On Economic Efficiency under Non-Convexity

by

Jean-Paul Chavas

and

Walter Briec *

September 9, 2009

Abstract: This paper investigates economic efficiency under non-convexity. The analysis relies on a generalization of the separating hyperplane theorem under non-convexity. The concept of zero-maximality is used to characterize Pareto efficiency under non-convexity. We show the existence of a separating hypersurface that can be used to provide a dual characterization of efficient allocations. When the separating hypersurface is non-linear, this implies that non-linear pricing is an integral part of economic efficiency. Implications for the decentralization of economic decisions under non-convexity are discussed.

Keywords: efficiency, zero-maximality, non-convexity, nonlinear pricing.

JEL: C02, D5, D6.

*Respectively, Professor, University of Wisconsin-Madison, Taylor Hall, Madison, WI 53706, USA, Email: jchavas@wisc.edu; and Professor, University of Perpignan, 52 avenue de Villeneuve, Perpignan 66860, France, Email: briec@univ-perp.fr

On Economic Efficiency under Non-Convexity

1 Introduction

The analysis of economic efficiency relies extensively on convexity assumptions: convexity of production sets and convexity of preferences (e.g, Debreu [10]). Yet, such assumptions are often not satisfied. Indeed, non-convexity can arise under a number of circumstances, including externalities, increasing returns, and non-divisibility (e.g., Guesnerie [11], Radner [24]). Local characterizations of Pareto efficiency under non-convexity have been explored by Bonnisseau and Cornet [6], Jofre and Cayupi [12], and Mordukhovich [21]. Relying on subgradient, they extended the Second Welfare theorem by investigating the first-order necessary conditions and the associated local prices supporting a Pareto efficient allocation under non-convexity. However, a global analysis of pricing under non-convexity is more challenging. Economists have come to rely on the separating hyperplane theorem to identify competitive prices supporting an efficient allocation (e.g., [10]). But non-convexity means that the separating hyperplane theorem can no longer be used. Moreover, non-convexity means that, in general, knowing local prices (as analyzed in [6], [11], [12], [21]) is no longer sufficient to provide a global characterization of Pareto efficiency. This is an important issue since non-convexity is known to have adverse effects on the ability of competitive markets and decentralized decisions to support an efficient allocation [11]. This has stimulated much research on nonlinear pricing schemes and their effects on efficiency. As stressed by Wilson ([30]), nonlinear pricing is commonly used in both regulated and unregulated industries. Starting with Ramsey pricing ([26]), the efficiency aspects of nonlinear pricing have been explored in the context of monopoly pricing (e.g., Boiteux [5] , Mirrlees [19]) and regulatory policies (e.g., Laffont and Tirole [13]). The analysis has also covered the efficiency of nonlinear tariffs (e.g., Ordover and Panzar [22], Amstrong and Vickers [2], Oren et al. [23]) and of bundle pricing (e.g., McAfee et al. [18], Armstrong [1]).

The objective of this paper is to develop a refined analysis of efficient allocation under non-convexity. Our approach relies on a generalized separation theorem under non-convexity. First, following Luenberger ([17]), we start with zero-maximal allocations as a representation of Pareto efficiency. Second, we use zero-maximal allocations to show the existence of a separating hypersurface that supports a dual characterization of Pareto efficiency under non-convexity. This extends the well-known separating hyperplane theorem (which applies under convexity). We show how replacing the separating hyperplane by a separating (non-linear) hypersurface provides the required analytical insights to analyze economic efficiency under non-convexity. Our main result is the establishment of a dual characterization of zero-maximality that remains valid under non-convexity. In this context, we show how the separating hypersurface provides information on nonlinear prices supporting efficient allocations. Implications for the decentralization of economic decisions under non-convexity are discussed.

2 The Model

2.1 Concepts and Notations

Consider an economy involving n individuals consuming m commodities. Let $[n] = \{1, ..., n\}$ denote the set of individuals. The i-th individual consumes $x_i \in \mathcal{X}_i \subset \mathbb{R}^m$, where $\mathcal{X}_i$ is the feasible set for x_i, $i \in [n]$. For each $i \in [n]$, we make the following assumptions:

A.1 *$\mathcal{X}_i$ is closed.*
A.2 *$\mathcal{X}_i$ has a lower bound.*

While these assumptions include $\mathcal{X}_i = \mathbb{R}^m_+$ as a special case, note that assumptions A.1 and A.2 do not require the set $\mathcal{X}_i$ to be convex.

Let $\Pi = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n$ and $\Sigma = \mathcal{X}_1 + \mathcal{X}_2 + \cdots + \mathcal{X}_n$. The vector of all consumptions is denoted by $X = (x_1, x_2, \ldots, x_n) \in \Pi \subset \mathbb{R}^{nm}$. The goods $X \in \Pi$ are produced. Denote the vector of aggregate netputs produced by $y \in \mathcal{Y} \subset \mathbb{R}^m$,[1] where $\mathcal{Y}$ is the feasible set representing the aggregate technology. We make the following assumptions:

B.1 *$\mathcal{Y}$ is closed.*
B.2 $\mathcal{Y} \cap \Sigma \neq \varnothing$.

Importantly, we do not assume that the set $\mathcal{Y}$ is convex nor that it exhibits free disposal.[2] This allows for a technology exhibiting increasing returns to scale and/or non-divisibility.

An allocation (X, y) involves consumption $X = (x_1, x_2, ..., x_n) \in \mathbb{R}^{mn}$ along with production $y \in \mathbb{R}^m$. It is feasible if it satisfies $X \in \Pi$, $y \in \mathcal{Y}$, and

$$\sum_{i \in [n]} x_i \leq y. \tag{2.1}$$

Note that assumption B.2 guarantees that a feasible allocation always exists.

We assume that each consumer has a preference ordering that can be represented by a utility function. The utility function of the i-th individual is $u_i : \mathcal{X}_i \longrightarrow \mathbb{R}$, $i \in [n]$. We make the following assumption:

[1] In the netput notation, outputs are positive and inputs are negative.
[2] $\mathcal{Y}$ would exhibit free disposal if it satisfied $\mathcal{Y} = \mathcal{Y} - \mathbb{R}^m_+$.

C.1 The utility function u_i is continuous on $\mathcal{X}_i$ for each $i \in [n]$.

Note that we do not assume that u_i is quasi-concave on $\mathcal{X}_i$, $i \in [n]$. Under a state-contingent approach, this allows for risk loving behavior. For the i-th consumer, define the range of attainable utility $\mathcal{U}_i = \{u_i(x_i) : x_i \in \mathcal{X}_i\} \subset \mathbb{R}$, $i \in [n]$ and $\mathcal{U} = \prod_{i\in[n]} \mathcal{U}_i$.

Our evaluation of consumer preferences will rely on the benefit function introduced by Luenberger ([14]). Consider a reference bundle $g \in \mathbb{R}^m_+$ with $g \neq 0$. For the i-th consumer, the benefit function is defined as follows:

Definition 2.1.1 *The i-th consumer's benefit function b_i is defined on $\mathcal{X}_i \times \mathbb{R}$ by:*

$$b_i(x_i, U_i) = \begin{cases} \sup_\beta \{\beta : \beta \in \Delta_i(x_i, U_i)\} & \text{if } \Delta_i(x_i, U_i) \neq \varnothing, \\ -\infty & \text{otherwise,} \end{cases} \tag{2.2}$$

where $\Delta_i(x_i, U_i) = \{\beta \in \mathbb{R} : u_i(x_i - \beta g) \geq U_i, x_i - \beta g \in \mathcal{X}_i\}$.

The benefit function defined in (2.2) measures the number of units of the reference bundle g the i-th consumer is willing to give up starting from utility level U_i to obtain x_i.

The properties of the benefit function have been investigated by Luenberger ([14]). We say that the reference bundle g is good for the i-th consumer if, for any $x_i \in \mathcal{X}_i$, we have $x_i + \alpha g \in \mathcal{X}_i$ and $u_i(x_i + \alpha g) > u_i(x_i)$ for all $\alpha > 0$. As shown in ([14]), for $i \in [n]$, the benefit function $b_i(x_i, U_i)$ has the following properties:

(a) $b_i(x_i, U_i)$ is upper semi-continuous in x_i and U_i, and non-increasing in U_i.

(b) $b_i(x_i, U_i)$ is non-decreasing in x_i if $u_i(x_i)$ is non-decreasing in x_i.[3]

(c) When g is good for the i-th consumer, $u_i(x_i) = U_i$ implies that $b_i(x_i, U_i) = 0$.

(d) When x_i is in the interior of $\mathcal{X}_i$, $b_i(x_i, U_i) = 0$ implies that $u_i(x_i) = U_i$.

(e) The benefit function satisfies the translation property $b_i(x_i + \alpha g, U_i) = \alpha + b_i(x_i, U_i)$ for any $\alpha \in \mathbb{R}$.

In general, note that $\{x_i : u_i(x_i) \geq U_i, x_i \in \mathcal{X}_i\} \subset \{x_i : b_i(x_i, U_i) \geq 0, x_i \in \mathcal{X}_i\}$. However, combining (*c*) and (*d*), it follows that $\{x_i : u_i(x_i) \geq U_i, x_i \in \mathcal{X}_i\} = \{x_i : b_i(x_i, U_i) \geq 0, x_i \in \mathcal{X}_i\}$ when g is good for the i-th consumer and x_i is in the interior of $\mathcal{X}_i$.

Next, following [17], define the shortage function

[3] As shown in ([14]), $b_i(x_i, U_i)$ would be concave in x_i if $u_i(x_i)$ is quasi-concave on $\mathcal{X}_i$ and $\mathcal{X}_i$ is a convex set. However, given our focus on non-convexity, we do not assume that $u_i(x_i)$ is quasi-concave on $\mathcal{X}_i$, or that $\mathcal{X}_i$ is a convex set. In other words, our analysis applies in situations where the benefit function $b_i(x_i, U_i)$ is not concave in x_i.

Definition 2.1.2 *The shortage function is defined on $\mathbb{R}^m$ by*

$$S(y) = \begin{cases} \inf_{\sigma}\{\sigma : y - \sigma g \in \mathcal{Y}\} \text{ if there is a } \sigma \text{ such that } y - \sigma g \in \mathcal{Y}, \\ +\infty \text{ otherwise.} \end{cases} \tag{2.3}$$

The shortage function in (2.3) measures how far short (measured in number of units of the reference bundle g) is point y from the frontier of the aggregate technology (represented by the feasible set $\mathcal{Y}$). In general, note that $S(y) \leq 0$ when $y \in \mathcal{Y}$, implying that $\mathcal{Y} \subset \{y : S(y) \leq 0\}$. In addition, the shortage function satisfies the translation property $S(y + \alpha g) = \alpha + S(y)$ for any $\alpha \in \mathbb{R}$.[4] Below, we will make extensive use of the benefit function (2.2) and of the shortage function (2.3) in the analysis of efficiency.

2.2 Efficiency

In our analysis of economic efficiency, we rely on the classical Pareto criterion: a feasible allocation is Pareto efficient if no individual can be made better off without making anyone else worse off. For some $U \equiv (U_1, U_2, ..., U_n) \in \mathcal{U}$, our analysis will rely on the following optimization problem[5]

$$V(U) = \sup_{\substack{X \in \Pi \\ y \in \mathcal{Y}}} \Big\{ \sum_{i \in [n]} b_i(x_i, U_i) : \sum_{i \in [n]} x_i \leq y \Big\}. \tag{2.4}$$

Below, we focus our attention on situations where the supremum in (2.4) is a maximum.[6] Then, following [17], an allocation (X, y) is said to be maximal if it solves (2.4). And it is zero-maximal if, in addition to being maximal, U is chosen to satisfy $V(U) = 0$.

Next, as in [15] and [17], we relate zero-maximality to Pareto efficiency. In what follows, for a given $X^* = (x_1^*, x_2^*, ..., x_n^*)$, we adopt the notation $U_i^* = u_i(x_i^*)$ for $i \in [n]$, and $U^* = (U_1^*, U_2^*, ..., U_n^*)$. Luenberger ([17], p. 190) showed the following proposition.

Proposition 2.2.1 *Assume that the reference bundle g is good for at least one consumer. If a feasible allocation $(X^*, y^*) \in \Pi \times \mathcal{Y}$ is Pareto efficient, then it is zero maximal.*

Also, Luenberger ([15], p. 231) proved the following result.

[4]Note that $S(y)$ would be convex in y if $\mathcal{Y}$ is a convex set ([17]). However, given our focus on non-convexity, we do not assume that $\mathcal{Y}$ is a convex set. In other words, our analysis applies to situations where the shortage function $S(y)$ is not convex.

[5]Under assumption B.2, the optimization problem (2.4) has necessarily a solution for some U.

[6]Under assumption C.1 and from Weierstrass' theorem, a sufficient (but not necessary) condition for (2.4) to correspond to a maximum is that the set $\Sigma \cap \mathcal{Y}$ be compact.

Proposition 2.2.2 *If an allocation* $(X^*, y^*) \in \Pi \times \mathcal{Y}$ *is zero-maximal, then it is Pareto efficient compared to all allocations where* x_i *is in the interior of* $\mathcal{X}_i$ *for each* $i \in [n]$.

Propositions 2.2.1 and 2.2.2 establish the relationship (under some regularity conditions) between zero-maximality and Pareto efficiency.[7] This relationship is intuitive. First, it associates Pareto efficiency with the maximization of aggregate benefit. Second, interpreting $V(U)$ as aggregate surplus, it implies the complete redistribution of social surplus. Finally, the utilities U satisfying $U \in \{U' : V(U') = 0\}$ characterize the Pareto utility frontier (Samuelson [28]). Importantly, these results hold without assuming that the production set $\mathcal{Y}$ is convex or that the utility functions $u_i(x_i), i \in [n]$, are quasi-concave.

The following alternative characterization of maximality will prove useful. (All proofs are presented in the Appendix).

Lemma 2.2.3 *Let* V *be the map defined in (2.4). We have*[8]

$$V(U) = \sup_{\substack{X \in \Pi \\ y \in \mathbb{R}^m}} \Big\{ \sum_{i \in [n]} b_i(x_i, U_i) - S(y) : \sum_{i \in [n]} x_i \leq y \Big\}. \tag{2.5}$$

Lemma 2.2.3 shows that a maximal allocation can be obtained by maximizing net benefit, i.e. aggregate benefit minus shortage. Note that the optimization with respect to production activities y is unrestricted in (2.5). This means that the shortage function $S(y)$ in (2.5) automatically imposes the feasibility condition $y \in \mathcal{Y}$.[9] In other words, in maximal allocations, the shortage function $S(y)$ captures all the relevant information about technology with or without a convex technology. From propositions 2.2.1 and 2.2.2 (and under the regularity conditions stated in these propositions), it follows that the optimization problem (2.5) defines a Pareto efficient allocation when $U \in \{U' : V(U') = 0)\}$. Below, we will make extensive use of this result.

3 Separation

3.1 Existence of a separating hypersurface

Let $(X^*, y^*) \in \Pi \times \mathcal{Y}$ be a zero-maximal allocation. From equation (2.5), it follows that

$$\sum_{i \in [n]} b_i(x_i, U_i^*) - S(y) \leq \sum_{i \in [n]} b_i(x_i^*, U_i^*) - S(y^*) = 0 \tag{3.1}$$

[7] Note that Proposition 2.2.2 can be shown to hold under slightly more general conditions. Indeed, Luenberger's proof ([15], p. 231) remains valid when compared to all allocations $x_i \in \mathcal{X}_i$ such that $x_i - \beta_i g \in \mathcal{X}_i$ for some $\beta_i > 0, i \in [n]$.

[8] Note that the supremum in (2.5) can be replaced by a maximum when (2.4) corresponds to maximum.

[9] In addition, it implies that a maximal allocation always satisfies $S(y) = 0$.

for all $y \in \mathbb{R}^m$ and all $X \in \Pi$. Define on $\mathbb{R}^m \times \mathcal{U}$ the function

$$B(y,U) = \sup_X \Big\{ \sum_{i\in[n]} b_i(x_i,u_i) : \sum_{i\in[n]} x_i \leq y, X \in \Pi \Big\}. \tag{3.2}$$

Noting that $b_i(x_i,U_i)$ defined in (2.2) is upper semi-continuous in (x_i,u_i), it follows that the function $B(y,U)$ in (3.2) is upper- semi-continuous in (y,U). In addition, considering maximal allocations, we deduce from (2.4) and (2.5) that

$$V(U) = \sup_y \{B(y,U) : y \in \mathcal{Y}\} = \sup_y \{B(y,U) - S(y) : y \in \mathbb{R}^m\} \tag{3.3}$$

which has for solution

$$Y^*(U) = \arg\max_y \{B(y,U) - S(y) : y \in \mathbb{R}^m\}, \tag{3.4}$$

for some $U \in \mathcal{U}$. Under the conditions stated in propositions 2.2.1 and 2.2.2, it follows that Pareto efficiency is obtained when U satisfies $V(U) = B(y,U) - S(U) = 0$ and $y \in Y^*(U)$.

Below, we consider the case where U satisfies $Y^*(U) \neq \varnothing$ and $\mathbb{R}^m \backslash Y^*(U) \neq \varnothing$. We also consider the map $h^* : \mathbb{R}^m \longrightarrow \mathbb{R}$ defined as

$$h^*(y) = \theta^*[B(y,U^*) - S(y)] + S(y), \tag{3.5}$$

for some $\theta^* \in]0,1[$. The function h^* in (3.5) is used to establish the following proposition.

Proposition 3.1.1 *(Separation theorem). Let $(X^*, y^*) \in \Pi \times \mathcal{Y}$ be a zero-maximal allocation and let $U^* = (u_1(x_1^*), ..., u_n(x_n^*))$. Then, there exists a function $h^* : \mathbb{R}^m \to \mathbb{R}$ which satisfies*

$$h^*(y) = B(y,U^*) = S(y) = 0, \tag{3.6}$$

for all $y \in Y^*(U^*)$ and

$$B(y,U^*) < h^*(y) < S(y) \tag{3.7}$$

for all $y \in \mathbb{R}^m \backslash Y^*(U^*)$. Moreover, $V(U^*) = 0$.

The function $h^* : \mathbb{R}^m \to \mathbb{R}$ defines a hypersurface $\{y : h^*(y) = 0, y \in \mathbb{R}^m\}$ that separates the sets $\{y : B(y,U^*) \geq 0, y \in \mathbb{R}^m\}$ and $\{y : S(y) \leq 0, y \in \mathbb{R}^m\}$. Noting that $\mathcal{Y} \subset \{y : S(y) \leq 0\}$ and $\{y : u_i(x_i) \geq U_i, x_i \in \mathcal{X}_i,\ i \in [n]; \sum_{i\in[n]} x_i \leq y,\ y \in \mathbb{R}^m\} \subset \{y : \sum_{i\in[n]} b_i(x_i,u_i) \geq 0,\ x_i \in \mathcal{X}_i, i \in [n]; \sum_{i\in[n]} x_i \leq y,\ y \in \mathbb{R}^m\} \subset \{y : B(y,U) \geq 0, y \in \mathbb{R}^m\}$, it follows that $\{y : h^*(y) = 0, y \in \mathbb{R}^m\}$ is also a separating hypersurface between $\mathcal{Y}$ and $\{y : u_i(x_i) \geq u_i^*, x_i \in \mathcal{X}_i, i \in [n];\ \sum_{i\in N} x_i \leq y,\ y \in \mathbb{R}^m\}$.

The separating hypersurface $\{y : h^*(y) = 0, y \in \mathbb{R}^m\}$ is in general not unique. To see that, simply consider choosing a different $\theta \in]0, 1[$ in (3.5), which would generate a different function h^*. This issue of choosing among the possible separating hypersurfaces will be investigated below. In general, we take h^* to be an upper semi-continuous function (e.g., by letting $\theta^* \to 1$ in (3.5)). Below, we will explore conditions under which the function h^* can be chosen to be continuous.

Note that the analysis becomes much simpler under convexity assumptions. Then, the separating hyperplane theorem applies, and h^* can be chosen to be linear. It this case, $\{y : h^*(y) = 0, y \in \mathbb{R}^m\}$ is a separating hyperplane and its gradient identifies competitive prices supporting an efficient allocation (e.g., [10]). However, non-convexity means that a separating hyperplane may no longer exist. Yet, the above result shows that a separating hypersurface continues to exist in general under non-convexity. The key to its existence relies on allowing the function h^* to be non-linear. As discussed below, the separating hypersurface $\{y : h^*(y) = 0, y \in \mathbb{R}^m\}$ will be of considerable interest in the characterization of efficient allocations under non-convexity. We will argue that $h^*(y)$ can be interpreted as the aggregate value of y, its gradient defining social prices supporting a Pareto efficient allocation.

3.2 Nonlinear price equilibrium

Interpreting $h^*(y)$ as the aggregate value of y, we now consider a class of allocations defined by the concept of a nonlinear price equilibrium. It includes as a special case the standard characterization of competitive equilibrium obtained when prices are uniform (e.g., [10], [17]).

Definition 3.2.1 *A triplet (X^*, y^*, h^*) consisting of an allocation $X^* = (x_1^*, ..., x_n^*)$, an aggregate production plan y^* and a real valued function h^* defined on $\mathbb{R}^m$, is called a nonlinear price equilibrium of the economy if:*

(a) $\sum_{i \in [n]} x_i^* \in \mathcal{Y}$ (feasibility).

(b) If $x_i \in \mathcal{X}_i$ and $u_i(x_i) > u_i(x_i^*)$ for some $i \in [n]$, then $h^*(x_i + y^* - x_i^*) > h^*(y^*)$ (preference maximization).

(c) $h^*(y^*) \geq h^*(y)$ for all $y \in \mathcal{Y}$, (aggregate profit maximization).

In the case of a convex economy (where the separating hyperplane theorem applies), note that we can set $h^*(y) = p^* \cdot y$, where p^* can be interpreted as the price vector for the m commodities. Then, our non-linear price equilibrium reduces to the classical market equilibrium. Indeed, in this case, condition (b) means that if $x_i \in \mathcal{X}_i$ and $u_i(x_i) > u_i(x_i^*)$, then $p \cdot x_i > p \cdot x_i^*$ for $i \in [n]$. Similarly, condition (c) becomes $p^* \cdot y^* \geq p^* \cdot y$ for all $y \in \mathcal{Y}$.

The following two propositions establish the close linkages existing between a nonlinear price equilibrium and zero-maximal allocations.

Proposition 3.2.2 *Assume that g is good for at least one consumer, and that all utility functions satisfy local non-satiation. Suppose there is some $\theta^* \in]0,1[$ such that the triple (X^*, y^*, h^*) is a nonlinear price equilibrium with $h^*(y) = \theta^*[B(y, U^*) - S(y)] + S(y)$ for all $y \in \mathbb{R}^m$. Then, (X^*, y^*) is zero maximal.*

Proposition 3.2.3 *Suppose all utility functions are continuous. Suppose that (X^*, y^*) is zero-maximal and that g is good for at least one consumer. Then, there is a map $h^* : \mathbb{R}^m \to \mathbb{R}$ such that:*

(a) For each $i \in [n]$, $x_i \in \mathcal{X}_i$ and $u_i(x_i) \geq U_i^*$ implies $h^*(x_i + y^* - x_i^*) \geq h^*(y^*)$.

(b) For all $y \in \mathcal{Y}$, $h^*(y) \leq h^*(y^*)$.

(c) If h^* is continuous[10] *and satisfies $h^*(y^*) \neq \min\{h^*(x_i + y^* - x_i^*) : x_i \in \mathcal{X}_i\}$ for all $i \in [n]$, then (X^*, y^*, h^*) is a nonlinear price equilibrium.*

Propositions 3.2.2 and 3.2.3 show the relationships between zero-maximality and a non-linear price equilibrium. Using propositions 2.2.1 and 2.2.2 (and under the conditions stated in these propositions), this also establishes relationships between non-linear price equilibrium and Pareto efficiency. Such results apply under a non-convex $\mathcal{Y}$ and non-quasi-concave preferences. They include as a special case the classical welfare theorems. Indeed, under convexity assumptions (where the separating hyperplane theorem applies), they reduce to the standard relationships between competitive equilibrium and Pareto efficiency (e.g., [10], [17]).

One can show that the convex case can be retrieved as a special case. In particular, if h^* is differentiable the gradient calculated at the equilibrium yields a price equilibrium.

Corollary 3.2.4 *Suppose that $\mathcal{Y}$ and $\mathcal{X}_i$ are convex for all $i \in [n]$. Moreover, suppose also that the utility functions are continuous and quasi-concave. Suppose that (X^*, y^*) is zero-maximal and that g is good for at least one consumer. Suppose that map h^* is continuous and differentiable at y^*. We have:*

(a) For each $i \in [n]$, $x_i \in \mathcal{X}_i$ and $u_i(x_i) \geq U_i^*$ implies $\nabla h^*(y^*) \cdot x_i \geq \nabla h^*(y^*) \cdot x_i^*$.

(b) For all $y \in \mathcal{Y}$, $\nabla h^*(y^*) \cdot y \leq \nabla h^*(y^*) \cdot y^*$.

(c) If h^* is continuous and satisfies $h^*(y^*) \neq \min\{h^*(x_i + y^* - x_i^*) : x_i \in \mathcal{X}_i\}$ for all $i \in [n]$, then $(X^*, y^*, \nabla h^*(y^*))$ is a price equilibrium.

[10] Conditions under which h^* is continuous are analyzed below.

4 Duality

In this section, we investigate a dual representation of zero-maximal allocations. Our duality relationships are obtained under the following additional assumption on consumer preferences:

D.1 (monotonicity). For each $i \in [n]$, u_i is non-decreasing on $\mathcal{X}_i$.

As noted above, the monotonicity assumption D.1 implies that $b_i(x_i, U_i)$ is non-decreasing in x_i. We assume that this condition holds through the rest of the paper.

Again, our analysis is presented without assuming that the utility function $u_i(x_i)$ is quasi-concave on $\mathcal{X}_i$, $i \in [n]$. And importantly, we do not assume that the set $\mathcal{Y}$ is convex.

4.1 Continuity

Above, we have seen that the benefit function b_i is in general upper semi-continuous. We have suggested that the function h^* defining the separating hypersurface can also be chosen to be upper semi-continuous. This section establishes conditions under which the stronger property of continuity applies to the benefit function b_i, the shortage function S as well as the function h^*. Such properties will prove important in our duality analysis.

First, consider the benefit function $b_i(x_i, U_i)$. We want to explore conditions under which b_i is continuous in x_i and $U_i, i \in [n]$. Note that Luenberger ([14]) explored this issue after making convexity assumptions. While establishing conditions for the continuity of the benefit function, some of Luenberger's results ([14]) do not hold under non-convexity.[11] This means that establishing continuity under non-convexity requires additional assumptions.

We introduce the following definition. For each $i = [n]$, we say that the direction of g is interior to the consumption set $\mathcal{X}_i$ if for all $x_i \in \mathcal{X}_i$ and all $\beta_i > 0, x_i + \beta_i g$ lies in the interior of $\mathcal{X}_i$. As shown below, this condition is needed to establish the lower semi-continuity of the benefit function. Note that this condition does not appear to be overly restrictive. For example, when $\mathcal{X}_i$ satisfies $\mathcal{X}_i = \mathcal{X}_i + \mathbb{R}^m_+$, then the direction of g is always interior to $\mathcal{X}_i$ if $g > 0$.[12]

The next result establishes conditions under which the benefit function is continuous.

Proposition 4.1.1 *Assume that assumptions A.1, A.2 and C.1 hold, and that g is good. For each $i \in [n]$, the benefit function $b_i(x_i, U_i)$ is jointly*

[11]For example, it can be shown that the benefit function is not always lower semi-continuous under non-convexity of the consumption set.

[12]When some elements of g are null, then the direction g being interior to the consumption set does restrict g to point away from the lower bound of $\mathcal{X}_i$.

continuous with respect to x_i and U_i in the interior of the region of $\mathcal{X}_i \times \mathcal{U}_i$ where $b_i(x_i, U_i)$ is finite if either one of the following two conditions holds:
(a) the set $\mathcal{X}_i$ is convex,
(b) the direction of g is interior to $\mathcal{X}_i$.

Proposition 4.1.1 establishes conditions under which the benefit function $b_i(x_i, U_i)$ is continuous in (x_i, U_i). These conditions are rather mild. First, they require that the reference bundle g be good. Second, they focus the analysis on points away from the region where the benefit function is infinite. This does not seem overly restrictive. Third, they include either the convexity of the set $\mathcal{X}_i$ (condition (a)), or that the direction of g be interior to the consumption set $\mathcal{X}_i$ (condition (b)). The former imposes a convexity restriction on the set $\mathcal{X}_i$, while the latter does not. However, condition (b) does impose some restrictions on the reference bundle g. As noted above, condition (b) is always satisfied if g is strictly positive and $\mathcal{X}_i = \mathcal{X}_i + \mathbb{R}^m_+$. And it remains satisfied when the bundle g points away from the lower bound of the set $\mathcal{X}_i$. This provides useful information that can help in the choice of g.

Below, we assume that the conditions stated in Proposition 4.1.1 are satisfied. This means that, under the stated regularity conditions, we take the benefit function to be continuous in x_i. Together with condition D.1, this means that we assume that the benefit function $b_i(x_i, U_i)$ is continuous in (x_i, U_i) and non-decreasing in x_i. These properties will be used in our duality analysis.

Can we also establish monotonicity and continuity properties of the shortage function $S(y)$? As noted above, we consider the general case where we do not restrict the set $\mathcal{Y}$ to be convex or to satisfy free disposal. In this context, the shortage function S may not be monotone nor continuous. However, define $\mathcal{Y}^+ \equiv \mathcal{Y} - \mathbb{R}^m_+$ as the free-disposal hull of $\mathcal{Y}$, and let S^+ be the shortage function obtained under $\mathcal{Y}^+$. Under free disposal, note that $S^+(y)$ is non-decreasing in y and satisfies $\mathcal{Y} \subset \mathcal{Y}^+ = \{y : S^+(y) \leq 0\}$ (see [3]; [17], p. 20). When $\mathcal{Y} \neq \mathcal{Y}^+$, it follows that the set $\mathcal{Y}^+ - \mathcal{Y}$ gives the region where free disposal does not hold. It is in that region that $S(y)$ fails to be non-decreasing in y. But under the monotonicity assumption D.1 (which implies that $B(y, U)$ is monotonic in y), the separation property stated in proposition 3.1.1 means that efficient allocations would never locate in this region. This suggests that we can ignore this region without affecting the identification of a separating hypersurface in (3.5). In other words, without a loss of generality, we can replace the function $S(y)$ by its associated free-disposal shortage function $S^+(y)$ in (3.5).[13]

[13]Note that this argument applies only to the identification of the separating hypersurface h^* (and not to the identification of maximal allocations). Thus, it does not mean that $\mathcal{Y}$ can be replaced by its associated free-disposal set $\mathcal{Y}^+$ in (2.4), or that S can be replaced by S^+ in (2.5). Indeed, doing so could incorrectly identify maximal allocations that are not feasible if they are located in $\mathcal{Y}^+ - \mathcal{Y}$.

Paralleling our earlier definition, we say that the direction of g is interior to the technology $\mathcal{Y}$ if for all $y \in \mathcal{Y}$ and all $\sigma > 0$ $y - \sigma g$ lies in the interior of $\mathcal{Y}$. As shown next, this condition plays an important role in establishing the continuity property of the shortage functions S and S^+.

Proposition 4.1.2 *Suppose that $\mathcal{Y}$ has an upper bound and that the direction of g is interior to $\mathcal{Y}$. Then,*
(a) the shortage function S is continuous on $\mathbb{R}^m$ in the interior of the region where it is finite;
(b) the free disposal shortage function S^+ is continuous on $\mathbb{R}^m$ in the interior of the region where it is finite.

Proposition 4.1.2 states conditions under which the shortage functions $S(y)$ and S^+ are continuous in y. These conditions seem rather mild. First, they restrict the feasible set to have an upper bound. Second, they focus the analysis away from regions where the shortage function is infinite. This does not appear unduly restrictive. Third, proposition 4.1.2 restricts the direction of g to be interior to $\mathcal{Y}$. This restriction is needed to establish the upper semi-continuity of the shortage functions S and S^+. For S^+, note that this condition is always satisfied when g is strictly positive. And when some elements of g are null, this restricts g to point away from the boundary of the set $\mathcal{Y}$. Again, this provides useful information that can help in the choice of g.

Below, we will assume that the conditions stated in Proposition 4.1.2 are satisfied. When applied to the free-disposal shortage function, this means that under the stated regularity conditions, S^+ is non-decreasing and continuous. Adding this to our analysis of the benefit function presented in Proposition 4.1.1 will prove useful in our duality analysis. Indeed, below, we assume that the conditions stated in propositions 4.1.1 and 4.1.2 are satisfied. This means that we proceed assuming that $B(y, U)$ and $S^+(y)$ are each non-decreasing and continuous in y.

As discussed above, after substituting $S^+(y)$ for $S(y)$, it follows from (3.5) that one can find a map h^* that is continuous and non-decreasing in y. On that basis, we now focus our attention on separating hypersurfaces characterized by continuous and non-decreasing functions h^* .

4.2 Generalized Lagrangian

Define Φ as the set of continuous and non-decreasing functions from $\mathbb{R}^m$ to $\mathbb{R}$ and satisfying the translation property: $f(y+\alpha g) = \alpha + f(y)$ for any $y \in \mathbb{R}^m$ and any $\alpha \in \mathbb{R}$. Consider the penalty functional

$$P\big(f(y), f(z)\big) = f(y) - f(z), \tag{4.1}$$

where $P : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $f \in \Phi$. The penalty functional $P(f(\cdot), f(\cdot))$ satisfies $P(f(y), f(y)) = 0$ for all $y \in \mathbb{R}^n$.

Consider the functional L defined on $\Pi \times \mathbb{R}^m \times \mathcal{U} \times \Phi$.

$$L\Big(X, y, U, f\Big) = \sum_{i\in[n]} b_i(x_i, U_i) - S(y) + P\Big(f(y), f\big(\sum_{i\in[n]} x_i\big)\Big), \tag{4.2}$$

where $X \equiv (x_1, ..., x_n) \in \Pi$, $U \equiv (U_1, ..., U_n) \in \mathcal{U}$, and $y \in \mathbb{R}^m$. Note that L in (4.2) can be interpreted as a generalized Lagrangian. Indeed, it would reduce to a standard Lagrangian when f is linear. And the translation property of the functions b_i and S and f implies that $L\Big(X, y, U, f\Big)$ is invariant to translation, with translation defined as y being replaced by $y + \sum_i \alpha_i g$, and x_i being replaced by $x_i + \alpha_i g$, $\alpha_i \in \mathbb{R}$, $i \in [n]$.

In the remainder of the paper, we focus our attention on situations where zero-maximal allocations exist. Under the regularity conditions stated in propositions 2.2.1 and 2.2.2, these allocations are Pareto efficient. We investigate the dual characterization of such allocations. While the dual characterization of Pareto efficiency is well-known under convexity (e.g., [16]), we seek to extent the dual interpretation of Pareto efficiency under non-convexity.

Consider

$$L^*(U) = \inf_f \sup_{X,y} \Big\{ L\big(X, y, U, f\big) : x \in \Pi, y \in \mathbb{R}^m, f \in \Phi \Big\}. \tag{4.3}$$

Note that the feasibility restriction $\sum_{i\in[n]} x_i \leq y$, is not imposed in (4.3).

We are interested in establishing linkages between (4.3) and the maximal allocation (2.5).

Lemma 4.2.1 *(Weak Duality) For $U \in \mathcal{U}$, let*

$$L^\#(U) = \sup_{x,y} \inf_f \Big\{ L\big(X, y, U, f\big) : X \in \Pi, y \in \mathbb{R}^m, f \in \Phi \Big\}.$$

We have,

$$L^*(U) \geq L^\#(U) \geq V(U). \tag{4.4}$$

The inequalities (4.4) show that, in general, $L^*(U)$ is an upper bound of both $L^\#(U)$ and $V(U)$. Situations where this upper bound is reached are of considerable interest (e.g., [4], [29]). The next proposition presents key results supporting a dual characterization of Pareto efficiency.

Proposition 4.2.2 *Assume that a maximal allocation exists for some $U \in \mathcal{U}$,*[14] *and that $L^*(U) = L^\#(U)$. Then, there is a saddle-point $(X^*, y^*, f^*) \in \Pi \times \mathbb{R}^m \times \Phi$ of the generalized Lagrangian $L(X, y, U, f)$ in (4.2) satisfying for all $X \in \Pi, y \in \mathbb{R}^m$ and $f \in \Phi$*

$$(a) \quad L(X, y, U, f^*) \leq L(X^*, y^*, U, f^*) \leq L(X^*, y^*, U, f), \tag{4.5}$$

[14] As noted above, under assumptions B.2 and C.1, a sufficient (but not necessary) condition for the existence of a maximal allocation is that the set $\mathcal{Y} \cap \Sigma$ be compact.

$$(b)\quad \sum_{i\in[n]} x_i^* \le y^*, \tag{4.6}$$

$$(c)\quad P\Big(f^*(y^*), f^*(\sum_{i\in[n]} x_i^*)\Big) = f^*(y^*) - f^*(\sum_{i\in[n]} x_i^*) = 0, \tag{4.7}$$

$$(d)\quad (X^*, y^*) \in \arg\max_{\substack{y\in\mathbb{R}^m\\ X\in\Pi}} \Big\{ \sum_{i\in[n]} b_i(x_i, U_i) - S(x) : \sum_{i\in[n]} x_i \le y \Big\}, \tag{4.8}$$

$$(e)\quad L^*(U) = V(U). \tag{4.9}$$

The above proposition establishes sufficient conditions for the existence of a saddle-point of the generalized Lagrangian as given in (4.5). Equation (4.6) states that (X^*, y^*) in the saddle point problem (4.5) is always at a feasible point where aggregate consumption does not exceed aggregate production: $\sum_{i\in[n]} x_i^* \le y^*$. At that point, equation (4.7) shows that the penalty function is always zero. Comparing it with (2.5), (4.8) imply that (X^*, y^*) corresponds to a maximal allocation. It means that y^* is necessarily feasible and satisfies $y^* \in \mathcal{Y}$. Finally, (4.9) states that $L^*(U) = V(U)$.

While $L^*(U) \ge V(U)$ in general (from (4.4)), equation (4.9) states that $L^*(U) = V(U)$ when a saddle-point exists. The condition $L^*(U) = V(U)$ has been called a condition of zero duality gap. The linkages between $L^*(U) = L^{\#}(U)$ and $L^*(U) = V(U)$ are presented next.

Lemma 4.2.3 *For $U \in \mathcal{U}$, we have $L^*(U) = L^{\#}(U)$ if and only if $L^*(U) = V(U)$.*

This shows that a zero duality gap, $L^*(U) = V(U)$, is equivalent to the existence of a saddle point, with $L^*(U) = L^{\#}(U)$. But under what conditions does a zero duality gap hold? To address that question, let $\gamma \in \mathbb{R}^m$, and rewrite (2.5) as

$$W(U,\gamma) \equiv \sup_{X,y} \Big\{ \sum_{i\in[n]} b_i(x_i, U_i) - S(x) : \sum_{i\in[n]} x_i \le y + \gamma, y \in \mathbb{R}^m, X \in \Pi \Big\}. \tag{4.10}$$

Comparing (2.5) and (4.10), it is clear that $W(U, 0) = V(U)$. It follows that the solution of (4.10) for (x, y) is a maximal allocation (assuming that it exists) when $\gamma = 0$. In general, $W(U,\gamma)$ is non-decreasing in γ. The next two propositions establish the linkages between a zero-duality gap and the continuity properties of $W(U,\gamma)$ with respect to γ.

Proposition 4.2.4 *Assume a zero duality gap: $L^*(U) = V(U)$. Then $W(U,\gamma)$ is upper semi-continuous in γ at $\gamma = 0$.*

Proposition 4.2.5 *Assume that $W(U,\gamma)$ is upper semi-continuous in γ at $\gamma = 0$. Then, there is a zero duality gap: $L^*(U) = V(U)$.*[15]

[15] Note that a similar result would apply to a gap between $L^*(U)$ and $L^{\#}(U)$

Propositions 4.2.4 and 4.2.5 show that a necessary and sufficient condition for a zero duality gap is that the function $W(U,\gamma)$ be upper semi-continuous in γ at $\gamma = 0$. This is the condition required to support a dual representation of Pareto efficiency, as discussed next.

4.3 Nonlinear expenditure and profit functions

Note that the choice of X in (4.3) suggests the introduction of the function E defined on $\Phi \times \mathcal{U}$ by

$$E(f,U) \equiv \inf_X \Big\{ f(\sum_{i\in[n]} x_i) - \sum_{i\in[n]} b_i(x_i, U_i) : X \in \Pi \Big\}. \tag{4.11}$$

We show next that the function $E(f,U)$ can be interpreted as an aggregate expenditure function.

Lemma 4.3.1 *For all $(f,U) \in \Phi \times \mathcal{U}$, we have*

$$E(f,U) = \inf_X \big\{ f(\sum_{i\in[n]} x_i) : u_i(x_i) \geq U_i, i \in [n], X \in \Pi \big\}. \tag{4.12}$$

Similarly, the choice of y in (4.3) implies the following maximization problem

$$\pi(f) \equiv \sup_y \{ f(y) - S(y) : y \in \mathbb{R}^m \}, \tag{4.13}$$

which defines the function π on Φ. We show next that $\pi(f)$ in (4.13) can be interpreted as an aggregate profit function.

Lemma 4.3.2 *For all $f \in \Phi$ we have*

$$\pi(f) = \sup_y \{ f(y) : y \in \mathcal{Y} \}. \tag{4.14}$$

Both $E(f,U)$ and $\pi(f)$ are conditional on the function f, which provides a measure of revenue for production activities, $f(y)$, and of expenditures for consumption activities, $f(\sum_{i\in[n]} x_i)$. Using (4.11) and (4.13) , the saddle-point problem (4.3) implies that

$$L^*(U) = \inf_f \{ \pi(f) - E(f,U) : f \in \Phi \}. \tag{4.15}$$

This shows that, when a saddle-point (4.5) exists, solving the problem (4.15) is dual to solving the primal problem (2.5).

Proposition 4.3.3 *Assume that there is a zero duality gap and that a maximal allocation exists for some $U \in \mathcal{U}$. Then, there is a dual function f^* satisfying*

$$f^* \in \arg\min_f \{ \pi(f) - E(f,U) : f \in \Phi \}. \tag{4.16}$$

In the remainder, we define

$$H(U) \equiv \arg\min_f \{\pi(f) - E(f, U) : f \in \Phi\}. \tag{4.17}$$

The above results mean that, under some regularity conditions, the function $f^* \in H(U)$ is associated with a dual characterization of maximal allocations. In general, f^* defines a continuous and non-decreasing separating hypersurface. It provides a general characterization of the separating hypersurface first considered in (3.5). In other words, h^* in (3.5) can be chosen more generally such that $h^* \in H(U^*)$, where $U_i^* = u(y_i^*)$ for $i \in [n]$. In other words, under zero duality gap, the separating hypersurface that defines the efficient aggregate revenue/cost can be obtained as the solution of $L^*(U) = \inf_f \{\pi(f) - E(f, U) : f \in \Phi\}$ in (4.15), with U being chosen such that $L^*(U) = 0$. Combining (4.15) and (4.17) with Propositions 2.2.1, 2.2.2, 4.2.2 and 4.3.3, we obtain the following result.

Proposition 4.3.4 *Assume that there is a zero duality gap. Then, a dual representation of Pareto efficiency is given by (4.12), (4.14) and (4.15), where U is chosen to satisfy $L^*(U) = 0$.*

Note that $L^*(U) = 0$ in (4.15) can be interpreted as the aggregate budget constraint stating that aggregate profit $\pi(f)$ must be entirely redistributed to consumers. This gives the following intuitive interpretation of efficiency: after maximizing aggregate profit and aggregate net benefit, and choosing f according to (4.15), Pareto efficiency is obtained by a complete redistribution of aggregate profit among the n consumers.[16]

It is important to stress that these results apply without assuming that set $\mathcal{Y}$ is convex nor that the functions u_i are quasi-concave on $\mathcal{X}_i$, $i \in [n]$. They generalize well-known results obtained under convexity. Indeed, under convexity, from the separating hyperplane theorem ([4]), f can be taken to be linear, the penalty function (4.1) becomes linear, and (4.2) becomes the standard Lagrangian. Then the gradient of f^* measures the market prices supporting an efficient allocation ([10]). However, (4.16) applies under non-convexity when f^* is possibly nonlinear. The implications of this generalization are discussed next.

5 Implications

Under non-convexity, we allowed the function $f \in \Phi$ to be non-linear. Under zero-maximality (where U is chosen such that $L^*(U) = 0$), equation (4.16) means that the function $f(x)$ can be used to define a separating hypersurface,

[16] Note that this does not say how aggregate profit gets distributed among the n consumers. While different distributions of profit will lead to a move along the Pareto utility frontier, all points along this frontier satisfy the Pareto efficiency criterion.

where f is a continuous and non-decreasing function on $\mathbb{R}^m$ and satisfies $f \in H(U^*)$.

In general, the function $f \in H(U^*)$ can take many forms. In this section, we explore alternative forms of f that can support an efficient allocation. We start with the simplest case by asking the question: can $f \in H(U^*)$ be an affine function? It can if there exists $(\alpha, p) \in \mathbb{R} \times \mathbb{R}^m_+$ satisfying $f(y) = \alpha + p \cdot y$, with

$$\alpha + p \cdot y = B(y, U^*) = S(y) \tag{5.1}$$

for all $y \in Y^*(U^*)$ and

$$B(y, U^*) < \alpha + p \cdot y < S(y) \tag{5.2}$$

for all $y \in \mathbb{R}^m \backslash Y^*(U^*)$. This includes two important special cases: the case of uniform pricing (UP) when $\alpha = 0$; and the case of two-part tariffs (TPT) when $\alpha \neq 0$. They are discussed next.

5.1 The case of uniform pricing (UP)

When $\alpha = 0$ and $f \in H(U^*)$ is defined by $f(y) = p \cdot y$, f characterizes a separating hyperplane. From the separating hyperplane theorem ([4]), it is well known that, under convexity assumptions, there always exists some $p \in \mathbb{R}^m_+$ such that f characterizes a separating hyperplane satisfying equations (5.1) and (5.2) with $\alpha = 0$. In this context, p can be interpreted as the social prices of the m commodities. And under competitive markets, they are the market-clearing prices, generating the classical results that competitive markets support an efficient allocation (e.g., [10]). This makes it clear that $f(y) = p \cdot y$ identifies scenarios of uniform pricing (UP), where the prices p are the same for all agents in the economy.

But can uniform pricing still apply in the presence of non-convexity? The answer is: yes, provided that (5.1) and (5.2) hold with $\alpha = 0$. This requires that p satisfies two conditions. First, given $\alpha = 0$, (5.1) implies that $p \cdot x$ must be tangent to both $\{y \in \mathbb{R}^n : B(y, U^*) \geq 0\}$ and $\{y \in \mathbb{R}^n : S(y) \leq 0\}$, for all $y \in Y^*(U^*)$. When the set $Y^*(U^*)$ has a single point, this means that $y \mapsto p \cdot y$, $y \mapsto B(y, U^*)$, as well as $y \mapsto S(y)$ must go through that point. And when the set $Y^*(U^*)$ has more than one point, this requires that $B(y, U^*)$ and $S(y)$ are both linear in y and with identical derivatives on $Y^*(U^*)$. Note that this rules out the presence of non-convexity within the set $Y^*(U^*)$

Second, given $\alpha = 0$, (5.2) implies that the map $y \mapsto p \cdot y$ must strictly separate $\{y \in \mathbb{R}^m : B(y, U^*) > 0\}$ and $\{y \in \mathbb{R}^n : S(y) < 0\}$. This allows for non-convexity within the set $\mathbb{R}^m \backslash Y^*(U^*)$. These non-convexities can be associated with $B(y, U^*)$ being a non-concave function of y and/or $S(y)$ being a non-convex function of y, for $y \in \mathbb{R}^m \backslash Y^*(U^*)$. However these non-convexities should remain mild so that (5.2) is not violated. This corresponds to situations where $conv\{y \in \mathbb{R}^m \backslash Y^*(U^*) : B(y, U^*) \geq 0\} \cap conv\{y \in \mathbb{R}^m \backslash Y^*(U^*) : S(y) \leq 0\} = \varnothing$, where $conv\{A\}$ denotes the convex hull of A. In this case, the non-convex maximal allocation can be 'convexified' and treated as if it were a standard convex problem.[17]

Under such conditions and given $\alpha = 0$, (5.1) and (5.2) remain satisfied even under non-convexity. It means that a separating hyperplane exists. Again, its gradient $p \in \mathbb{R}^m_+$ provides a measure of social prices supporting an efficient allocation. Importantly, with $f(y) = p \cdot y$, equation (4.11) becomes $E(p, U) = \sum_{i \in [n]} E_i(p, U_i)$, where

$$E_i(p, U_i) = \inf_{x_i} \{p \cdot x_i) - b_i(x_i, U_i) - : x_i \in \mathbb{R}^m\}, \tag{5.3}$$

$E_i(p, U_i)$ being the classical expenditure function for the i-th consumer, $i \in [n]$ (see [14]). In this context, it follows from (5.3) that $-E_i(p, u_i)$ provides a measure of the net benefit (net of cost) obtained by the i-th consumer.

Similarly, with $f(y) = p \cdot y$, equation (4.13) becomes

$$\pi(p) = \sup_y \{p \cdot y - S(y) : y \in \mathbb{R}^m\}, \tag{5.4}$$

where $\pi(p)$ is the classical aggregate profit function (see [17]). In this case, the prices p are the same for all agents and standard efficiency results apply. First, conditional on p, efficient consumption decisions can be decentralized, where each consumer behaves so as to minimize its expenditures. Second, equation (5.4) states that efficiency implies the maximization of aggregate profit, conditional on prices p. This applies as well when production activities involve a set of firms. And it applies even in the presence of production externalities among firms. As such, equation (5.4) is just a restatement of the Coase theorem (see [9])).

While the Coase theorem implies aggregate profit maximization, it does not imply firm-level profit maximization in the presence of externalities. Indeed, ruling out firm externalities is in general required for decentralized profit maximization to be consistent with economic efficiency. To see that, assume that k firms are involved in the production of y. In the absence of production externalities, the feasible set can be written as $\mathcal{Y} = \sum_{j=1}^k \mathcal{Y}_j$, where $\mathcal{Y}_j$ is the feasible set for the j-th firm. Then, equation (5.4) can be written as $\pi(p) = \sum_{j \in [k]} \pi_j(p)$, where $\pi_j(p) = \sup\{p \cdot y_j : y_j \in \mathcal{Y}_j\}$ is the indirect profit function for the j-th firm, $j = 1, ..., k$. When f is linear, this reduces to the standard result that decentralized profit maximization (conditional on prices p, outputs being chosen to satisfy the marginal cost pricing rule) is consistent with economic efficiency (see [10]). Under such a scenario, non-convexities do not invalidate standard results concerning the efficiency of decentralized decisions under competitive markets.

[17] This includes the case of pseudoconcavity/pseudoconvexity as discussed by Mangasarian [20] .

5.2 The case of two-part tariffs (TPT)

Now consider the case where $f(x) = \alpha + p \cdot x$, with $\alpha \neq 0$. Interpreting p as the marginal unit prices of the m commodities, it follows that α can be interpreted as an aggregate fixed fee that is both paid by the consumers and received by the producers. This is the basic two-part tariff (TPT) scheme, representing the simplest possible form of non-linear pricing. This raises the question : what do we gain by going beyond uniform pricing and introducing the fixed fee α?

The usefulness of TPT pricing is that the fixed fee α can help find efficient pricing scheme $f(x) = \alpha + p \cdot x$ that satisfies equations (5.1) and (5.2). Below, we consider two scenarios where $\alpha > 0$: the case of perfect price discrimination; and the case of non-convexity associated with increasing returns to scale (IRTS).

First, consider the situation where $\alpha = \sum_{i \in [n]} E_i(p, u_i)$, where $E_i(p, u_i)$ is defined in equation (5.3). With α being the aggregate fee paid by consumers and received by producers, this means that the aggregate net consumer benefits are completely extracted by the producers. This is a situation of perfect price discrimination corresponding to an efficient allocation where aggregate benefits are maximized but where such benefits also totally captured by producers (see [25]). This generates the largest possible aggregate profit. It is efficient in the sense of satisfying the Pareto efficiency criterion. But is it equitable? It depends on how the aggregate profit is redistributed among the n consumers. Note that efficient price discrimination could be desirable on equity ground if the aggregate profit is distributed in a way that generates a more equitable distribution of purchasing power among consumers. But this begs the question: why go through the trouble of extracting consumer benefits to give it back to them? This indicates that price discrimination schemes are likely to arise either if they are associated with a redistribution of purchasing power among consumers, or if they are a necessary part of implementing an efficient allocation. This later case can arise under non-convexity, as illustrated next.

Second, consider the case of increasing returns to scale (IRTS), where $kF \subset F$ for all $k > 1$. Assume that $0 \in F$. Under uniform pricing $p \in \mathbb{R}^m_+$, consider the profit function $\pi^+(p) = \sup_y \{p \cdot y : y \neq 0, y \in \mathcal{Y}\}$. It is well known that, under IRTS and uniform pricing, $\pi^+(p) < 0$. This means that uniform pricing gives no incentive to produce under IRTS. The profit-maximizing solution then is $y = 0$, which is in general inefficient. In this case, the problem is not with profit maximization, but rather with uniform pricing (UP). This is a scenario where UP cannot support an efficient allocation. But TPT pricing can. To illustrate, consider the case where $\alpha = -\pi^+(p) > 0$. Then, profit maximization under (TPT) implies that $\pi(p) = sup_y\{\alpha + p \cdot y : y \in \mathcal{Y}\} = 0$. This identifies $\alpha = -\pi^+(p) > 0$ as the smallest fixed fee paid by consumers to provide an incentive for producers to produce and support an efficient allocation.[18] Under this TPT scheme, the fixed payment α generates zero aggregate profit, meaning that there is nothing to redistribute from producers as consumers efficiently capture all the economic surplus.

These two scenarios generate efficient allocations under two extreme situations: the aggregate surplus is captured entirely by producers in the first case, but by consumers in the second case. The general TPT case includes scenarios where both producers and consumers share the aggregate surplus. In situations where $\sum_{i\in[n]} E_i(p, u_i) > -\pi^+(p)$, the fixed payment α can be chosen such that $\sum_{i\in[n]} E_i(p, u_i) > \alpha > -\pi^+(p)$. Under Pareto efficiency, the aggregate profit is $\alpha + \pi^+(p) > 0$ is then redistributed to the n consumers. Again, the nature of the redistribution affects the equity but not the efficiency of the allocation. This indicates that TPT schemes provide some flexibility in Pareto efficiency. And this illustrates that TPT cannot always reduce to UP (e.g., the case of IRTS) and still support efficient allocations. This makes it clear that Pareto efficiency should be evaluated under the broader context of non-linear pricing.

5.3 The general case of non-linear pricing (NLP)

While IRTS is a well-known example where the convexity of the production set does not hold, our analysis applies under general non-convexity. The general case of non-convexity covers situations where equations (5.1) and (5.2) are not satisfied for any $\alpha \in \mathbb{R}$ and $p \in \mathbb{R}^m_+$. Then, non-convexity means that neither UP nor TPT pricing schemes can support an efficient allocation. It means that we must now work with a separating hypersurface, where $f \in H(U^*)$ is necessarily non-linear. While $f \in H(U^*)$ remains in general continuous and non-decreasing on $\mathbb{R}^m_+$, it may not be differentiable. In this context, it can be useful to analyze its properties using subgradient (see [8], [27]). Indeed, elements of the subgradient of f provide a measure of local prices. When evaluated at a Pareto efficient point, these local prices must be tangent to both the benefit function and the shortage function. This is consistent with the analysis of first-order necessary conditions presented by Bonnisseau and Cornet [6], Jofre and Cayupi [12], and Mordukhovich [21] in their characterization of Pareto efficiency. However, when f is non-linear, these local prices vary with the evaluation point. As such, they are not globally valid. Our analysis extends previous research by showing how a non-linear $f \in H(U^*)$ provides information about both local and global prices supporting an efficient allocation under non-convexity. It provides new and useful insights about the efficiency of non-linear pricing [30].

The nonlinearity of $f \in H(U^*)$ can arise from two sources. First, f is nonlinear when equation (5.1) fails to hold. This means that $B(x, U^*) =$

[18] The fixed fee $\alpha > 0$ amounts to a subsidy to producers. This argument has been used to argue that a subsidy to infant industries facing IRTS can help support an efficient allocation. However, note that if α represents this subsidy, it amounts to an income transfer to producers and not a price subsidy.

$S(x)$ is nonlinear for $y \in Y^*(U^*)$. This is relevant only if $Y^*(U^*)$ is not a singleton. This corresponds to situations where the functions $B(y, U^*)$ and $S(y)$ are both nonlinear and tangent to each other at more than one point. Although this scenario can occur, we conjecture that it may not very common. Second, $f \in H(U^*)$ must be nonlinear when (5.2) fails to hold. This requires that $B(y, U^*)$ be non-concave in y on $\mathbb{R}^m \backslash Y^*(U^*)$ and/or that $S(y)$ be non-convex in y on $\mathbb{R}^m \backslash Y^*(U^*)$. Such non-convexities can arise in situations of indivisibility and/or of increasing returns.

The separating function $f \in H(U^*)$ being non-linear implies non-linear pricing. Note that this does not invalidate expenditure minimization or profit maximization. As in the convex case, efficient consumption decisions involve the minimization of aggregate expenditures (4.12). However, note that the presence of nonlinearity in $f \in H(U^*)$ means that the decentralization of consumer decisions is no longer consistent with Pareto efficiency. The reason is that the marginal value of consumer goods (as measured by the subgradient of $f(\sum_{i \in [n]} x_i)$) now depends on the consumption level of all consumers. In other words, under non-convexity and nonlinear pricing, the general consistency of decentralized consumption decisions with Pareto efficiency fails to hold.

And as in the convex case, efficient production decisions involve the maximization of aggregate profit (4.14). This is a generalization of the Coase theorem under non-convexity. In the presence of multiple firms, it applies whether or not production externalities exist among firms. However, $f \in H(U^*)$ being non-linear means that commodity prices are no longer constant. This means that, under multiple firms, the decentralization of production decisions is no longer consistent with Pareto efficiency. The reason is that the marginal value of aggregate production (e.g., as measured by the subgradient of f) now depends on the production level of all firms. In other words, under non-convexity and nonlinear pricing, the general consistency of decentralized production decisions with Pareto efficiency fails to hold.

These arguments suggest that developing efficient nonlinear pricing schemes can be quite challenging. Indeed, the identification of market clearing prices becomes complex and the decentralization of allocation decisions becomes problematic. However, evolutionary selection may help in the long run. To see that, consider experimentations with alternative nonlinear pricing. Such experimentations can identify which scheme generates higher aggregate surplus. To the extent that evolutionary selection favors the schemes generating larger surpluses, the pricing system may evolve toward a steady state equilibrium that is Pareto efficient (e.g., Courtault and Tallon [7]). This would occur in the long run if the decision makers in charge of the experimentation are in a position to capture the increase in aggregate surplus. Then, they would have incentives to select the pricing schemes that increase aggregate surplus. In the long run, that process would converge to an efficient allocation. In this case, while efficiency may not hold in the short run, it could arise in the long run without explicit coordination.

6 Conclusion

Our analysis has examined economic efficiency under non-convexity. It relied on a generalized separation theorem under non-convexity. Starting with zero-maximal allocations as a representation of Pareto efficiency, we showed the existence of a separating hypersurface that supports a dual characterization of Pareto efficiency under non-convexity. We showed how replacing the standard separating hyperplane (which always exists under convexity) by a separating (non-linear) hypersurface provides the required analytical insights to analyze economic efficiency under non-convexity. Our main result is the establishment of a dual characterization of zero-maximality that remains valid under non-convexity. In this context, we show how the separating hypersurface provides information about pricing supporting efficient allocations. When the separating hypersurface is non-linear, this implies that non-linear pricing is an integral part of economic efficiency. Finally, we explored the adverse effects of non-convexity on the ability of decentralized decision-making to support Pareto efficiency.

Under non-convexity and nonlinear pricing, finding that complete decentralization is no longer efficient is important. This points to the limitations of competitive markets in their ability to support economic efficiency. More generally, it stresses the importance of coordination schemes needed to implement Pareto efficiency. Such coordination schemes can in principle be provided by contracts implemented under various institutional environments. The challenge then is to identify the nature and types of centralization schemes that would implement Pareto efficient allocation. Alternatively, we argue that the prospects from uncovering efficient nonlinear pricing schemes may improve in the long run if evolutionary selection tends to favor the schemes that contribute to increasing aggregate surplus.

Appendix

Proof of lemma 2.2.3. Noting that $y \in \mathcal{Y}$ implies that $S(y) \leq 0$, we have

$$\begin{aligned} V(U) &= \sup_{X,y} \Big\{ \sum_{i \in [n]} b_i(x_i, U_i) : X \in \Pi, y \in \mathcal{Y}, \sum_{i \in [n]} x_i \leq y \Big\} \\ &\leq \sup_{X,y} \Big\{ \sum_{i \in [n]} b_i(x_i, U_i) - S(y) : X \in \Pi, y \in \mathcal{Y}, \sum_{i \in [n]} x_i \leq y \Big\} \\ &\leq \sup_{X,y} \Big\{ \sum_{i \in [n]} b_i(x_i, U_i) - S(y) : X \in \Pi, y \in \mathbb{R}^m, \sum_{i \in [n]} x_i \leq y \Big\}. \end{aligned}$$

We now need to show that $V(U) \geq \sup_{X,y}\{\sum_{i \in [n]} b_i(x_i, U_i) - S(y) : X \in \Pi, y \in \mathbb{R}^m, \sum_{i \in [n]} x_i \leq y\}$. From (2.5), this inequality clearly holds if $\sum_{i \in [n]} b_i(x_i, U_i) - S(y) = -\infty$. Consider the case where $\sum_{i \in [n]} b_i(x_i, U_i) - S(y) > -\infty$. Letting $y' = y - S(y)g$, we have $y' \in \mathcal{Y}$, and $S(y') = 0$. Let

$x'_i = x_i - \beta_i g,\ i \in [n]$, where the $\beta'_i s$ satisfy $\sum_{i\in[n]} \beta_i = S(y)$. Using the translation property of both b and S, we obtain

$$
\begin{aligned}
&\sup_{X,y} \Big\{ \sum_{i\in[n]} b_i(x_i, U_i) - S(y) : X \in \Pi, y \in \mathbb{R}^m, \sum_{i\in[n]} x_i \le y \Big\} \\
&= \sup_{X,X',y,y'} \Big\{ \sum_{i\in[n]} b_i(x'_i, U_i) - S(y') : X' \in \Pi, \sum_{i\in[n]} x'_i \le y', \\
&\qquad\qquad y' = y - S(y)g \in \mathcal{Y},\ x'_i = x_i - \beta_i g, i \in [n], \sum_{i\in[n]} \beta_i = S(y) \Big\} \\
&\le \sup_{X',y'} \Big\{ \sum_{i\in[n]} b_i(x'_i, U_i) : X' \in \Pi, y' \in \mathcal{Y}, \sum_{i\in[n]} x'_i \le y' \Big\} \\
&= V(U),
\end{aligned}
$$

which concludes the proof. ■

Proof of proposition 3.1.1. From (3.5), pick some $\theta^* \in]0, 1[$ and consider the map $h^* : \mathbb{R}^m \longrightarrow \mathbb{R}$ defined by

$$h^*(y) = \theta^*[B(y, U^*) - S(y)] + S(y).$$

If $y \in Y^*(U^*)$ then $B(y, U^*) = S(y) = 0$, hence $h^*(y) = 0$. Suppose that $y \notin Y^*(U^*)$. From (3.4), $B(y, U^*) - S(y) < 0$ for $y \in \mathbb{R}^m \backslash Y^*(U^*)$. Hence $h^*(y) = \theta^*[B(y, U^*) - S(y)] + S(y) < S(y)$. Furthermore, since $0 < \theta^* < 1$ and $[B(y, U^*) - S(y)] < 0$, we deduce that $\theta^*[B(y, U^*) - S(y)] > B(y, U^*) - S(y)$. It follows that $B(y, U^*) < \theta^*[B(y, U^*) - S(y)] + S(y) = h^*(y)$ for all $y \in \mathbb{R}^m \backslash Y^*(U^*)$, which ends the proof. ■

Proof of proposition 3.2.2. We need to prove that $B(y^*, U^*) = S(y^*) = 0$. If (X^*, y^*, h^*) is a nonlinear price equilibrium, by definition, $y^* \in \mathcal{Y}$. It follows that $S(y^*) \le 0$. Moreover, we have from Proposition 3.1.1

$$B(y^*, U^*) \le \theta^*[B(y^*, U^*) - S(y^*)] + S(y^*) \le S(y^*).$$

Consequently, we have $B(y^*, U^*) \le 0$. Given $u_i(x_i^*) = U_i^*$ for each $i \in [n]$, it follows that $\sum_{i\in[n]} b_i(x_i, U_i^*) \ge 0$. Therefore $B(y^*, U^*) \ge 0$. Consequently, we have $B(y^*, U^*) = 0$. This implies that $S(y^*) \ge 0$ from Proposition 3.1.1. Thus $S(y^*) = 0$, which ends the proof. ■

Proof of proposition 3.2.3. Let $\mathcal{L}(U^*) = \{y : y = \sum_{i\in[n]} x_i, b_i(x_i, U_i^*) \ge 0, i \in [n]\}$. From Proposition 3.1.1 and letting $\theta^* \to 1$ in (3.5), there exists a map $h^* : \mathbb{R}^m \longrightarrow \mathbb{R}$ such that

$$\mathcal{Y} \subset \{y : h^*(y) \le 0\} \quad \text{and} \quad \mathcal{L}(U^*) \subset \{y : h^*(y) \ge 0\}.$$

This immediately proves condition (b). Now, fix i and consider $x_i \in \mathcal{X}_i$. If $b_i(x_i, U_i^*) \ge 0$, we have $y_i = x_i + \sum_{j\neq i} x_j^* \in \mathcal{L}(U^*)$. Hence, $h^*(x_i + \sum_{j\neq i} x_j^*) \ge$

$h^*(y^*)$. Thus, $b_i(x_i, U_i^*) \geq 0$ implies $h^*(x_i + \sum_{j\neq i} x_j^*) = h^*(x_i + y^* - x_i^*) \geq h^*(y^*)$. Now, for any x_i with $u_i(x_i) \geq u_i(x_i^*)$ we have $b_i(x_i, U_i^*) \geq 0$. Hence, using the above result, we deduce that $u_i(x_i) \geq U_i^*$ implies that $h^*(x_i + y^* - x_i^*) \geq h^*(y^*)$. This prove the condition (a).

Now suppose that $h^*(y^*) \neq \min\{h^*(x_i + y^* - x_i^*) : x_i \in \mathcal{X}_i\}$, for each i. Assume that $x_i \in \mathcal{X}_i$ with $u_i(x_i) > U_i^*$. Then, from the above result, $h^*(x_i + y^* - x_i^*) \geq h^*(y^*)$. Suppose $h^*(x_i + y^* - x_i^*) = h^*(y^*)$. When $h^*(y^*) \neq \min\{h^*(x_i + y^* - x_i^*) : x_i \in \mathcal{X}_i\}$ and by continuity, there is some x_i' near x_i such that $u_i(x_i') > U_i^*$ and $h^*(x_i' + y^* - x_i^*) < h^*(x_i + y^* - x_i^*)$. It follows that $y_i' = x_i' + \sum_{j\neq i} x_j^* \in \mathcal{L}(U^*)$ and $h^*(y_i') < h^*(y^*)$. But this contradicts the separating property of h^*. Thus, $u_i(x_i) > U_i^*$ implies that $h^*(x_i + y^* - x_i^*) > h^*(y^*)$, which ends the proof. ■

Proof of corollary 3.2.4. Since the $\mathcal{X}_i$'s are convex and the u_i's are continuous and quasi-concave for each $i \in [n]$, it follows that the subset of $\mathcal{L}(U^*) = \{y : y = \sum_{i\in[n]} x_i, b_i(x_i, U_i^*) \geq 0, i \in [n]\}$ is closed and convex. All we need to prove is that the hyperplane $H \equiv \{y \in \mathbb{R}^m : \nabla h(y^*) \cdot y = \nabla h(y^*) \cdot y^*\}$ weakly separates $\mathcal{L}(U^*)$ and $\mathcal{Y}$. From proposition 3.1.1 $h^*(y^*) = \min\{h^*(y) : y \in \mathcal{L}(U^*)\} = 0$. Since $\mathcal{L}(U^*)$ is convex and h^* is differentiable at y^* we obtain the optimality condition $-\nabla h(y^*) \in N_{\mathcal{L}(U^*)}(y^*)$, where $N_{\mathcal{L}(U^*)}(y^*)$ is normal cone to $\mathcal{L}(U^*)$ at y^*. Moreover, we also have $h^*(y^*) = \max\{h^*(y) : y \in \mathcal{Y}\} = 0$. Consequently, since $\mathcal{Y}$ is closed and convex, the optimality conditions yield: $\nabla h(y^*) \in N_{\mathcal{Y}}(y^*)$. Let $T_{\mathcal{L}(U^*)}(y^*)$ and $T_{\mathcal{Y}}(y^*)$ denote the tangent cones to $\mathcal{L}(U^*)$ and $\mathcal{Y}$ at y^*, respectively. Since $\nabla h(y^*) \in N_{\mathcal{Y}}(y^*) \cap -N_{\mathcal{L}(U^*)}(y^*)$, the hyperplane H weakly separates $T_{\mathcal{L}(U^*)}(y^*)$ and $T_{\mathcal{Y}}(y^*)$. However, by definition $\mathcal{L}(U^*) \subset T_{\mathcal{L}(U^*)}(y^*)$ and $\mathcal{Y} \subset T_{\mathcal{Y}}(y^*)$. Consequently, H separates $\mathcal{L}(U^*)$ and $\mathcal{Y}$, which ends the proof. ■

Proof of proposition 4.1.1. Luenberger ([14], p. 465) proved result (a). To obtain result (b), first note that Luenberger ([14], p. 465) proved the upper semicontinuity of the benefit function. Thus, we just need to prove its lower semi-continuous. Assume that there is a sequence $\{(x_i^k, U_i^k)\}_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} (x_i^k, U_i^k) = (x_i, U_i) \quad \text{and} \quad b_i(x_i^k, U_i^k) > -\infty, \tag{6.1}$$

for all $k \in \mathbb{N}$. Using proof by contradiction, suppose that b_i is not lower semi-continuous. It follows that there is some $\gamma > 0$ such that

$$b_i(x_i^k, U_i^k) < b_i(x_i, U_i) - \gamma \tag{6.2}$$

for all $k \in \mathbb{N}$. Let $x_i^\gamma \equiv x_i - (b_i(x_i, U_i) - \gamma)g$. Since g is good, we have:

$$U_i \leq u_i\big(x_i - b_i(x_i, U_i)g\big) < u_i(x_i^\gamma) \tag{6.3}$$

for all $k \in \mathbb{N}$. Since the direction of g is interior to $\mathcal{X}_i$, we deduce that $x_i^\gamma = x_i - b_i(x_i, U_i)g + \gamma g$ lies in the interior of $\mathcal{X}_i$. Hence, there is some $\delta > 0$

such that $B(x_i^\gamma, \delta) \subset \mathcal{X}_i$, where $B(x_i^\gamma, \delta)$ is the open ball centered at x_i^γ of radius δ. Define

$$z^k \equiv x_i^k - (b_i(x_i, U_i) - \gamma)g, \tag{6.4}$$

which satisfies $\lim_{k \longrightarrow \infty} z^k = x_i^\gamma$. It follows that, for any $\delta > 0$, there exists k_δ such that, for all $k > k_\delta$,

$$\|z^k - x_i^\gamma\| < \delta. \tag{6.5}$$

In addition, since u_i is continuous, for any $\epsilon > 0$, there exists δ_ϵ such that

$$\|z - x_i^\gamma\| < \delta_\epsilon \Longrightarrow |u_i(z) - u_i(x_i^\gamma)| < \epsilon. \tag{6.6}$$

Using (6.1), for all $\epsilon > 0$, there exists k_ϵ such that, for all $k > k_\epsilon$,

$$|U_i^k - U_i| < \epsilon. \tag{6.7}$$

From (6.3), one can choose, $\epsilon > 0$ such that $u_i(x_i^\gamma) - \epsilon > U_i$. Pick some $k \in \mathbb{N}$ such that $k \geq \max\{k_\epsilon, k_{\delta/2}\}$.

Define $\bar{z}^k \equiv z^k - \frac{\delta}{3}\frac{g}{\|g\|}$. From equation (6.5), we have

$$\|\bar{z}^k - z^k\| \leq \delta/3 \Longrightarrow \|\bar{z}^k - x_i^\gamma\| \leq \delta/3 + \delta/2 < \delta.$$

Consequently $\bar{z}^k \in B(x_i^\gamma, \delta) \subset \mathcal{X}_i$. Moreover, equations (6.6) and (6.7) imply that $u_i(\bar{z}^k) > U_i^k$. From (6.4), we have

$$\bar{z}^k = x_i^k - (\frac{\delta/3}{\|g\|} + b_i(x_i, U_i) - \gamma)g.$$

Thus, $x_i^k - \big(\frac{\delta/3}{\|g\|} + b_i(x_i, U_i) - \gamma\big)g \in \mathcal{X}_i$ and

$$u_i\Big(x_i^k - (\frac{\delta/3}{\|g\|} + b_i(x_i, U_i) - \gamma)g\Big) > U_i^k.$$

It follows that $b_i(x_i^k - (\frac{\delta/3}{\|g\|} + b_i(x_i, U_i) - \gamma)g, U_i^k) \geq 0$. Using the translation property of the benefit function, we obtain $b_i(x_i^k, U_i^k) \geq b_i(x_i, U_i) + \frac{\delta/3}{\|g\|} - \gamma$. But this contradicts (6.2). ■

Proof of proposition 4.1.2. First, consider the shortage function $S(y)$. Define a sequence $\{y^k\}_{k \in \mathbb{N}}$ and some $\bar{y} \in \mathbb{R}^n$ such that $\lim_{k \to \infty} y^k = \bar{y}$ and $S(y^k)$ has finite values on $\mathbb{R}^m$ for all $k \in \mathbb{N}$. Since $\mathcal{Y}$ has an upper bound, the sequence $\{S(y^k)\}_{k \in \mathbb{N}}$ has a lower bound and $\liminf_{k \to \infty} S(y^k) > -\infty$. Let $\underline{\sigma} \equiv \liminf_{k \to \infty} S(y^k)$. Since $\mathcal{Y}$ is closed, $y - \underline{\sigma} g \in \mathcal{Y}$. Thus, $\underline{\sigma} \geq S(y)$. Hence S^+ is lower semi-continuous.

Next, we need to prove that S is upper semi-continuous, i.e. that $\overline{\sigma} \equiv \limsup_{k \to \infty} S(y^k) \leq S(y)$. Using proof by contradiction, suppose that S is not upper semi-continuous. It follows that there is some $\gamma > 0$ such that

$$S(y^k) > S(y) + \gamma \tag{6.8}$$

for all $k \in \mathbb{N}$. Let $y^\gamma \equiv y - (S(y) + \gamma)g$.

Since the direction of g is interior to $\mathcal{Y}$ and $y^\gamma = y - S(y)g - \gamma g$, y^γ lies in the interior of $\mathcal{Y}$. Hence, there is some $\delta > 0$ such that $B(y^\gamma, \delta) \subset \mathcal{Y}$ where $B(y^\gamma, \delta)$ is the open ball centered at y^γ of radius δ. Let

$$z^k \equiv y^k - (S(y) + \gamma)g, \tag{6.9}$$

which satisfies $\lim_{k \longrightarrow \infty} z^k = y^\gamma$. Hence, for all $\delta > 0$, there exists k_δ such that for all $k > k_\delta$

$$\|z^k - y^\gamma\| < \delta. \tag{6.10}$$

Define $\bar{z}^k \equiv y^k + \frac{\delta}{3} \frac{g}{\|g\|}$. From equation (6.10), we have

$$\|\bar{z}^k - z^k\| \leq \delta/3 \Longrightarrow \|\bar{z}^k - y^\gamma\| \leq \delta/3 + \delta/2 < \delta.$$

Consequently $\bar{z}^k \in B(y^\gamma, \delta) \subset \mathcal{Y}$. From (6.9), we have

$$\bar{z}^k = y^k - (-\frac{\delta/3}{\|g\|} + S(y) + \gamma)g.$$

It follows that $y^k - \big(-\frac{\delta/3}{\|g\|} + S(y) + \gamma\big)g \in \mathcal{Y}$, or $S(y^k - \big(-\frac{\delta/3}{\|g\|} + S(y) + \gamma\big)g) \leq 0$. Using the translation property of the shortage function, we obtain $S(y^k) \leq S(y) - \frac{\delta/3}{\|g\|} + \gamma$. But this contradicts (6.8).

Now consider the free-disposal shortage function S^+. First, remark that if $\mathcal{Y}$ has an upper bound then $\mathcal{Y}^+ \equiv \mathcal{Y} - \mathbb{R}^n_+$ has the same upper bound. Second, note that the logic used in the continuity proof of S applies to S^+. Thus, under the stated conditions, S^+ is continuous. ■

Proof of lemma 4.2.1. Note that

$$\sup_{X,y} \Big\{ L\big(X, y, U, f\big) : X \in \Pi, y \in \mathbb{R}^m \Big\} \geq \inf_f \Big\{ L\big(X, y, U, f\big) : f \in \Phi \Big\} \tag{6.11}$$

for all $X \in \Pi$, $y \in \mathbb{R}^m$, $f \in \Phi$. Using (4.3), (6.11) implies $L^*(U) \geq L^\#(U)$. This proves the first inequality in (4.4).

Suppose that $\sum_{i \in [n]} x_i > y$. Consider a sequence $f^k \in \Phi, k = 1, 2, ...,$ such that $f^k(v) = \alpha + k \sum_{j \in [m]} v_i$, with $v = (v_1, ..., v_m)$ and α being an arbitrary constant. This implies that for $k \geq 1$:

$$f^k(y) - f^k\big(\sum_{i \in [n]} x_i\big) = k\big(\sum_{j \in [m]} y_j - \sum_{j \in [m]} \sum_{i \in [n]} x_{ij}\big) < 0.$$

Consequently $\lim_{k \longrightarrow \infty} P\big(f^k(y) - f^k(\sum_{i \in [n]} x_i)\big) = -\infty$. Hence, the condition $\sum_{i \in [n]} x_i > y$ implies that $\inf_f \Big\{ L\big(X, y, U, f\big) : f \in \Phi \Big\} = -\infty$. It follows from (2.5) that

$$\begin{aligned} L^{\#}(U) &= \sup_{x,y} \inf_{f} \Big\{ L(X,y,U,f) : \sum_{i\in[n]} x_i \le y, X \in \Pi, y \in \mathbb{R}^m, f \in \Phi \Big\} \\ &\ge \sup_{X,y} \{ \sum_{i\in[n]} b_i(x_i, U_i) - S(y) : \sum_{i\in[n]} x_i \le y, X \in \Pi, y \in \mathbb{R}^m \} \qquad (6.12) \\ &= V(U), \end{aligned}$$

which proves the second inequality in (4.4). ■

Proof of proposition 4.2.2. Equation (4.5) follows from (4.4) and (4.3) when $L^*(U) = L^{\#}(U)$. The second inequality in (4.5) implies that

$$f^*(y^*) - f^*(\sum_{i\in[n]} x_i^*) \le f(y^*) - f(\sum_{i\in[n]} x_i^*), \quad \forall f \in \Phi. \tag{6.13}$$

Assume that $\sum_{i\in[n]} y_i^* > x^*$. Then, there exists a strictly increasing linear function f^a defined on $\mathbb{R}^m$ by $f^a(v) = \alpha + a \cdot v$ where $a \in \mathbb{R}^m_{++}$, satisfying $f^a(y^*) - f^a(\sum_{i\in[n]} x_i^*) < 0$. Also, letting $f^b(v) = \alpha + b \cdot v$, where $b > a$, we have $f^b(y^*) - f^b(\sum_{i\in[n]} x_i^*) < f^a(y^*) - f^a(\sum_{i\in[n]} x_i^*) < 0$. This implies that $f(y^*) - f(\sum_{i\in[n]} x_i^*)$ does not have a lower bound, which contradicts (6.13). This gives (4.6).

Note that $f^*(y^*) - f^*(\sum_{i\in[n]} x_i^*) = 0$ when $\sum_{i\in[n]} x_i^* = y^*$. Consider the case where $\sum_{i\in[n]} x_i^* \le y^*$ and $\sum_{i\in[n]} x_i^* \ne y^*$. Choosing $f^c \in \Phi$ such that $f^c(v)$ does not depend on v, (6.13) implies that $f^*(y^*) - f^*(\sum_{i\in[n]} x_i^*) \le 0$. The function $f^* \in \Phi$ being non-decreasing, it follows that $\sum_{i\in[n]} x_i^* \le y^*$ and $\sum_{i\in[n]} x_i^* \ne y^*$ imply that $f^*(y^*) - f^*(\sum_{i\in[n]} x_i^*) \ge 0$. Combining these results yields (4.7).

Assuming that a maximal allocation exists, using (4.6) and (4.7), the first inequality in (4.5) implies (4.8). Finally, using (4.5) and (2.5), equations (4.7) and (4.8) yield (4.9). ■

Proof of lemma 4.2.3. $L^*(U) = L^{\#}(U)$ implying $L^*(U) = V(U)$ was shown in the previous proposition. The converse follows from (4.4). ■

Proof of proposition 4.2.4. Assume that $W(U, \gamma)$ is not upper semi-continuous in γ at $\gamma = 0$. This means that there exist a $d > 0$ and a sequence $\{\gamma^k : k = 1, 2, ...\}$ satisfying $\lim_{k\to\infty} \gamma^k = 0$ and

$$W(U, \gamma^k) \ge W(U, 0) + d, k = 1, 2, ... \tag{6.14}$$

Under a zero duality gap (where $L^*(U) = V(U)$), (6.14) means that there exists a $f \in \Phi$ satisfying

$$
\begin{aligned}
W(U,0) &= V(U) \\
&> \sup_{\substack{y\in\mathbb{R}^m\\ X\in\Pi}} \Big\{ \sum_{i\in[n]} b_i(x_i,U_i) - S(y) + f(y) - f(\sum_{i\in[n]} x_i) \Big\} - d/2 \\
&\geq \sup_{\substack{y\in\mathbb{R}^m\\ X\in\Pi}} \Big\{ \sum_{i\in[n]} b_i(x_i,U_i) - S(y) + f(y) - f(\sum_{i\in[n]} x_i) \colon \sum_{i\in[n]} x_i \leq y+\gamma^k \Big\} - d/2.
\end{aligned}
$$

Choose $y^k \in \mathbb{R}^m$ and $X^k \in \Pi$ such that $\sum_{i\in[n]} x_i^k \leq y^k + \gamma^k$ and satisfying $\sum_{i\in[n]} b_i(x_i^k, U_i) - S(y^k) \geq W(U,\gamma^k) - d/4$. It follows that

$$
\begin{aligned}
W(U,0) &> \sup_{\substack{y\in\mathbb{R}^m\\ X\in\Pi}} \Big\{ W(U,\gamma^k) - d/4 + f(x) - f(\sum_{i\in[n]} x_i) \colon \sum_{i\in[n]} x_i \leq y+\gamma^k \Big\} - d/2, \\
&\geq W(U,\gamma^k) - 3d/4 + f(x^k) - f(\sum_{i\in[n]} x_i^k).
\end{aligned}
\tag{6.15}
$$

Combining (6.14) and (6.15), we obtain $W(U,\gamma^k) > W(U,\gamma^k) + d/4 + f(y^k) - f(\sum_{i\in[n]} x_i^k)$. Since $f \in \Phi$ is continuous and non-decreasing, letting $k \to \infty$ implies that $W(U,\gamma^k) > W(U,\gamma^k) + d/4$, a contradiction. ■

Proof of proposition 4.2.5. Suppose that a zero duality gap does not hold. From (4.4), this means that $V(U) < L^*(U)$. Thus, there exists a $\delta > 0$ such that

$$
\begin{aligned}
V(U) = W(U,0) &\leq L^*(U) - \delta, \\
&\leq \sup_{\substack{y\in\mathbb{R}^m\\ X\in\Pi}} \Big\{ \sum_{i\in[n]} b_i(x_i,U_i) - S(y) + f(y) - f(\sum_{i\in[n]} x_i) \Big\} - \delta
\end{aligned}
\tag{6.16}
$$

Consider a sequence $\{\gamma^k\}_{k\in\mathbb{N}}$ where $\gamma^k = \overline{\gamma}/k, \overline{\gamma} \in \mathbb{R}^m_{++}$. From (6.16), choose $y^k \in \mathbb{R}^m$ and $x^k \in \Pi$ such that $\sum_{i\in[n]} x_i^k = y^k + \gamma^k$ and satisfying

$$
W(U,0) \leq \sum_{i\in[n]} b_i(x_i^k, U_i) - S(x^k) + f(y^k) - f(\sum_{i\in[n]} x_i^k) - \delta/2. \tag{6.17}
$$

Note that

$$
\begin{aligned}
W(U,\gamma^k) &= \sup_{\substack{y\in\mathbb{R}^m\\ X\in\Pi}} \Big\{ \sum_{i\in[n]} b_i(x_i,U_i) - S(y) : \sum_{i\in[n]} x_i \leq y + \gamma^k \Big\} \\
&\geq \sum_{i\in[n]} b_i(x_i^k, U_i) - S(y^k).
\end{aligned}
\tag{6.18}
$$

Combining (6.17) and (6.18) yields $W(U,0) \leq W(U,\gamma^k) + f(y^k) - f(\sum_{i\in[n]} x_i^k) - \delta/2$. Letting $k \to \infty$ and using the continuity of f, this implies $W(U,0) \leq$

$\limsup_{k\to\infty} W(U,\gamma^k) - \delta/2$, which contradicts the upper semi-continuity of $W(U,\gamma)$ at $\gamma = 0$. ■

Proof of lemma 4.3.1. We have from (4.11):

$$\begin{aligned}
&\inf_X \{f(\sum_{i\in[n]} x_i) : u_i(x_i) \geq U_i, i \in [n], X \in \Pi\} \\
&\geq \inf_X \{f(\sum_{i\in[n]} x_i) - \sum_{i\in[n]} b_i(x_i, U_i) : u_i(x_i) \geq U_i, i \in [n], X \in \Pi\} \\
&\geq \inf_X \{f(\sum_{i\in[n]} x_i) - \sum_{i\in[n]} b_i(x_i, U_i) : X \in \Pi\} \\
&= E(f, U).
\end{aligned}$$

We now need to show that $E(f,U) \geq \inf_X\{f(\sum_{i\in[n]} x_i) : u_i(x_i) \geq U_i, i \in [n], X \in \Pi\}$. From (4.11), this inequality clearly holds if $\sum_{i\in[n]} b_i(x_i, U_i) = -\infty$. Consider the case where $\sum_{i\in[n]} b_i(x_i, U_i) > -\infty$. Letting $x_i' = x_i - b_i(x_i, U_i)g$, we have $x_i' \in \mathcal{X}_i$, $u_i(x_i') \geq U_i$, and $b_i(x_i', U_i) = 0$ for all $i \in [n]$. Using the translation property of both b_i and f, we obtain

$$\begin{aligned}
E(f,U) &= \inf_X \left\{ f(\sum_{i\in[n]} x_i) - \sum_{i\in[n]} b_i(x_i, U_i) : X \in \Pi \right\} \\
&= \inf_{X,X'} \left\{ f(\sum_{i\in[n]} x_i') - \sum_{i\in[n]} b_i(x_i', U_i) : x_i' = x_i - b_i(x_i, U_i)g \in \mathcal{X}_i, i \in [n], X \in \Pi \right\} \\
&= \inf_{X,X'} \left\{ f(\sum_{i\in[n]} x_i') : x_i' = x_i + b_i(x_i, U_i)g) \in \mathcal{X}_i, u_i(x_i') \geq U_i, i \in [n], X \in \Pi \right\} \\
&\geq \inf_{x'} \left\{ f(\sum_{i\in[n]} x_i') : u_i(x_i') \geq U_i, x_i' \in \mathcal{X}_i, i \in [n] \right\},
\end{aligned}$$

which concludes the proof. ■

Proof of lemma 4.3.2. We have

$$\begin{aligned}
&\sup_y \{f(y) : y \in \mathcal{Y}\} \\
&\leq \sup_y \{f(y) - S(y) : y \in \mathcal{Y}\} \\
&\leq \sup_x \{f(y) - S(y) : y \in \mathbb{R}^m\} \\
&= \pi(f).
\end{aligned}$$

We now need to show that $\pi(f) \leq \sup_y\{f(y) : y \in \mathcal{Y}\}$. From (4.13), this inequality clearly holds if $S(y) = +\infty$. Consider the case where $S(y) < +\infty$. Letting $y' = y - S(y)g$, we have $y' \in \mathcal{Y}$, and $S(y') = 0$. Using the translation property of both S and f, we obtain

$$\begin{aligned}\pi(f) &= \sup_y \{f(y) - S(y) : y \in \mathbb{R}^m\} \\ &= \sup_{y,y'} \{f(y') - S(y') : y' = y - S(y)g \in \mathcal{Y}, y \in \mathbb{R}^m\} \\ &= \sup_{y,y'} \{f(y') : y' = y + S(y)g \in \mathcal{Y}, y \in \mathbb{R}^m\} \\ &\leq \sup_{y'} \{f(y') : y' \in \mathcal{Y}\},\end{aligned}$$

which concludes the proof. ■

Proof of proposition 4.3.3. Equation (4.16) is obtained directly from (4.15). ■

References

[1] Armstrong, M. Multiproduct Nonlinear Pricing, *Econometrica* 64(1996): 51-75.

[2] Armstrong, M. amd J. Vickers. Competitive Price Discrimination, *Rand Journal of Economics* 32(2001): 579-605.

[3] Berge, C. *Topological Spaces* Oliver and Boyd, Edinburgh, 1963.

[4] Bertsekas, D.P. *Nonlinear Programming* Athena Scientific, Belmont, MA, 1995.

[5] Boiteux, M. Sur la Gestion des Monopoles Publics Astreints l'Equilibre Budgetaire, *Econometrica* 24(1956): 22-40.

[6] Bonnisseau, J.M. and B. Cornet. Valuation Equilibrium and Pareto Optimum in Non-Convex Economies. *Journal of Mathematical Economics* 17(1988): 293-308.

[7] Courtault, J.M. and J.M. Tallon. Allais's Tading Process and the Dynamic Evolution of a Market Economy, *Economic Theory* 16(2000): 477-481.

[8] Clarke, F.H. *Optimization and Nonsmooth Analysis* Classics Appl. Math. 5, SIAM, Philadelphia, 1990.

[9] Coase, R.H. The Problem of Social Cost, *Journal of Law and Economics* 3(1960): 1-44.

[10] Debreu, G. *Theory of Value* Wiley, New York, 1959.

[11] Guesnerie, R. Pareto Optimality in Non-Convex Economies, *Econometrica* 43,1 (1975): 1-29.

[12] Jofre, A. and J.R. Cayupi. A Nonconvex Separation Property and Some Applications, *Mathematical Programming, Series A* 108(2006): 37-51.

[13] Laffont, J.j. and J. Tirole. *A Theory of Incentives in Regulation and Procurement.* MIT Press, Cambridge, MA, 1992.

[14] Luenberger, D.G. Benefit Functions and Duality, *Journal of Mathematical Economics* 21(1992a): 461-481.

[15] Luenberger, D.G. New Optimality Principles for Economic Efficiency and Equilibrium, *Journal of Optimization Theory and Applications* 75, 2 (1992b): 221-264.

[16] Luenberger, D.G. Optimality and the Theory of Value, *Journal of Economic Theory*, 63, no. 2 (1994), 147-169.

[17] Luenberger, D.A. *Microeconomic Theory.* McGraw-Hill, New York, 1995.

[18] McAfee, R. P., J. McMillan and M. Whinston. Multiproduct Monopoly, Commodity Bundling and Correlation of Values, *Quarterly Journal of Economics* 114(1989): 371-384.

[19] Mirlees, J.A. Optimal Tax Theory: A Synthesis, *Journal of Public Economics* 6(1976): 327-358.

[20] Mangasarian, O.L. *Nonlinear Programming.* McGraw-Hill, New York, 1969.

[21] Mordukhovich, B.S. An Abstract Extremal Principle with Applications to Welfare Economics *Journal of Mathematical Analysis and Applications.* 251(2000): 187-216.

[22] Ordover, J.A. and J.C. Panzar. On the Nonexistence of Partoa Suyperior Outlay Schedules, *Bell Journal of Economics* 11(1980): 351-354.

[23] Oren, S.S., S.A. Smith, and R.B. Wilson. Competitive Nonlinear Tariffs, *Journal of Economic Theory* 29(1982): 49-71.

[24] Radner, R. Competitive Equilibrium under Uncertainty, *Econometrica* 36(1968): 31-56.

[25] Tirole, J. *The Theory of Industrial Organization.* MIT Press, Cambridge, MA, 1992.

[26] Ramsey, F.P. A Contribution to the Theory of Taxation, *Economic Journal* 37(1927): 47-61.

[27] Rockafellar, R.T. and R.J. Wets, *Variational Analysis.* Springer-Verlag, Berlin, 1998.

[28] Samuelson, P.A. Social Indifference Curves, *Quarterly Journal of Economics* 70(1956): 1-22.

[29] Rubinov, A.M., X.X. Huang and X.Q. Yang. The Zero Duality Gap Property and Lower Semicontinuity of the Perturbation Function, *Mathematics of Operations Research* 27(2002): 775-791.

[30] Wilson, R.B. *Nonlinear Pricing.* Oxford University Press, Oxford, 1993.