Pareto Optimality in Non-Convex Economies and Marginal Cost Pricing Equilibria

Kwan Koo Yun*

I. Introduction

The Second welfare theorem in general equilibrium theory states that, under convexity assumptions, a Pereto optimal allocation may be realized as a competitive equilibrium after a suitable redistribution of income. When individual convexity assumptions are violated, as in the case of the increasing returns to scale, Guesnerie (1) formulated and proved a modified version of the second welfare theorem. The theorem states that at a Pareto optimal allocation, one may find a price vector such that, at each consumer's consumption vector, the first order conditions for expenditure minimization in reaching the utility level of the consumption are satisfied (with respects to the price vector) and at each producer's production vector, the first order conditions for profit maximization are satisfied. Such a state of an economy will be called a marginal cost pricing equilibrium in this paper. Guesnerie's arguments critically depend upon replacing the convexity assumptions by the convexity assumptions of certain approximating cones to the sets under consideration. In this paper, we show that such assumptions are largely unnecessary. In fact, our approach to the problem is very different. We first establish some necessary conditions for a Pareto optimal solution for maximizing n (real valued) functions under general constraints. We also discuss constraint qualifications for such problems. This is the objective of the first part of the paper. In the second part of the paper, we describe an economy in which non-convexities may be present. The results in the first part are then applied to obtain characterizations of a Pareto optimal allocation. In particular, it is shown that a Pareto optimal allocation may be realized as a marginal cost pricing equilibrium under very general conditions. Finally, some comments are made on possible directions in which the analyses may be extended.

*State University of New York at Albany, U.S.A. This paper was presented at The International Convention of Korean Economists, Seoul, 20-25 Aug., 1984.

II. Necessary Conditions for Pareto Optimality

We condiser the following problem of maximizing n functions under constraints.

$$\underset{s \in X}{\text{maximize}}\ f_1, f_2, \ldots, f_n \tag{A}$$

subject to $g(s) \leqslant e$ and $s \in C$.

Here, X is an open set in R^p, C is a closed (in R^p) subset of X and $f: X \rightarrow R^n$, $g: X \rightarrow R^\ell$ are continuously differentiable. f_i in the above problem denotes the ith component function of f.

The *feasible set* F is defind by $F = \{s \in X \mid g(s) \leqslant e \text{ and } s \in C\}$. $\geqslant$, $>$, $\gg$denote weak, semi-strict, and strict vector inequalities respectively.

One may consider different concepts of solution to the above problem. $s \in F$ is a *Pareto optimum* if there does not exist $s' \in F$ such that $f(s') > f(s)$. $s \in F$ is a *weak Pareto optimum* if there does not exist $s' \in F$ such that $f(s') \gg f(s)$. Note that weak Pareto optimality is a more general concept than Pareto optimality. If $s \in F$ is a Pareto optimum with respect to a neighborhood of it in F, it is called a *local Pareto optimum*. A *weak local Pareto optimum* is defined similarly.

In the description of necessary conditions for Pareto optimality that follows, we need the concepts of conical approximation of a closed set at a point in the set. $A \subset R^n$ is a *cone* if $x \in A$ implies $\lambda \cdot x \in A$ for all $\lambda > 0$. A is a *convex cone* if $x, y \in A$ implies that $\alpha \cdot x + \beta \cdot y \in A$ for all $\alpha, \beta > 0$. Given a set A in R^n, its (negative) *polar cone* A^0 is defined by $A^0 = \{v \in R^n \mid v \cdot w \leqslant 0, \text{ for all } w \in A\}$. A^0 is a closed convex cone. Given two sets A, B in R^n such that $A \supset B$, we have $B^0 \supset A^0$. Also, $A^0 = (\bar{A})^0$. Here, $\bar{A}$ denotes the closure of A. Let B be the open unit ball in R^n and C a closed set in R^n.

Given $x \in C$, we define:

Hypertangent cone (of C at x);

$H_c(x) = \{v \in R^n \mid \exists\ \varepsilon > 0 \text{ such that } y + t \cdot w \in C \text{ for all } y \in (x + \varepsilon \cdot B) \cap C,\ w \in v + \varepsilon \cdot B,\ t \in (0, \varepsilon)\}$.

Clark's tangent cone;

$T_C(x) = \{v \in R^n \mid \text{for all } \{x_i\} \subset C \text{ converging to } x \text{ and } \{t_i\} C(0, \infty) \text{ decreasing to } 0,\ \exists\ \{v_i\} \text{ in } R^n \text{ converging to } v \text{ such that } x_i + t_i \cdot v_i \in C \text{ all } i\}$.

curvilinear tangent cone;

$T_c'(x) = \{v \in R^n \mid$ there is a continuously differentiable function $c: (0, a) \rightarrow C$, for some $a > 0$ such that $c(0) = x$ and $DC(0) = v\}$. D stands for the derivative.

contingent cone;

$K_c(x) = \{v \in R^n |$ there is a sequence $\{x^i\} \subset C$ converging to x and a sequence $\{t_i\} \subset (0, \infty)$ decreasing to 0 such that $(x_i - x)/t_i \to v\}$.

Finally, when the smallest linear subspace of R^n containing $C - \{x\}$ is R^n, we define:

cone of interior displacement;

$k(C, x) = \{ v \in R^n | \exists\ \varepsilon > 0$ such that $x + t \cdot w \in C$ for all $w \in v + \varepsilon \cdot B$ and $t \in (0, \varepsilon)\}$.

When there is a proper subspace of R^n containing $C - \{x\}$, we redefine B in the above definition as the open unit ball in the subspace. It is immediate from their definitions that $k(C, x) \subset T'_C(x)$ and $H_C(x) \subset T_C(x) \subset K_C(x)$. Also, $T'_C(x) \subset K_C(x)$. $H_C(x)$ and $K(C, x)$ are open cones and $T_C(x)$ is a closed convex cone. $K_C(x)$ is a closed cone but need not be convex. If $H_C(x)$ is not empty, $H_C(x)$ = interior of $T_C(x)$ (see Clarke (2), Theorem 2.4.8.). Since the closure of the interior of a convex body is equal to the closure of the convex body and since $T_C(x)$ is closed, we have the following proposition.

Proposition 1: If $H_C(x) \neq \phi$, $T_C(x)^0 \supset T'_C(x)^0$

Proof: $T_C(x)^0 = \text{int}\ \overline{T_C(X)}^0 = \text{int}\ T_C(x)^0 = H_C(x)^0 \supset T'_C(x)^0$. Q.E.D. Here, int denotes the interior operation relative to usual topology in R^n, and the upper bar in second term denotes the closure.

Henceforth, we shall write $T_C(x)^0$ as $N_C(x)$.

Given $S \subset R^n$, we define:

$$C(S) = \{ \sum_{i=1}^{k} \lambda_i \cdot s_i \mid k \in \{1, 2, \ldots\},$$
$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k) > 0 \text{ and } s_i \in S, \text{ all } i\}.$$

C(S) is different from the conical hull of S in that λ is not allowed to be the zero vector. C(S) is a convex cone which may or may not contain zero.

The following Lemmas are fundamental in what follows.

Lemma 1: Consider $V = \{v_1, v_2, \ldots, v_m\} \subset R^n$ and a closed convex cone K in R^n. If $C(V) \cap K = \phi$, then there exist a hyperplane through 0 separating C(V) and K such that C(V) is contained in an open half space.

Proof: Given $\varepsilon > 0$ and the unit open ball B in R^n, it is straightforward to prove (using the convexity of B):

$$\bigcup_{v'_i \in (v_i + \varepsilon \cdot B)} C[v'_1, v'_2, \ldots, v'_m] = C[\bigcup_i (v_i + \varepsilon \cdot B)].$$

The second set is a convex cone containing C(V) in its interior. We now show that for $\varepsilon > 0$, small enough, the first set is disjoint from K. Suppose not. Then, there are $v^{\nu} = (v_1^{\nu}, v_2^{\nu}, \ldots, v_m^{\nu})$ converging to $v = (v_1, v_2, \ldots, v_m)$ and $\lambda^{\nu} > 0$ such that $\sum_{i=1}^{m} \lambda_i^{\nu} \cdot v_i^{\nu} \in K$, for each ν. Since K is a cone and $\lambda^{\nu} > 0$, we may assume $\| \lambda^{\nu} \| = 1$. Here, $\| \;\|$ denotes the Euclinean norm. Thus, $\{ \lambda^{\nu} \}$ has a convergent subsequence and we may assume, without loss of generality, that λ^{ν} converges to $\lambda > 0$. Then, $\sum_{i=1}^{m} \lambda_i \cdot v_i \in$ K by the closedness of K and contradicts our hypothesis. We can now apply a separation theorem to obtain a hyperplane through zero separating $C\{ \cup(v_i + \varepsilon \cdot B)\}$ from K when $\varepsilon > 0$ is small enough. Q.E.D.

Lemma 2: Consider a non-empty convex cone A in R^n and $V = \{v_1, v_2, \ldots, v_m\}$ in R^n. Suppose there does not exist $a \in A$ such that $v_i \cdot a > 0$, all i, then one can find $\lambda > 0$ such that $\sum_{i=1}^{m} \lambda_i \cdot v_i \in A^0$.

Proof: Suppose $C(V) \cap A^0 = \phi$. Then, by Lemma 1, there exist a hyperplane separating C(V) from A^0 such that C(V) is contained in an open half space. Let b be a non-zero normal vector to the hyperplane pointing inward to the half space containing V. Then $b \in A^{00}$. Since A is convex, $\bar{A} = \bar{A}^{00} = A^{00}$. Thus, $b \in \bar{A}$. Moreover, $v_i \cdot b > 0$, all $i = 1, 2, \ldots m$, since V is contained in an open half space. This means that, for all b′ close enough to b, $v_i \cdot b' > 0$, for all i. Thus, there exists $a \in A$ such that $v_i . a > 0$, all i. Thus a contradiction. Q.E.D.

Now we prove the main mathematical theorem in this paper.

Theorem 1: Suppose s is a weak local Pareto optimum for the problem (A). Then, givne any convex set $A \subset K_C(s)$, there exists $(\lambda, r) > 0$ such that

$\sum_{i=1}^{n} \lambda_i \cdot Df_i(s) - \sum_{k=1}^{\ell} r_k \cdot D_{g_k}(s) \in A^0$ and $r_k . g_k(s) = O$, all k.

Here, $Df_i(s)$ (resp. $Dg_k(s)$) denotes the derivative of f_i (resp. g_k) at s.

Proof: Let I(s) be $\{k \mid g_k(s) = e_k\}$. Since s is a weak local Pareto optimum, there does not exist $c \in K_C(s)$ such that $Df_i(s).c > 0$, all i and $-Dg_k(s).c > 0$, all $k \in I(s)$. If there is such a c in $K_C(s)$, it is easy to show that there is s′ arbitarily close to s such that $s' \in C$, $f_i(s') > f_i(s)$, all i and $g_k(s') < e_k$ for each $k \in I(s)$, violating the local weak Pareto optimality of s (This follws directly from the definition of the derivative and the definition of $K_C(s)$). Now, let

$V = \{ Df_i(s) \mid \text{all } i\} \cup \{ -Dg_k(s)\ k \in I(s)\}$.

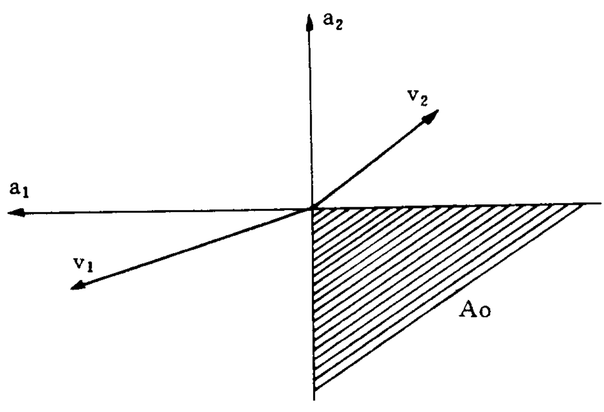
By Lemma 2, there are non negative numbers, $\{\lambda_i\}_{i=1}^{n}$ and $\{r_{k_0}\}$ $k \in I(s)$, not all zero such that $\sum_{i=1}^{n} \lambda_i \cdot Df_i(s) - \sum_{k\in I(S)} r_k \cdot Dg_k(s) \in A^\circ$. By defining $r_k = 0$, for $k \notin I(s)$, we obtain the theorem. Q.E.D.

Corollary 1: Suppose s is a weak local Pareto optimum for the problem (A). Then, there exists $(\lambda, \gamma) > 0$ such that $\sum_{i=1}^{n} \lambda^i \cdot Df_i(s) - \sum_{k=1}^{\ell} r_{k'} \cdot Dg_k(s) \in N_C(s)$, and $r_k . g_k(s) = 0$, all k.

Proof: $T_C(s)$ is always a non-empty convex cone contained in $K_C(s)$. Q.E.D.

Corollary 2: If $K_C(s)$ is convex at a weak local Pareto optimum for the problem (A), the Lagrangian expression in Theorem 1 may be stated in terms of $K_C(s)^0$. Similarly, if $T_C'(s)$ (or k (C, s)) is non-empty, convex at a weak local Pareto optimum, then the Lagrangian expression in Theorem 1 may be stated in terms of $T_C'(s)^0$ (or k $(C, s)^0$).

Guesnerie (1) uses the assumption that k(C, s) is nonempty, convex at a Pareto optimum for his characterizations of the Pareto optimum. Theorem 1 gives a nice perspective on this assumption. Roughly speaking, theorems that are obtained in the second part of this paper follow from Theorem 1. Corresponding to different choices of the set A in theorem 1, we obtain similar but different results. The assumption that k(C, s) is non-empty, convex at a Parto optimum is just one of many possible choices. In this paper, we shall use the Clarke tangent cone because it is always a convex subset of the contingent cone. Also, the Clarke tangent cone coincides with other natural concepts of tangent cones when C is nice at s (e.g., when C is (locally) convex at s or smooth).



[Fig. 1]

The convexity in the choice of A in Lemma 2 (and consequently, in theorem 1) is indispensable. [Fig. 1] gives a counterexample, when $A = \{a_1\ a_2\}$, $V = \{v_1, v_2,\}$.

We now present an alternative way of obtaining Corollary 1. We start from the necessary conditions for a maximum of a mathematical programming problem involving locally Lipschitz objective function. Then, we convert the problem of maximizing objective function. Then, we convert the problem of maximizing n functions under constraints into an equivalent mathematical programming problem with a locally Lipschitz objective function and obtain the results. Our ultimate goal here is to discuss Clarke's contributions to "constraint qualifications", developed in connection with mathematical programming and then to apply them to the Pareto optimality case.

Let f be a locally Lipschitz real valued function defined on an open set U in R^n. If Ω_f denotes the set of Lebesgue measure zero in U on which f is not differentiable and S any set of measure zero, the *generalized gradient* of f, at x, $\partial f(x)$, is defined by:

$$\partial f(x) = \mathrm{co}\{\lim Df(x_i) \mid x_i \to x,\ x_i \notin \Omega_f \cup S\}.$$

Here, co denotes the convex hull. The generalized gradient is not empty since Df(x) is bounded (whenever it is defined) on a neighborhood of x by the Lipschitz property of f. $\partial f(x)$ reduces to Df(x) if f is continuously differentiable at x. Also, it is easily verified that $\partial(-f(x)) = -\partial f(x)$.

Instead of developing necessary conditions for Pareto optimum from scratch, one can start from the following mathematical programming problem and necessary conditions for a solution of such a problem given by Clarke.

mathematical programming

$$\max_{s \in X} \text{imize } f \quad \text{subject to } g(s) \leqslant e \text{ and } s \in C. \tag{B}$$

The interpretations of the symbols are the same as in the problem (A) except that; in the above problem, there is one real valued function f to maximize and f is assumed to be only locally Lipschitz.

Clarke's Theorem 1 (Clarke (2), Theorem 6.1.1.)

If s is a local maximum for the problem (B), there is

$(\lambda, r) \in R_+^{\ell+1} - \{0\}$ such that $0 \in -\lambda \cdot \partial f(s) + \sum_{i=1}^{\ell} r_k \cdot Dg_k(s) + N_C(s)$ and

$(r_k \cdot g_k(s) = O,\ k = 1, 2, \ldots, \ell.$

In the above, $R_+{}^{\ell+1}$ dnotes the nonnegative orthant of $R^{\ell+1}$ and $Dg_k(s)$

denotes the derivative of g^k at s.

We obsrve now that s is a weak local Pareto optimum of the problem (A) if and only if s is a local maximum of the following problem.

$$\text{maximize } \min \{ f_1(s) - f_1(\bar{s}),\ f_2(s) - f_2(\bar{s}), \ldots, f_n(s) - f_n(\bar{s}) \} \quad \text{(C)}$$
$$\text{subject to } g(s) \leqslant e \text{ and } s \in C.$$

min denotes the minimum operation.

Let h(s) represent min $\{f_1(s) - f_1(\bar{s}),\ f_2(s) - f_2(\bar{s}), \ldots,\ f_n(s) - f_n(\bar{s})\}$ and $h_i = f_i(s) - f_i(\bar{s})$.

Then, $h(s) = -\max_i \{ -h_1(s), -h_2(s), \ldots, -h_n(s)\}$.

By a proposition of Clarke ((2), Proposition 2.3.12), $\partial h(s) - co\{ -\partial h_i(s) \mid i = 1, 2, \ldots, n\}$. Since h_i is continuously differentiable, $\partial h_i(s) = Df_i(s)$ and $\partial h_i(s) \subset \{ \Sigma \mu_i \cdot Df_i(s) \mid \mu_i \geqslant 0,\ i = 1, 2, \ldots, n \text{ and } \Sigma \mu_i = 1\}$. These observations together with Clarke's Theorem 1 prove Corollary 1.

We have already observed that s is a weak local Pareto optimum of the problem (A) if and only if it is a local maximum of the problem (C). We now observe that s is a local Pareto optimum of the problem (A) if and only if it is a local maximum of the following n problems simultaneously.

For each $i = 1, 2, \ldots, n$;

$$\text{maximize}_{s \in X}\ f_i$$
$$\text{subject to}$$
$$f_{i'}(\bar{s}) \geqslant f_{i'}(\bar{s}), \text{ all } i' = i, \quad \text{(D)}$$
$$g(s) \leqslant e \text{ and } s \in C.$$

If $\lambda = 0$ in Corollary 1, the Lagrangian expression is not very useful. So, now we seek conditions undr which λ in Corollary 1 may be chosen as a non-zero vector. For the mathematical programming problem (B), Clarke (3) introduced the calmness condition for such a purpose. Let F_e denote the feasible set in the problem (B) corresponding to the given vector e. We now difine function $\phi : R^{\ell} \to [\infty, \infty]$ by $\phi(e) = \sup_{s \in F_e} f(s)$. $\phi(e)$ is defined to be $-\infty$ if $F_e = \phi$. Problem (B) is defined to be *calm* if $\phi(e)$ is finite and $\limsup_{e' \to e} (\phi(e') - \phi(e))/\|e'\| < \infty$. Calmness conditions are satisfied by most of the familiar constraints qualifications.

We now state:

Clarke's Theorem 2 (Clarke (3))

Suppose that the problem (B) is calm (with respect to the given e), then λ in the statement of the Clarke's Theorem 1 may be chosen to be positive. Moreover, given a neighborhood of E of e on which ϕ is finite, the problem (B) is calm with respect to almost all (in Lebesgue sence) e ' in E.

In order to use the above theorem in the context of a weak local Pareto optimum, we have to localize the concept of calmness. We shall say that the problem (B) is *locally calm* at $s \in F_e$ if one may find a closed δ-ball (around zero), $\bar{B}_\delta$, $\delta > 0$ such that the problem (B) is calm with respect to the feasible set;

$$F_{s,\delta}(e) = \{s' \in X | s' \in s + \bar{B}_\delta,\ g(s') \leqslant e \text{ and } s' \in C\}.$$

Suppose $\bar{s}$ is a weak local Pareto optimum for the problem (A) and thus is a local maximum for the problem (C) and suppose problem (C) is locally calm at s. Then, the arguments in the derivation of Theorem 1 make it clear that one may choose $\lambda > 0$ in the Lagrangian expression in Theorem 1. Suppose now that $\bar{s}$ is a local Pareto optimum and thus is a local maximum for each of the n problems in (D). If each of the problems is locally calm at s, we may choose for each i, $(\lambda^i, r^i) \geqslant 0$, $\lambda^i_i > 0$ such that $\sum_{j=1}^{n} \lambda^i_j Df_j(\bar{s}) + \sum_{k=1}^{\ell} r^i_K Dg_k(\bar{s}) \in N_c(\bar{s})$ and $r^i_k \cdot g_k(\bar{s}) = 0$, for all k. If we choose $\lambda_j = \sum_{j=1}^{n} \lambda^i_j/n$, $j = 1,2,\ldots, n$ and $r_k = \sum_i r^i_k/n$, then by the convexity of $n_C(\bar{s})$, $\sum_{j=1}^{n} \lambda_j Df_j(\bar{s}) + \sum_{k=1}^{\ell} r_k Dg_k(\bar{s}) \in N_C(\bar{s})$. Also, $r^k \cdot g^k(\bar{s}) = 0$ for all k. Thus, we have the following proposition.

Proposition 2: If $\bar{s}$ is a weak local Pareto optimum for the problem (A) and problem (C) is locally calm at $\bar{s}$, λ in the statement of Corollary 1 may be chosen as a non-zero vector. If $\bar{s}$ is a local Pareto optimum and each of the problems in (D) is locally calm at $\bar{s}$, λ may be chosen as a strictly positive vector.

Calmness condition is interesting not only for its generality but also it leads to a proposition that most non-linear programming problems are calm. It would be desirable if a similar proposition may be proven with respect to maximizing n functions. Since the problems in (C) or (D) involve a local Pareto optimum $\bar{s}$ in the statement of the problem, "genericity" of the calmness condition does not follow in any obvious way in the case of maximizing n functions. We hope to answer this question in a future study.

III. Weak Local Pareto Optimum in a Non-Convex Economy and Marginal Cost Pricing Equilibrium

We now describe an economy with production where consumers' preferred sets or producers' productin sets may not be convex. We specify the conditions under which the attainable set of the economy is compact. This, together with the continuity of proferences, guarantees the existence of a weak Pareto optimal allocation. We then characterize the weak Pareto optimal allocation using Theorem 1. A result, sharper than Guesnerie's, is obtained under weaker conditions as a consequence.

X_i denotes the consumption set of consumer $i = 1,2,\ldots, n$. Y_j is the production set of producer $j = 1,2,\ldots, m$. There are ℓ commodities. X_i, Y_j are subsets of R^ℓ for all i, j. The set of *allocations*, C, is defined by $C = X^1 \times X^2 \times \ldots X_n \times Y_1 \times Y_2 \times \ldots \times Y_m$; $e \in R^\ell_+$ is the initial endowment for the economy. $s \in C$ is a *feasible allocation* if $g(s) \leqslant e$, where $g(s) = \sum_j x_i - \sum_j y_j$. The set of feasible allocations is denoted by F. $u_i : C \rightarrow R$ is the utility function of consumer i. Let $\pi_i : C \rightarrow X_i$ denote the projection map of C onto X_i. If u_i depends upon $x_i (= \pi_i(s))$ only, then u_i naturally induces a function $\tilde{u}_i : X_i \rightarrow R$ defined by $\tilde{u}_i(x_i) = u_i(s)$, for any $s \in \pi_i^{-1}(x_i)$. We shall use the following assumptions on the economy:

Assumption 1: (i) X_i is closed for each i.
Y_j is closed for each j.
(ii) $X_i \subset R^\ell_+$ for each i.
(iii) $(\sum_j Y_j) \supset -R^\ell_+$ (free disposability).
(iv) $A(\sum_j Y_j) \cap A(-\sum_j Y_j) = \{0\}$ (irreversibility).

Assumption 2: u_i: $C \rightarrow R$ is continuously differentiable, for each i. $D_{x_i} u_i(s) > 0$, for all i (weak monotonicity).

In Assumption 1, $A(\sum_j y_j)$ denotes the asymptotic cone of $\sum_j Y_j$. As usual, u_i is defined to be continuously differentiable on a closed set C if it may be extended to a continuously differentiable function on a neighborhood of C. $D_{x_i} u_i(s)$ denotes the derivative of u_i with respect to x_i.

Proposition 2: Under Assumtions 1 and 2, there exists a Pareto optimal allocation.

Proof: Under Assumpion 1, the feasible set of the economy, F, is compact (see Brown and Heal (4: Lemma 3)). Thus, the problem of maximizing $\sum_j u_i(s)$ subject to $s \in F$ has a solution. This solution is easily seen to be Pareto optimal. Q.E.D.

Brown and Heal uses a theorem of Hurwicz and Reiter (5), to show the compactness of F under Assumption 1. Hurwicz and Reiter defines the feasible set as $\{s \in c \mid g(s) = e\}$. But this definition of the feasible set is equivalent to our definition under the free disposability assumption (Assumption 1-(iii)).

Now, consider the following problem.

$$\begin{aligned} &\text{maximize } u_1, u_2, \ldots, u_n \\ &\text{subject to } g(s) \leqslant e \text{ and } s \in C. \end{aligned} \tag{E}$$

Suppose $\bar{s} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_m)$ is a weak local Pareto optimum of the above problem. By applying Corollary 1 and by observing that $N_C(\bar{s}) = N_{x_1}(\bar{x}_1) \times N_{x_2}(x_2) \times \cdots \times N_{x_n}(x_n) \times N_{Y_1}(y_1) \times \cdots \times N_{Y_m}(y_m)$, (this follows from the definition of the Clarke tangent cone and the definition of C), we obtain:

there exists $(\lambda, r) \in R_+^{n+\ell} - \{0\}$ such that

$$\sum_i \lambda_i D_{x_h} u_i(\bar{s}) - r \in N_{X_h}(\bar{x}_h) \text{ for } h=1, 2, \ldots, n,$$

$$(*) \quad \sum_i \lambda_i D_{y_f} u_i(\bar{s}) + r \in N_{Y_f}(\bar{y}_f) \text{ for } f=1, 2, \ldots, m,$$

$$\text{and } r_k \cdot g_k(\bar{s}) = 0,\ k=1, 2, \ldots, \ell.$$

These necessary conditions are quite general. We now specialize the situation to obtain more intuitive conditions:

Theorem 2: Assume A1 − (i) and A2 and assume that u_i depends upon x_i only. If $\bar{s} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_m)$ is a weak local Pareto optimum such that $\bar{x}_i \in \text{int } X_i$ for each i, then there are $\lambda >> 0$ and $r > 0$ such that $\lambda_i \cdot D\tilde{u}_i(\bar{x}_i) = r$ for $i = 1, 2, \ldots, n$,

$$r \in N_{Y_j}(\bar{y}_j) \text{ for } j=1, 2, \ldots, m, \text{ and}$$

$$r_k \cdot g_k(\bar{s}) = 0,\ k = 1, 2, \ldots, \ell.$$

Proof: If u_i depends only on x_i, then $D_{x_h} u_i(\bar{s}) = 0$ whenever $i \neq h$. Then, we may use $D\tilde{u}_i(\bar{x}_i)$ instead of $D_{x_i} u_i(\bar{s})$ for notational simplicity. If $\bar{x}_i \in \text{int } X_i$, $N_{X_i}(\bar{x}_i) = \{0\}$. Thus, the expression in Theorem 2 follow from (*). We know $(\lambda, r) \neq 0$. If $r > 0$ then $D\tilde{u}_i(\bar{x}_i) > 0$ implies $\lambda_i > 0$, all i. If $\lambda_i > 0$ for some i, again $D\tilde{u}_i(\bar{x}_i) > 0$ implies $r > 0$ and thus $\lambda >> 0$. Q.E.D.

We now formulate a theorem that may be considered a counterpart of the second welfare theorem in a non-convex economy. Given a utility allocation $u = (u_1, u_2, \ldots, u_n)$, we define:

$$X_i(u_i) = [x_i \in X_i \mid \tilde{u}_i(x_i) \geqslant u_i];\ i=1, 2, \ldots, n.$$

$$C(u) = X_1(u_1) \times X_2(u_2) \times \cdots \times X_n(u_n) \times Y_1 \times \cdots \times Y_n.$$

$X_i(u_i)$, $C(u)$ are closed sets.

If $\bar{s}$ is a weak local Pareto optimum, it is east to see that $\bar{s}$ is a weak local Pareto optimum with respect to the following problem:

$$\begin{aligned} &\text{maximize } u_1, u_2, \ldots, u_n \\ &\text{subject to } g(s) \leq e \text{ and } s \in C(\bar{u}), \text{ where} \\ &\bar{u} = (u_1(\bar{s}), u_2(\bar{s}), \ldots, u_n(\bar{s})). \end{aligned} \qquad \text{(F)}$$

Theorem 3: Assume A1 – (i) and A2 and assume that u_i depends upon x_i only. If $\bar{s} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_m)$ is a weak local Pareto optimum and if, for each i, there is $v \in T_{X_i(\bar{u}_i)}(\bar{x}_i)$ such that $D\tilde{u}_i(x_i) . v > 0$, then there exists $r > 0$ such that $-r \in N_{X_i(\bar{u}_i)}(\bar{x}_i)$ all i and $r \in N_{Y_j}(y_j)$, all j and $r_k . g_k(\bar{s}) = 0$, $k = 1, 2, \ldots, \ell$.

Proof: Necessary conditions for Pareto optimum $\bar{s}$ for problem F are: there exists $(\lambda, r) \in R_+^{n+\ell} - \{0\}$ such that $\lambda_i\, D\tilde{u}_i(\bar{x}_i) - r \in N_{X_i(\bar{u}_i)}(\bar{x}_i)$ $i = 1, 2, \ldots, n$, $r \in N_{Y_j}(y_j)$, $j = 1, 2, \ldots, m$, and $r_k . g_k(\bar{s}) = 0$, $k = 1, 2, \ldots, \ell$. We show $-D\tilde{u}_i(\bar{x}_i) \in N_{X_i(\bar{u}_i)}(\bar{x}_i)$, for all i. Take any $v \in T_{X_i(\bar{u}_i)}(\bar{x}_i)$. By the definition of the Clarke tangent cone, for each $\{x^\nu\} \subset X_i(\bar{u}_i)$ converging to $\bar{x}_i$ and each t^ν decreasing to 0, there is a sequence $\{v^\nu\}$ converging to v and $x^\nu + t^\nu \cdot v^\nu \in X_i(\bar{u}_i)$, all ν. Taking a constant sequence $x^\nu = \bar{x}_i$, all ν, we have

$$\frac{\tilde{u}_i(\bar{x}_i + t^\nu . v^\nu) - \tilde{u}_i(\bar{x}_i)}{t^\nu}$$

$$\frac{\tilde{u}_i(\bar{x}_i + t^\nu \cdot v) - \tilde{u}_i(\bar{x}_i)}{t^\nu} + \frac{\tilde{u}(\bar{x}_i + t^\nu \cdot v^\nu) - \tilde{u}(\bar{x}_i + t^\nu \cdot v)}{t^\nu} \geqslant 0$$

(Recall $\tilde{u}_i$ is defined at $\bar{x}_i + t^\nu \cdot v$ for all large enough ν, by the continuous differentiability of $\tilde{u}_i$). By the continuous differentiability of $\tilde{u}_i$, given any $\varepsilon > 0$ and for all ν large enough,

$$|\tilde{u}_i(\bar{x}_i + t^\nu \cdot v^\nu) - \tilde{u}_i(\bar{x}_i + t^\nu \cdot v) - D\tilde{u}_i(\bar{x}_i)\,(t^\nu \cdot (v^\nu - v))|$$

$\leqslant \varepsilon \cdot |t^\nu \cdot (v^\nu - v)|$. This implies that

$$\frac{\tilde{u}_i(\bar{x}_i + t^\nu \cdot v^\nu) - \tilde{u}_i(\bar{x}_i + t^\nu . v)}{t^\nu} \to 0 \text{ as } \nu \to \infty$$

and we have $D\tilde{u}_i(x_i) \cdot v \geqslant 0$. thus, $-D\tilde{u}_i(x) \in N_{X_i(\bar{u}_i)}(\bar{X}_i)$.

If $r = 0$, there exists i such that $\lambda_i > 0$, and $\lambda_i \cdot D\tilde{u}_i(\bar{x}_i)\, N_{X_i(\bar{u}_i)}(\bar{x}_i)$. This implies that $T_{X_i}(\bar{u}_i)(\bar{x}_i)$ is contained in a hyperspace normal to $D\tilde{u}_i(\bar{x}_i)$ and contradicts an assumption in Theorem 3. Q.E.D.

In the above theorem, r was chosen as a semi-strictly positive vector rather than simply a non-zero vector. This is the consequence of our assumptions of free disposability and $D\tilde{u}_i(x_i) > 0$. If we remove these assumptions, the theorem may be stated in terms of $r \neq 0$.

It would be instructive to compare Theorem 1 with the theorems in Guesnerie (1) in detail. This, in fact, is possible through the concepts and the results introduced in the first part of this paper. Such comparisons woul involve more detailed discussion of the Guesnerie's model than seems desirable here, however, will be postponed for a future occasion.

In the first part of this paper, we considered the calmness condition in the context of a mathematical programming. One may consider a class of mathematical programming problems in which the data enter in more general ways and ask whether a theorem such as Clarke's Theorem 2 may be proven in such situations. One may also consider the formulation of the counterpart of the calmness condition in the problem of maximizing n functions and its possible applications to economics.

The results in the second part of this paper may be extended in several directions. It is not difficult to show that continuous differentiability of utility functions may be replaced by local Lipschitz continuity to obtain analogous results. Also, one may introduce externalities. The constraint set C was given in the above as a product set of consumption and production sets. One may consider a more general set and still obtain useful necssary conditions for a Pareto optimum. It is very likely that the existence of tax-subsidy equilibrium supporting a Pareto optimum allocation may be derived from such necessary conditions in an environment more general than the one considered by Osana (6).

References

1. Guesnerie, R., "Pareto Optimality in Non-Convex Economics," *Econometrica,* Jan. 1975.
2. Clarke, F., *Optimization and Non-Smooth Analysis,* J. Wiley, 1983.
3. ________, "A New Approach to Lagrange Multipliers," *Mathematics of Operations Research,* May 1976.
4. Brown, D. and Heal, G., "Existence, Local Uniqueness, and Optimality of a Marginal Cost Pricing Equilibrium in an Economy with Increasing Returns," *Social Science Working Paper,* 415, California Institute of Technology, Jan. 1982.
5. Hurwicz, L. and Reiter S., "On the Boundedness of the Feasible Set Without Convexity Assumptions," *International Economic Review,* Vol. 14, 1973.
6. Osana, H., "Optimal Tax-Subsidy System for an Economy with Marshallian Externalities," *Econometrica,* March 1977.
7. Smale, S., "Global Analysis and Economics V—Pareto Theory With Constraints," *Journal of Mathematical Economics,* Vol. I, 1974.
8. Ekeland, I., "Non-Convex Minimization Problems," *Bulletin of the American Mathematical Society,* May 1979.
9. Hestenes, M., *Optimization Theory - The Finite Dimensional Case,* John Wiley, 1975.
10. Rockafellar, R. T., *The Theory of Subgradients and Its Applications to Problems of Optimization, Convex and Non-Convex Functions,* Heldermann Verlag Berlin, 1981.

11. Stoer, J. and Witzgall, C., *Convexity and Optimization in Finite Dimensions I*, Springer-Verlag, 1970.
12. Beato, P., "The Existence of Marginal Cost Pricing Equilibria with Increasing Returns," *Quarterly Journal of Economics*, Nov. 1982.
13. Beato, P. and Mas-Colell, A., "Marginal Cost Pricing and Aggregate Production Efficiency an Example," Working Paper, Universitat Autonoma de Barcelona, 1982.
14. Cornet, B., "Existence of Equilibria in Economies with Increasing Returns," Berkeley Paper No. IP 311, Aug. 1982.
15. Brown, D. and Heal, J., Khan, A. and Vohra, R., "On a General Existence Theorem for Marginal Cost Pricing Equilibria," Working Paper 83-15, School of Business, Columbia Univ., Jan. 1983.