_2$." Here the notation refers to marginal rate of substitution, a notion which contains partial derivatives that would have to be meaningful on the boundary but need not be so for the Cobb-Douglas family. Yet Example 5.1 on p. 132 discusses application of the same first-order conditions to the Cobb-Douglas utility function.

Example 6. (Varian [14]) *Microeconomic Analysis* [14] contains no first-order optimality conditions that are especially made to suit the general utility maximization problem. Rather, for each separate application the author refers to the Kuhn-Tucker theorem on p. 503. However, this result is only formulated for functions with all of $\mathbb{R}^{\ell}$ as their domain of definition. Hence, applications to the Cobb-Douglas instance (p. 111) and the CES instance (p. 55) are out of order from a formal viewpoint (on p. 112 utility maximization for the CES family is treated by means of earlier results for the producer that are handled by duality). Of course, the use of the logarithmic transformation to the Cobb-Douglas function on p. 111 could be exploited very easily to rectify the application of the Kuhn-Tucker theorem, because the transformed optimization problem takes place on $\mathbb{R}^2_{++}$, which is an open set (see Remark 2). In contrast to all other references under review, [14, p. 57] contains a rather complete derivation (informal, because it is based only on applying first-order necessary conditions) of a corner point solution. This was already observed in section 3.

Example 7. (Henderson-Quandt [7]) *Microeconomic Theory: a Mathematical Approach* [7] offers more evidence of the aforementioned confusion. The conditions for the utility function, stated in the lines following (2-1) on p. 8, fail to include

a proper specification of the domains of definition, differentiability and continuity. General existence of an optimal solution seems to have been ignored. Worse, because the conditions on p. 8 also require the partial derivatives of the utility function to be *strictly* positive, the applications to the Cobb-Douglas utility function in [7, pp. 18-19] can only be explained if the main results of sections 2-2 and 2-6 were meant to deal with $\mathbb{R}^{\ell}_{++}$ as the domain of the utility function. However, that clashes with the corner solution depicted in figure 2-4(b) and the nonnegativity formulas that precede it. The lines following (2-1) also require the utility function to be regular strictly quasiconcave, which is in terms of second-order partial derivatives (see footnote 1 on p. 8); thus, some mention can be made in [7] about sufficiency in the sense of part (c) of Theorem 1. However, a coherent and complete method of utility maximization is not presented and, in fact, because of the deficiencies observed above, already section 2-2 (which is for $\ell = 2$) is unable to deal with any of the five instances (i) to (v) in a mathematically correct way.

References

1. Aliprantis, C.D.: Problems in Equilibrium Theory. Springer, New York (1996)
2. Aliprantis, C.D., Brown, D.J., Burkinshaw, O.: Existence and Optimality of Competitive Equilibria. Springer, New York (1989)
3. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming (2nd ed.). Wiley, New York (1993)
4. Besanko, D., Braeutigam, D.R.: Microeconomics (2nd ed.). Wiley, New York (2005)
5. Crouzeix, J.-P.: Conditions for convexity of quasiconvex functions. Math. Oper. Res. **5**, 120-125 (1980)
6. Edwards, C.H.: Advanced Calculus of Several Variables. Academic Press, New York (1973)
7. Henderson, J.M., Quandt, R.E.: Microeconomic Theory: a Mathematical Approach (3rd ed.). McGraw-Hill, New York (1980)
8. Jehle, G.A., Reny, P.J.: Advanced Microeconomic Theory (2nd ed.). Wiley, New York (2001)
9. Luenberger, D.G.: Microeconomic Theory. McGraw-Hill, New York (1995)
10. Mas-Colell, A., Whinston, M.D., Green, J.R.: Microeconomic Theory. Oxford University Press, New York (1995)
11. Simon, C.P., Blume, L.: Mathematics for Economists. Norton, New York (1993)
12. Sydsæter, K., Hammond, P., Seierstad, A., Strøm, A.: Further Mathematics for Economic Analysis. Prentice Hall, Harlow, England (2005)
13. Takayama, A.: Mathematical Economics (2nd ed.). Cambridge University Press, Cambridge (1996)
14. Varian, H.R.: Microeconomic Analysis (3rd ed.). Norton, New York (1992)