Efficient, Pareto improving processes[1]

Kwan Koo YUN
Economics Department, SUNY at Albany, Albany, NY 12222
yun@albany.edu

April 17, 2004

[1] I would like to thank Robert Wilson, Kenneth Arrow, David Luenberger and Wanjin Kim for helpful comments.

Abstract

We give two procedures for determining whether efficient Pareto improving local changes are possible. When they are, the procedures compute for them. Any procedure generating efficient and Pareto improving changes can be replicated by these procedures. The two programs form a striking duality. We apply the procedures to Pareto improving exchange processes, Pareto-improving tariff-tax reforms and to the problem of constrained Pareto optimum where informational constraints are present.

keywords: Pareto improvement, constrained Pareto optimum
JEL Classification: C61, C78, D51, D82, H21

1 Introduction

Suppose that representatives from n countries negotiate piecemeal changes in multilateral tariffs. If they fail to reach an agreement, status quo prevails. A reasonable requirement of any agreement is that all participants benefit from the change. If larger tariff changes are more costly, whether economically or politically, it is desirable that such changes are efficient also. There are usually many efficient Pareto improving directions of changes when Pareto improving changes are possible. An equity criterion further narrows the choices. A large country may argue that it should benefit more than smaller countries. One way to express the equity considerations is through the share ratios of the total increase in aggregate welfare that goes to individuals. The MDP exchange process [4], [12] is such a process. Another way is through the exponents in the generalized Nash product of benefit increases[13].

If the share ratios of the aggregate welfare increase are prescribed, one can attempt to find an efficient local change that produces the share ratio. We modify a method in Yun [25] and use it to test whether there is an efficient and Pareto improving local change that generates the share ratios and to compute the direction when it exists. If positive exponents are chosen in the generalized Nash product, we maximize the product over Pareto improving changes of unit length. When a Pareto improvement is possible, the product is maximized at a unique efficient and Pareto improving direction.

We represent the feasible local changes by a non-empty, closed convex cone K in R^l and criteria by vectors $\{v_i\}_{i=1}^n$ in R^l. We interpret $\{v_i\}$ as the gradients of some criteria functions. A local change d in K is

improving if $v_i \cdot d > 0$, all i. When there is a Pareto improving change, $d \in K\backslash\{0\}$ is *efficient* if there does not exist $d' \in K$ of equal size ($|d'| = |d|$) such that $v_i \cdot d' > v_i \cdot d$, for all i.

Denote the standard unit simplex in R^n as $\Delta \equiv \{x \in R^n | x \geq 0, \sum x_i = 1\}$ and its relative interior as Δ° and the unit disk in R^l by $D \equiv \{d \in R^l | d \cdot d \leq 1\}$. Given v in R^l, $\pi_K\ v$ is the orthogonal projection of v to K.

Consider the following two problems:

Program 1. Given positive numbers $\{c_i\}$, $i = 1, \cdots, n$,

$$\min_{\lambda \in \Delta} \left| \pi_K \sum_i \lambda_i c_i v_i \right|. \tag{1}$$

Program 2. Given λ in Δ°,

$$\max_d \Pi_i\ (v_i \cdot d)^{\lambda_i}$$

$$\text{subject to } d \in K \cap D,\ v_i \cdot d \geq 0,\ i = 1, \cdots, n.$$

The two problems share criterion vectors $\{v_i\}$ and the feasible directions K.

The positive numbers $\{c_i\}$ specify the first problem and the positive exponents $\{\lambda_i\}$ specify the second. The minimum of program 1 or the maximum of program 2 is zero if and only if a Pareto improving change does not exist. A solution to either program gives a Pareto improving and efficient change when a Pareto improvement is possible. Conversely, given any Pareto improving and efficient local change, one can specify program 1 or Program 2 in such a way that the given direction solves the problems. The two programs form a striking duality. Suppose program 1 is defined with positive $\{c_i\}$,where $c = (c_1, \cdots, c_n)$ is positively proportional to $(\frac{1}{s_1}, \cdots, \frac{1}{s_n})$ for some $s \in \Delta^\circ$ and it has a strictly positive solution λ. If we define program 2 with the exponents $\{\lambda_i\}$, then the solution d yields share ratios s, i.e., $s_i = \frac{v_i \cdot d}{\sum_i v_i \cdot d}$, each i. Conversely, suppose program 2 is defined with positive exponents $\lambda \in \Delta$ and the solution d yields share ratios $\{s_i\}$. Then, if program 1 is defined with c, that is positively proportional to $(\frac{1}{s_1}, \cdots, \frac{1}{s_n})$, $\{\lambda_i\}$ is a solution to the program 1. In both cases, $\pi_K \sum_i \lambda_i \frac{1}{s_i} v_i$ given by the solution λ of the program 1 is positively proportional to the solution d of the program 2. When there is a single function f whose gradient is v, the duality reduces to the fact that f increases fastest in the direction of the projected gradient $\pi_K\ v$.

Figures 1 - 3 illustrate the above programs and their duality when $n = 2$, $K = R^2$ and $c_1 = c_2 = 1$. In Figure 1, two gradients $v_1 = a = (\frac{3}{4}, \frac{5}{4}),\ v_2 = b =$

$(\frac{7}{4}, \frac{1}{4})$ are shown. The vector $v = (1,1) = \frac{3}{4}a + \frac{1}{4}b$ solves the first program. Since v is orthogonal to $a - b$, $a \cdot v = b \cdot v > 0$. So, the share ratios are $s = (\frac{1}{2}, \frac{1}{2})$. The arc connecting e to f (excluding e or f) consists of vectors in the unit circle that form acute angles with both a and b. The inner product of the vectors on the arc with a and b respectively are plotted by the heavy curve in Figure 2. The part of the heavy curve in Figure 2 connecting A to B is 'efficient.' The efficient part is obtained by taking the vectors on the unit circle between a and b in Figure 1 and taking inner products with a and b. If we maximize $w = (a \cdot d)^{\frac{3}{4}}(b \cdot d)^{\frac{1}{4}}$ by choosing d among unit vectors, we reach the maximum at C in Figure 2 at $d = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. At this direction of change, $a \cdot d = b \cdot d = \sqrt{2}$ and the resulting share ratios are $s = (\frac{1}{2}, \frac{1}{2})$. Figure 3

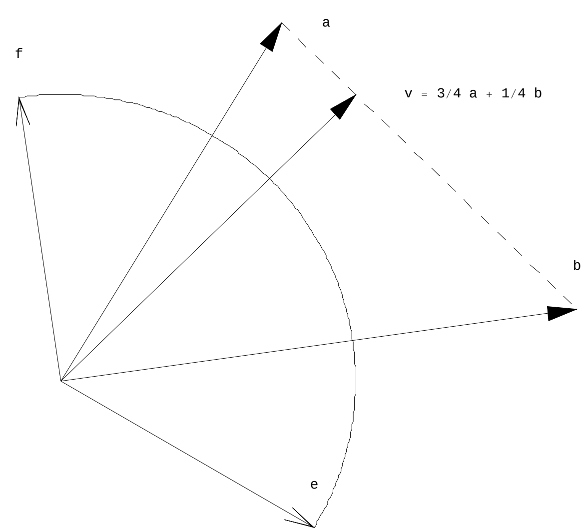

Figure 1:

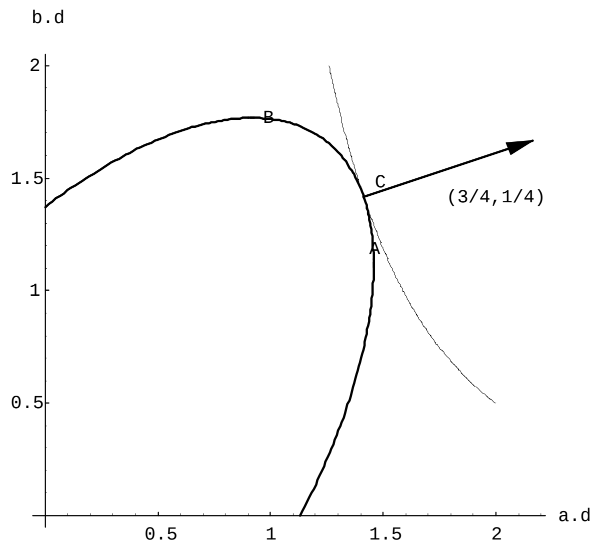

Figure 2:

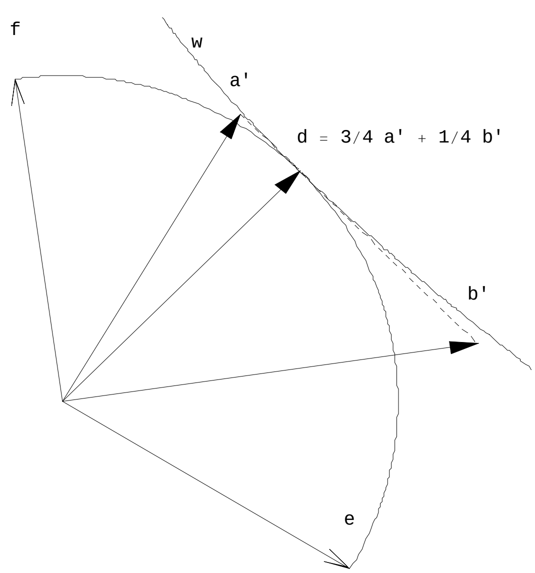

Figure 3:

transports Figure 2 to Figure 1. Here, the level curve of w corresponding to $d = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, $a' = \frac{1}{\sqrt{2}}a$ and $b' = \frac{1}{\sqrt{2}}b$ are shown. The gradient of $\ln w$ is equal to $d = \frac{3}{4}\frac{1}{\sqrt{2}}a + \frac{1}{4}\frac{1}{\sqrt{2}}b$. Also, $(\frac{1}{\sqrt{2}}a - \frac{1}{\sqrt{2}}b) \cdot d = 0$. From the geometry given by the previous two equations, $\lambda = (\frac{3}{4}, \frac{1}{4})$ minimizes $\left|\lambda_1 \frac{1}{\sqrt{2}}a + \lambda_2 \frac{1}{\sqrt{2}}b\right|$ over $\lambda \in \Delta$. Thus, when c is positively proportional to $(\frac{1}{s_1}, \frac{1}{s_2}) = (2, 2)$, $\lambda = (\frac{3}{4}, \frac{1}{4})$ solves program 1.

Dixit[5], Guesnerie[8],Turunen-Red and Woodland[21], study conditions under which Pareto improving piecemeal tariff and tax reforms exist. Weymark[20] shows that when the tangent cone K is a half space of R^l and $Df_i(x)$ was a gradient of f_i at x, $i = 1, \cdots, n$, a feasible direction is efficient if and only if it can be expressed as $\pi_K \sum_{i=1}^n \lambda_i Df_i(x)$ for some nonzero $\lambda \geq 0$. When there is one welfare function whose gradient is not zero, the gradient direction is the only efficient direction of change and is welfare improving. D'Aspremont and Tulkens[4], Tirole and Guesnerie [18] use the gradient of a weighted sum of welfare functions in studying exchange processes and tax reforms respectively. When there are more than one maximand, the gradient method gives an efficient direction but is not necessarily Pareto improving. Tulkens and Zamir[19] study local cooperative games with transferable utility in the context of dynamic exchange processes.

We examine what an efficient and Pareto improving exchange process looks like in general. By varying the equity criterion, we can generate all efficient and Pareto improving exchange processes. We show that the $M70$ process is efficient while the MDP process is not. We then give an example of a tariff reform. Our method not only checks the feasibility of a particular type of Pareto improving tariff reform but it also computes, whenever possible, the direction of an efficient and Pareto improving tariff reform corresponding to a choice of equity criterion.

Next, we apply our analysis to models where information constraints are present. In the standard principal-agent model, the principal has all the bargaining power and the agent solves an optimization problem within the principal's problem (the incentive constraint). When both parties have some bargaining power and when some constraints are informational, constrained Pareto optimum – not Pareto optimum nor the principal's solution – is the relevant concept. We characterize a constrained Pareto optimum and show how to find (constrained) Pareto improving directions when the current position is not a constrained Pareto optimum.

1.1 Mathematical Preliminaries

A subset A of R^l is a *cone* if $v \in A$, $\alpha > 0$ imply $\alpha v \in A$. For a non-empty subset S of R^l, its (negative) normal cone is $S^- = \{v \in R^l | v \cdot w \leq 0$, for all w in $S\}$. Given a non-empty closed set C in R^l and x in C, the tangent cone is $T_C(x) = \{v \in R^l |$ there is a sequence $\{x_i\}$ in C converging to x and a positive sequence $\{t_i\}$ decreasing to zero such that $(x_i - x)/t_i \to v\}$. The tangent cone is non-empty, closed. If C is convex, $T_C(x)$ is convex. We often write $T_C(x)^-$ as $N_C(x)$. Suppose C and D are closed convex sets in R^l and $x \in C \cap D$. If $T_C(x) - T_D(x) = R^l$, then $T_{C\cap D}(x) = T_C(x) \cap T_D(x)$ and $N_{C\cap D}(x) = N_C(x) + N_D(x)$. The condition $T_C(x) - T_D(x) = R^l$ is equivalent to $N_C(x) \cap -N_D(x) = \{0\}$([1],[24]).

Let K be a non-empty, closed convex cone in R^l. The orthogonal projection to K, π_K, maps v in R^l to the unique vector in K minimizing the Euclidean distance from v. For $v \in R^l$ and $c > 0$, $\pi_K cv = c\pi_K v$ and π_K is linear if K is a linear subspace. Given any $v \in R^l$, $\{v = w + z, w \cdot z = 0, w \in K, z \in K^-\}$ has a unique solution $\{w = \pi_K v,\ z = \pi_{K^-} v\}$. Thus, if $w = \pi_K v$, $v \cdot w = w \cdot w$ and for $d \in K$, $v \cdot d \leq w \cdot d$. Given $d \in K$, $u \in N_K(d)$ if and only if $u \in K^-$ and $u \cdot d = 0$. The functions $v \in R^n \mapsto |\pi_K v|$ and $\lambda \in \Delta \mapsto |\pi_K \sum_i \lambda_i v_i|$ are convex.

Given v, w in R^l, we write: $v \propto w$ if there is $c > 0$ such that $v = cw$; $v > 0$ if $v \geq 0$ and $v \neq 0$; $v >> 0$ if $v^i > 0$, all i; and $|v|$ is the Euclidean norm of v.

Let F be the cone of Pareto improving directions for some $\{v_i\}_{i=1}^n$ in R^l and λ a vector in Δ°. If F is not empty, $d \in F \to \ln(v_1 \cdot d)^{\lambda_1}(v_2 \cdot d)^{\lambda_2} \cdots (v_n \cdot d)^{\lambda_n}$ is a concave function. From this and the geometry of F, Program 2 has a unique solution when Pareto improving directions exist.

2 Main Results

Let K be a non-empty, closed convex cone of feasible changes in R^l. When there is a single objective function with the gradient v, a feasible change d satisfying $v \cdot d > 0$ does not exist if and only if $\pi_K\ v = 0$. When $\pi_K\ v$ is not zero, $\pi_K\ v$ is the efficient direction that improves the objective function. Theorem 3 extends this observation to the case of n objective function in terms of program 1 and program 2. Let $\{v_i\}$ be in R^l, $i = 1, \cdots, n$.

Lemma 1 *A Pareto improving change does not exist if and only if there is λ in Δ satisfying $\pi_K \sum_i \lambda_i v_i = 0$. A Pareto improving change $v \in K$ is efficient if and only if there is $\lambda > 0$ such that $v = \pi_K \sum_i \lambda_i v_i$.*

Proof. Suppose there is no d in K satisfying $v_i \cdot d > 0$, all i. Applying a necessary condition for a Pareto optimum [2][24], $\sum_i \lambda_i v_i \in K^-$ or equivalently, $\pi_K \sum_i \lambda_i v_i = 0$ for some λ in Δ. Conversely, suppose there is λ in Δ satisfying $\sum_i \lambda_i v_i \in K^-$. Since $\sum_i \lambda_i v_i \cdot d \leq 0$ for any d in K, there cannot be d in K satisfying $v_i \cdot d > 0$, all i. If $v \in K \backslash \{0\}$ is efficient, there is no u in K such that $|u| = |v|$ and $v_i \cdot u > v_i \cdot v$, all i. A necessary condition for this is: There is $(\lambda, \alpha) > 0$ such that $\sum_i \lambda_i v_i - \alpha v \in N_K(v)$ [2][24], or equivalently, $\sum_i \lambda_i v_i - \alpha v \in K^-$ and $(\sum_i \lambda_i v_i - \alpha v) \cdot v = 0$ (since K is a closed, convex cone). Thus, $v = \pi_K(\sum_i \lambda_i v_i - \alpha v + v)$. For any $\lambda > 0$, $\sum_i \lambda_i v_i \notin K^-$ since a Pareto improving change is possible. Since $N_K(v) \subset K^-$, $\sum_i \lambda_i v_i \notin N_K(v)$. We conclude that α cannot be zero. By choosing $\alpha = 1$, $v = \pi_K \sum_i \lambda_i v_i$, where $\lambda > 0$. Conversely, suppose $v = \pi_K \sum_i \lambda_i v_i$ for some $\lambda > 0$. Since $\sum_i \lambda_i v_i \notin K^-$, $v \neq 0$. The orthogonal decomposition of $\sum_i \lambda_i v_i$ gives $\sum_i \lambda_i v_i - v = u$, where $u \in K^-$ and $v \cdot u = 0$. Consider $w \in K$ with $|w| = |v|$. Since $(\sum_i \lambda_i v_i - v) \cdot v = 0$ and $(\sum_i \lambda_i v_i - v) \cdot w \leq 0$, $(\sum_i \lambda_i v_i - v) \cdot (w - v) \leq 0$. From $v \cdot v \geq v \cdot w$, $(\sum_i \lambda_i v_i) \cdot (w - v) \leq 0$. Thus, it is not possible that $v_i \cdot w > v_i \cdot v$, all i. ■

Since the definitions of a Pareto improving change and an efficient change are independent of the lengths of $\{v_i\}$, characterizations in Lemma 1 are independent of them.

Lemma 2 *Define program 1 with strictly positive* $\{c_i\}$. *Then,* $\lambda \in \Delta$ *is a solution to program 1 if and only if* $(*)$ $v = \pi_K \sum_i \lambda_i c_i v_i$ *satisfies* $c_i v_i \cdot v \geq v \cdot v$, *all* i *and* $c_i v_i \cdot v = v \cdot v$ *whenever* $\lambda_i > 0$. *Assume that a Pareto improving change exists and define program 2 with a strictly positive* $\lambda \in \Delta$. *Then,* d *in* K, *where* $|d| = 1$, *is the solution to program 2 if and only if* $(**)$ $v_i \cdot d > 0$ *for each* i *and* $d = \pi_K \sum_i \lambda_i \frac{1}{v_i \cdot d} v_i$.

Proof. Suppose that λ solves program 1 and let $w = \sum_i \lambda_i c_i v_i$ and $v = \pi_K\, w$. Then, $w = v + u$, where $u \in K^-$ and $v \cdot u = 0$. Since $|v| \leq |\pi_K \ \sum_i \lambda'_i c_i v_i|$ for any $\lambda' \in \Delta$, $|v| \leq |\pi_K\ (w + t(c_i v_i - w))| \leq |w + t(c_i v_i - w)) - u| = |v + t(c_i v_i - w))|$ for each i and $t \in [0, 1]$. The second inequality follows since $|\pi_K\ (w + t(c_i v_i - w))|$ is the minimum distance from $w + t(c_i v_i - w)$ to K^- and u is in K^-. Thus, $\frac{d}{dt}(v + t(c_i v_i - w)) \cdot (v + t(c_i v_i - w))\,|_{t=0} = 2v \cdot (c_i v_i - w) \geq 0$ for each i. We have: $c_i v_i \cdot v \geq w \cdot v = v \cdot v$, all i. Since $\sum_i \lambda_i c_i v_i \cdot v = w \cdot v = \sum_i \lambda_i v \cdot v$, $\lambda_i c_i v_i \cdot v = \lambda_i v \cdot v$, all i and $c_i v_i \cdot v = v \cdot v$ if $\lambda_i > 0$. Conversely, suppose $(*)$ holds and let $v' = \pi_K \sum_i \lambda'_i c_i v_i$, for some $\lambda' \in \Delta$. There are u, u' in K^- such that $v = \sum_i \lambda_i c_i v_i - u$, $u \cdot v = 0$ and $v' = \sum_i \lambda'_i c_i v_i - u'$, $u' \cdot v' = 0$. Since u minimizes the distance from $\sum_i \lambda_i c_i v_i$ to K^- and $[u', u] \subset K^-$, $\frac{d}{dt}(t(u'-u)-v)\cdot(t(u'-u)-v)\,|_{t=0} = -v\cdot(u'-u) \geq 0$. Then, $v \cdot v' = v \cdot (-u' + \sum_i \lambda'_i c_i v_i) \geq v \cdot (-u + \sum_i \lambda'_i c_i v_i) \geq v \cdot v$. The second inequality is from $c_i v_i \cdot v \geq v \cdot v$ and $v \cdot u = 0$. Thus, $|v|\,|v'| \geq v \cdot v' \geq v \cdot v = |v| \cdot |v|$, implying $|v'| \geq |v|$.

Next, suppose that d solves program 2. Since Pareto improvement is possible, $v_i \cdot d > 0$, all i. Since for $\alpha > 0$, $\Pi_i\ (v_i \cdot \alpha d)^{\lambda_i} = \Pi_i\ \alpha(v_i \cdot d)^{\lambda_i}$, $|d| = 1$. Taking the log of the maximand, a necessary condition for optimum is $\sum_i \lambda_i \frac{1}{v_i \cdot d} v_i \in N_C(d)$, where $C = K \cap D$. We show that $N_C(d) = N_D(d) + N_K(d)$. It is sufficient to show $T_D(d) - T_K(d) = R^l$, or equivalently, $N_D(d) \cap -N_K(d) = \{0\}$. Now, $N_D(d) = \{cd | c \geq 0\}$ and $N_K(d) = \{u | u \in K^-, u \cdot d =$

$0\}$. If $cd = -u$ for $u \in N_K(d)$ and $c \geq 0$, then $cd \cdot u = -u \cdot u = 0$, implying $u = 0$. Thus, $\sum_i \lambda_i \frac{1}{v_i \cdot d} v_i = \gamma d + u$ where $\gamma \geq 0$, $u \in K^-$ and $u \cdot d = 0$. Taking an inner product of d with both sides of the equation, $\gamma = 1$. Thus, $d = \pi_K \sum_i \lambda_i \frac{1}{v_i \cdot d} v_i$. Conversely, suppose $(**)$ holds and consider a Pareto improving $d' \in K \cap D$. Using the concavity of the maximand and writing $\sum_i \lambda_i \frac{1}{v_i \cdot d} v_i = d + u$, where $u \in K^-$ and $u \cdot d = 0$, $\sum_i \lambda_i \ln v_i \cdot d' - \sum_i \lambda_i \ln v_i \cdot d \leq \sum_i \lambda_i \frac{1}{v_i \cdot d} v_i \cdot (d' - d) = (d+u) \cdot (d' - d) \leq d \cdot (d' - d) \leq 0$. The last two inequalities follow from $u \cdot d' \leq 0$, $u \cdot d = 0$ and from $|d'| = |d|$. ■

Combining Lemma 1 and Lemma 2:

Theorem 3 *Let $v = \pi_K \sum_i \lambda_i c_i v_i$, where λ is a solution of program 1. Then, $v = 0$ if and only if a Pareto improving change does not exist. If $v \neq 0$, it is efficient and Pareto improving. If the maximum value at a solution d to Program 2 is positive, d is non-zero, efficient and Pareto improving. If the maximum at a solution d to Program 2 is zero, a Pareto improving direction does not exist and the zero vector is a solution.*

Proof. From Lemma 1, $v = 0$ if and only if Pareto improving direction does not exist and v is efficient if Pareto improving direction exists. From Lemma 2, v is Pareto improving since $c_i v_i \cdot v \geq v \cdot v > 0$, all i. Suppose the maximum value of Program 2 is positive at a solution d. If there were $d' \in K$ such that $|d'| = |d|$ and $v_i \cdot d' > v_i \cdot d$, all i, d would not maximize the generalized Nash product over the feasible set. Thus, d is efficient. The rest is clear. ■

Although Program 1 may not have a unique solution, the corresponding direction $\pi_K \sum_i \lambda_i c_i v_i$ is unique.

The cone of feasible changes K need not be all economically feasible directions but rather those the policy makers choose to restrict themselves to. For example, the policy makers may test whether proportional reductions of tariffs would increase welfare of all parties (Yun[23]).

A converse of Theorem 3 is:

Theorem 4 *A non-zero $v \in K$ is an efficient, Pareto improving direction if and only if there are positive numbers $\{c_i\}$, $i = 1, \cdots, n$ and λ in Δ satisfying $v = \pi_K \sum_i \lambda_i c_i v_i$ and $c_i v_i \cdot v = c_j v_j \cdot v$, all i, j. Program 1, defined with $\{c_i\}$, has λ as a solution and program 2, if defined with a strictly positive λ, has $d = \frac{v}{|v|}$ as its unique solution.*

Proof. Suppose that $v \in K$ is an efficient, Pareto improving direction. By Lemma 1, there is $\alpha \in \Delta$ such that $v \propto w = \pi_K \sum_i \alpha_i v_i$. We only consider the case of when $v = w$. There are $u \in K^-$, $u \cdot v = 0$ such that $v = -u + \sum_i \alpha_i v_i$. Choose positive numbers $\{c_i\}$ such that $c_i v_i \cdot v = c_j v_j \cdot v$, all i, j and $\sum_i \frac{\alpha_i}{c_i} = 1$ (for each i, choose $a_i = \frac{v_1 \cdot v}{v_i \cdot v}$ and let $c_i = k a_i$ where $k = \sum_i \frac{\alpha_i}{a_i}$). Then, $v = -u + \sum_i \lambda_i c_i v_i$, where $\lambda_i \equiv \frac{\alpha_i}{c_i}$ for each i and $\lambda \in \Delta$.

The converse follows from Lemma 1 and $v \cdot v = \sum_i \lambda_i c_i v_i \cdot v = c_i v_i \cdot v > 0$. By Lemma 2, λ is a solution of the program 1 defined with $\{c_i\}$. Next, let $d = \frac{v}{|v|}$. Then, $d = -\frac{u}{|v|} + \sum_i \lambda_i \frac{c_i}{|v|} v_i \in K$, where $|d| = 1$ and $c_i v_i \cdot d = c_j v_j \cdot d$, all i, j. Taking an inner product of d with both sides of the equation, $1 = \frac{c_i}{|v|} v_i \cdot d$ for each i. Thus, $d = \pi_K \sum_i \lambda_i \frac{1}{v_i \cdot d} v_i$. By Lemma 2, d is the solution of program 2 defined with λ. ■

In a two person bargaining problem, a utility allocation is a Nash solution if and only if there is an affine transformation of utilities such that the solution in the new units is simultaneously the egalitarian and the utilitarian solution. Myerson[13][16] considers a generalized Nash problem where individuals may carry different weights. In the present context, we have:

Corollary 5 *A feasible direction d of unit length is efficient and Pareto improving if and only if there are positive numbers $\{c_i\}$, $i = 1, \cdots, n$ and λ in Δ such that d maximizes $\sum_i \lambda_i c_i \cdot v$ over $v \in K \cap D$ and $c_i v_i \cdot d = c_j v_j \cdot d$, for each i, j.*

Proof. By Theorem 4, d in $K \backslash \{0\}$ is efficient and Pareto improving if and only if there are positive $\{c_i\}$ such that $d = \pi_K \sum_i \lambda_i c_i v_i$ and $c_i v_i \cdot d = c_j v_j \cdot d$ for each i, j. For any $v \in K \cap D$, $\sum_i \lambda_i c_i v_i \cdot d = d \cdot d = 1 \geq d \cdot v \geq \sum_i \lambda_i c_i v_i \cdot v$. The converse follows since $1 > d \cdot v$ for $v \neq d$ in the previous expression and thus d is the unique maximzer of $\sum_i \lambda_i c_i v_i \cdot v$ over $v \in K \cap D$. ■

2.1 Share ratios of aggregate welfare improvements

A Pareto improving change d generates a *share ratio* s in Δ where $s_i = \frac{v_i \cdot d}{\sum_i v_i \cdot d}$ for each i. When v_i is the domestic price vector for the economy i in terms of commodity l, $v_i \cdot d$ represents the rate of increase of welfare measured in commodity l in the economy. Given a strictly positive s in Δ, one can look for an efficient, Pareto improving change d generating the share ratio. If such a change exists, it can be found by minimizing $\left|\pi_K \sum \lambda_i \frac{1}{s_i} v_i\right|$ over $\lambda \in \Delta$. The following feasibility condition shows when an efficient, Pareto improving direction d can generate a strictly positive share ratios s.

$$\text{(a) } d \propto v = \pi_K \sum_i \frac{\lambda_i}{s_i} v_i, \text{ for an } \lambda \text{ in } \Delta. \text{ (b) } v_i \cdot v = s_i v \cdot v, \text{all } i. \tag{Condition F}$$

Condition F shows that the share ratio s_i is the Fourier coefficients $\frac{v_i \cdot v}{v \cdot v}$ in projecting v_i in the direction of v.

Lemma 6 *An efficient, Pareto improving direction d generates a strictly positive share ratios s if and only if Condition F is satisfied.*

Proof. We may replace the conditions in Theorem 4 by (a′) $v = \pi_K \sum_i \lambda_i c_i v_i$ and (b′) $v_i \cdot v = c_i v \cdot v$, all i since $v \cdot v = \sum_i \lambda_i c_i v_i = c_i v_i \cdot v$. A non-zero d satisfying Condition F satisfies (a′), (b′) with $c_i = \frac{1}{s_i}$ and generates s since $\sum_i v_i \cdot v = v \cdot v$. Conversely, if a non-zero d satisfies (a′) and (b′) and generates

s, then d satisfies Condition F with $d = kv$ and $s_i = k\frac{1}{c_i}$, where $k = 1/\sum_i \frac{1}{c_i}$.

■

The following duality theorem now follows from Theorem 4 and Lemma 6.

Theorem 7 *If Program* 1 *is defined with* $\{\frac{1}{s_i}\}$ *for some* s *in* Δ° *and has a solution* λ *in* Δ° *and* $\pi_K \sum_i \lambda_i \frac{1}{s_i} v_i \neq 0$, *then Problem* 2, *defined with* λ, *has a unique solution* d *that generates the share ratio* s. *Conversely, if Program 2 is defined with* λ *in* Δ° *and has a Pareto improving solution* d *that generates the share ratio* s, *then Problem 1, defined with* $\{\frac{1}{s_i}\}$, *has a solution* λ. *In both cases,* $d \propto \pi_K \sum_i \lambda_i \frac{1}{s_i} v$.

Proof. Suppose that program 1 is defined with $\{\frac{1}{s_i}\}$ for some s in Δ° and has a solution λ in Δ° and $v = \pi_K \sum_i \lambda_i \frac{1}{s_i} v_i \neq 0$. By Lemma 2, $v_i \cdot v = s_i v \cdot v$, all i. By Theorem 4, program 2 defined with λ has $d \propto v$ as the solution and thus generates s. If program 2 is defined with λ in Δ° and has a Pareto improving solution d, then by Lemma 2, $d = \pi_K \sum_i \lambda_i \frac{1}{v_i \cdot d} v_i$.where $v_i \cdot d > 0$,all i. Then, $v = \pi_K \sum_i \lambda_i \frac{1}{s_i} v_i = \left(\sum_i v_i \cdot d\right) d$. Since $|d| = 1$, $v \cdot v = \left(\sum_i v_i \cdot d\right)^2 = \frac{1}{s_i} v_i \cdot v$. By Theorem 4, program 1, defined with s, has a solution λ. ■

Schematically:

$$\begin{array}{lll} & \text{program 1} & \text{program 2} \\ \text{given} & s\ (c_i = \frac{1}{s_i}) & \lambda \\ \text{determines} & \lambda & s \\ \text{direction of change} & v = \pi_K \sum_i \lambda_i \frac{1}{s_i} v_i & d = \pi_K \sum_i \lambda_i \frac{1}{v_i \cdot d} v_i \propto v \end{array} \tag{2}$$

A Pareto improving direction given by Program 2 is invariant with respect to the lengths of $\{v_i\}$. This is clear since for any positive numbers $\{c_i\}$, $\Pi_i (c_i v_i \cdot d)^{\lambda_i} = c \Pi_i (v_i \cdot d)^{\lambda_i}$, where $c = \Pi_i c_i^{\lambda_i} > 0$. The invariance is important since the lengths of $\{v_i\}$ may not be economically meaningful. Program 1, by contrast, picks different Pareto improving directions when the lengths of $\{v_i\}$ are altered through multiplicative factors $\{\frac{1}{s_i}\}$. A duality between the procedures shows, however, that they are in a sense equivalent.

3 Applications and Examples

3.1 Exchange Processes

Consider an exchange economy with n traders and $l(\geq 2)$ goods. The consumption x_i of trader i and the aggregate endowment ω are strictly positive vectors in R^l. The consumption allocations are $\{x = (x_1, \cdots, x_n) | x >>$

$0, \sum_i x_i = \omega\}$. The utility function of trader i is $u_i(x) \equiv U_i(x_i)$, whose gradients are strictly positive vectors v_i and V_i. Then, $v_i = (0, \cdots, 0, V_i, 0, \cdots, 0)$, where V_i is the i_{th} l-vector of v_i. We normalize: $|V_i| = 1$, all i. The feasible directions at a consumption allocation x is $K = \left\{z = (z_1, \cdots, z_n) \,|z_i \in R^l, \text{ all } i \text{ and } \sum_i z_i = 0\right\}$ whose normal is $K^- = \left\{(r, \cdots, r) \in R^{nl} | r \in R^l\right\}$. Let $z = \pi_K \sum_i \lambda_i \frac{1}{s_i} v_i$, $\lambda \in \Delta$. By the orthogonal decomposition, $\sum_i \lambda_i \frac{1}{s_i} v_i = (\lambda_1 \frac{1}{s_1} V_1, \cdots, \lambda_i \frac{1}{s_i} V_i, \cdots, \lambda_n \frac{1}{s_n} V_n) = (z_1, \cdots, z_n) + (r, \cdots, r)$. From this, $r = \frac{1}{n} \sum_i \lambda_i \frac{1}{s_i} V_i$. For $i = 1, \cdots, n$,

$$z_i = \frac{\lambda_i}{s_i} V_i - \frac{1}{n} \sum_j \frac{\lambda_j}{s_j} V_j. \tag{3}$$

Total gain is: $\sum_i V_i \cdot z_i = \sum_j \frac{\lambda_j}{s_j} - \overline{V} \cdot \sum_j \frac{\lambda_j}{s_j} V_j$, where $\overline{V} \equiv \sum_j \frac{1}{n} V_j$. For each i,

$$\frac{\lambda_i}{s_i} - \frac{1}{n} V_i \cdot \sum_j \frac{\lambda_j}{s_j} V_j = s_i \left(\sum_j \frac{\lambda_j}{s_j} - \overline{V} \cdot \sum_j \frac{\lambda_j}{s_j} V_j \right). \tag{4}$$

3.1.1 exchange processes using program 1

Consider first the case of equal division of the aggregate gain; $s_i = \frac{1}{n}$, each i. By choosing $d = n^2 z$ and denoting $\widetilde{V} \equiv \sum_i \lambda_i V_i$, (3) and (4) become $d_i = n\lambda_i V_i - \widetilde{V}$ and $\lambda_i = \frac{1}{n}\left(1 + \left(V_i - \overline{V}\right) \cdot \widetilde{V}\right)$, $i = 1, \cdots, n$. From the Brouwer fixed point theorem, there is λ in Δ solving the latter equations.

Using a solution λ to evaluate $\widetilde{V}$,

$$d_i = \left(1 + \left(V_i - \overline{V}\right) \cdot \widetilde{V}\right) V_i - \widetilde{V},\ i = 1, \cdots, n \tag{5}$$

This process is Pareto improving (generates share ratios $s_i = \frac{1}{n}$, each i) whenever $d \neq 0$; that is, whenever $V_i \neq V_j$ for some i, j. In the followings, we assume that this is the case.

We present an efficient and Pareto improving process where one needs not solve for λ at each step. The process may be considered an approximation of the above process. Approximate $\widetilde{V}$ by $\overline{V}$ and assign values $\lambda_i = \frac{1}{n}\left(1 + \left(V_i - \overline{V}\right) \cdot \overline{V}\right)$ for each i. Then the process is given by (5). Since $\overline{V} \cdot \overline{V} < 1$ and $\{V_i\}$ are non-negative, $\lambda_i > 0$, all i. Also, $\sum_i \lambda_i = 1$ and $\sum d_i = 0$. This process is efficient since $d \propto \pi_K \sum \lambda_i v_i$. It is Pareto improving since $nV_i \cdot d_i = (1 - \overline{V} \cdot \overline{V})(1 - V_i \cdot \overline{V}) + \frac{1}{n}\left(\sum_j (1 - V_j \cdot \overline{V}) V_i \cdot V_j\right) > 0$. The inequality follows from $V_i \cdot \overline{V} \leq \left|\frac{1}{n} \sum_j V_j\right| < 1$.

The MDP process [12] is an exchange process that can generate any prescribed share ratios $\{s_i\}$. The price for such a flexibility is that the process is not efficient in general. Consider the case of two traders and three goods. Normalize the gradient of U_i by: $\pi_i(x_i) \equiv DU_i(x_i)/D_l U_i(x_i)$ and let $\overline{\pi}_i = (\pi_i^1, \pi_i^2)$, $\overline{\pi} \equiv \sum_i s_i \overline{\pi}_i$ and $\overline{d}_i = (d_i^1, d_i^2)$. According to the MDP process, $\overline{d}_i = s_i \Lambda(\overline{\pi}_i - \overline{\pi})$ and $d_i^3 = s_i w - \overline{\pi}_i \cdot \overline{d}_i$, $i = 1, 2$ where Λ is a diagonal matrix with positive entries and $w \equiv \sum_i \overline{\pi}_i \cdot \overline{d}_i$. Collecting the terms, $\overline{d}_1 = s_1 s_2 \Lambda(\overline{\pi}_1 - \overline{\pi}_2)$, $d_1^3 = -s_1 s_2 (s_2 \overline{\pi}_1 + s_1 \overline{\pi}_2)^t \Lambda(\overline{\pi}_1 - \overline{\pi}_2)$, $\overline{d}_2 = -\overline{d}_1$, $d_2^3 = -d_1^3$. The aggregate welfare increase, $w = s_1 s_2 (\overline{\pi}_1 - \overline{\pi}_2)^t \Lambda(\overline{\pi}_1 - \overline{\pi}_2)$ is positive unless $\overline{\pi}_1 = \overline{\pi}_2$. We give an example where Λ is the identity matrix and $U_1 = x_1 y_1 z_1$, $U_2 = x_2 y_2 z_2$. Holdings of traders are: $x_1 = (1, 3, 3)$, $x_2 = (3, 1, 1)$ and $s_1 = s_2 = \frac{1}{2}$. The MDP process gives the rate of welfare change $(.888, .888)$. Process 5, with the same norm of the rate of exchanges, gives the rate of welfare change $(1.1178, 1.1178)$.

3.1.2 exchange processes using program 2

Here, it is convenient to work with $d = \pi_K \sum_i \lambda_i \frac{1}{v_i \cdot d} v_i$ (Lemma 2). Suppose $\lambda \in \Delta^\circ$ is given. Following the arguments leading to (3), $d_i = \lambda_i c_i V_i - \frac{1}{n} \sum_j \lambda_j c_j V_j$, where $c_i = \frac{1}{V_i \cdot d_i}$. Multiplying both sides of the equation with V_i and writing $\widehat{V} = \sum_j \lambda_j c_j V_j$, $1 = c_i(\lambda_i c_i - \frac{1}{n} V_i \cdot \widehat{V})$, for $i = 1, \cdots, n$. After solving for positive numbers $\{c_i\}_{i=1}^n$ from these n equations, an efficient, Pareto improving change is given by: $d_i = \lambda_i c_i V_i - \frac{1}{n} \sum_j \lambda_j c_j V_j$.

Can we choose λ such that $\lambda_i c_i = \lambda_j c_j$ all i, j. In that case, $d_i \propto V_i - \frac{1}{n} \sum_j V_j$, all i. Such a choice is indeed possible ($\lambda_i = \frac{1 - V_i \cdot \overline{V}}{n(1 - \overline{V} \cdot \overline{V})}$, all i). Thus, this simple process, called the M70 process ([4]), is efficient and Pareto improving.

3.2 A Numerical Example of Tariff Changes

Consider an exchange international economy with two goods and two countries. Good 2 is a numeraire and p_1 is the world price of good 1 in terms of good 2. The consumption and the endowment of country i are $x_i = (x_i^1, x_i^2)$ and $\omega_i = (\omega_i^1, \omega_i^2)$. Welfare functions are: $U_1 = \sqrt{x_1^1 x_1^2}$ and $U_2 = \sqrt{x_2^1 x_2^2}$. Countries may impose specific tariffs on good 1 only. The tariff by country i is $t_i = (\tau_i, 0)$. Negative τ_i is subsidy. Domestic price vector in country i is $q_i \equiv p + t_i$. Equilibrium conditions are: $DU_i(x_i) \propto q_i$, $p \cdot (x_i - \omega_i) = 0$, $i = 1, 2$ and $\sum_i x_i = \sum_i \omega_i$.

Endowments are: $\{\omega_1 = (3, 1), \omega_2 = (1, 3)\}$. At a tariff profile of $(\tau_1, \tau_2) = (-0.05, 0.05)$, the equilibrium consumption allocations are $\{x_1 \approx (2.30, 1.48), x_2 \approx (1.70, 2.52)\}$ and p_1 is approximately 0.69. At the tariff equilibrium, country 1 exports good 1 and subsidizes its export. Country 2 imports good 1 and charges a tariff on it. At the equilibrium, we compute an efficient and Pareto improving change in the allocation space. Turunen-Red and Woodland[**?**] uses the Motzkin's Lemma to test whether a particular Pareto improving tariff reform exists. Our method not only tests the possibility but also computes for such a reform whenever they exist. The feasible allocations are $S = \{(x_1, x_2) | x_1 \geq 0, x_2 \geq 0, x_1 + x_2 = \omega_1 + \omega_2\}$. Given a strictly positive $x = (x_1, x_2)$, the tangent cone $T_S(x)$ is $K = \{d = (d_1, d_2) | d_1 = -d_2\}$.

Defining $u_i(x) \equiv U_i(x_i)$, $i = 1, 2$, let $v_1 = (DU_1(x_1), 0) \propto (q_1, 0)$ and $v_2 = (0, DU_2(x_2)) \propto (0, q_2)$. For any choice of strictly positive λ_1, λ_2 summing to 1, maximizing $\ln(v_1 \cdot d)^{\lambda_1}(v_2 \cdot d)^{\lambda_2}$ subject to $d \in K$, $|d| \leq 1$ is equivalent to maximizing $\ln(q_1 \cdot d_1)^{\lambda_1}(-q_2 \cdot d_1)^{\lambda_2}$ subject to $|d_1| \leq \frac{1}{\sqrt{2}}$. A necessary and sufficient condition for maximum is $\lambda_1 \frac{1/2}{q_1 \cdot d_1} q_1 + \lambda_2 \frac{1/2}{q_2 \cdot d_1} q_2 = d_1$. When $\{\lambda_1 = \frac{3}{4}$, $\lambda_2 = \frac{1}{4}\}$, d_1 is approximately $(-0.575, 0.412)$. In this computation, we just used the knowledge of q_1, q_2.

We can ask what tariff changes will produce the change d. This involves a general equilibrium computation that requires the knowledge of preferences around the current tariff equilibrium. Instead, we consider the welfare as a function of a tariff profile and compute an efficient and Pareto improving tariff changes. Writing x_i as a function of $\tau = (\tau_1, \tau_2)$, we redefine $v_i \equiv DU_i(x_i) Dx_i(\tau) \big|_{\tau=(-0.05,0.05)}$, $i = 1, 2$. We compute: $v_1 \approx (-0.210, -0.238), v_2 \approx (0.159, 0.115)$. Since we do not impose any restriction on the changes of tariffs, the tangent cone of feasible directions is R^2 and its normal cone is $\{0\}$. Maximizing $\ln(v_1 \cdot d)^{3/4}(v_2 \cdot d)^{1/4}$ subject to $d \cdot d \leq 1$ and $v_i \cdot d \geq 0$, $i = 1, 2$, the unique solution is $d \approx (0.631, -0.776)$. The welfare improvements are:$\{v_1 \cdot d \approx 0.52, v_2 \cdot d \approx 0.11\}$.

Now, we choose $c_1 = \frac{1}{v_1 \cdot d}, c_2 = \frac{1}{v_2 \cdot d}$ and minimize $|\lambda_1 c_1 v_1 + \lambda_2 c_2 v_2|$ over λ in Δ. The solution is $\lambda = (\frac{3}{4}, \frac{1}{4})$. These are precisely the weights used in the generalized Nash product.

3.3 Pareto improvement when informational constraints are present

In the principal-agent framework, the principal has all the bargaining power (see Grossman-Hart[[6]], Rogerson[[15]], Jewitt[[10]], for example). We may ask what the Pareto optimal solutions are when both parties have some bargaining power. Also, given a wage schedule w and associated action choice a by the agent, the principal can consider a piecemeal changes to w that produce Pareto improvement for both parties. We adopt Rogerson's notation closely. With a finite possible outcomes for an effort, the model is finite dimensional. Let $a \geq 0$ and $w_j \geq 0$, for $j = 1, \cdots, n$. The expected utility of the principal is given by: $U(w,a) = \sum_{j=1}^{n} p_j(a)u(x_j - w_j)$ and that of the agent by: $V(w,a) = \sum_{j=1}^{n} p_j(a)(v(w_j) - a)$. Here, $p_j(a)$ is the probability that the j_{th} outcome will occur when action a is chosen and x_j is the gross income to the principal in the event of the j_{th} outcome. When the expected utilities are differentiable, $DU(w,a) = (-p_1(a)u'(x_1 - w_1), \cdots, -p_n(a)u'(x_n - w_n), \sum_{j=1}^{n} p'_j(a)u(x_j - w_j))$ and $DV(w,a) = (p_1(a)v'(w_1), \cdots, p_n(a)v'(w_n), \sum_{j=1}^{n} p'_j(a)v(w_j) - 1)$. We say that $(w,a) \geq 0$ is Pareto optimal if there is no $(w',a') \geq 0$ such that $U(w',a') > U(w,a)$ and $V(w',a') > V(w,a)$. If $(w,a) >> 0$ is Pareto optimal, there is $\lambda \in [0,1]$ such that $\lambda DV(w,a) + (1-\lambda)\, DU(w,a) = 0$. That is, $\lambda v'(w_i) = (1-\lambda)\, u'(x_i - w_i)$, all

i and $\sum_{j=1}^{n} p_j'(a)\,[\lambda v(w_j) + (1-\lambda)\,u(x_j - w_j)] = \lambda$. Now, suppose that (w, a) needs to satisfy the participation constraint $((w, a) \in C_1 \equiv \{(w, a)|V(w, a) \geq \overline{V}\}$ and an incentive constraint.$((w, a) \in C_2 \equiv \{(w, a)|a \in \arg\max V(w, \cdot)\}$. (We need to introduce participation constraint for the principal, as well. We do not do so here for simplicity). While the wage schedule w needs to be agreed upon by both parties, the action a is chosen by the agent. The idea of constrained Pareto optimum repects this fact. We say that $(w, a) \geq 0, (w, a) \in C \equiv C_1 \cap C_2$ is *constrained Pareto optimal* if there is no $(w', a') \geq 0$, $(w', a') \in C_1 \cap C_2$ such that $U(w', a') > U(w, a)$ and $V(w', a') > V(w, a)$. A necessary condition that (w, a) is a constrained Pareto optimum is: there is $\lambda \in [0, 1]$ such that $\lambda DV(w, a) + (1-\lambda)DU(w, a) \in N_C(w, a)$. A constrained Pareto optimal (w, a) at which the participation constraint C_1 is binding is the solution for the principal in the standard principal-agent model. If the current allocation is not constrained Pareto optimal, we can consider a constrained Pareto improvement in w. That is, how can the wage schedule w be changed so that the welfare of both the principal and the agent improve while satisfying the incentive constraint?

For a constrained Pareto optimum that is different from a solution to the principal's problem, consider the case where the participation constraint is not binding. We consider a simple case that the optimum a for the agent is locally given by a differentiable function f; that is, $a = f(w)$, locally. Assuming $Df(w) \neq 0$, $T_C(w,a) = \{(dw, Df(w)dw) | dw \in R^n\}$ and $N_C(w,a) = \{(-cDf(w), c) | c \in R\}$. So, a necessary condition that (w,a) is a constrained Pareto optimum is that there is $\lambda \in [0,1]$ such that $\lambda DV(w,a) + (1-\lambda)DU(w,a) = (-cDf(w), c)$ for some $c \in R$. Writing out: there is $\lambda \in [0,1]$ and $c \in R$ such that $(p_1(a)(\lambda v'(w_1) - (1-\lambda)u'(x_1 - w_1)), \cdots, p_n(a)(\lambda v'(w_n) - (1-\lambda)u'(x_n - w$ $-cDf(w)$ and $\sum_{j=1}^{n} p_j'(a)[\lambda v(w_j) + (1-\lambda)u(x_j - w_j)] - \lambda = c$. If (w,a) is not a constrained Pareto optimum, we can compute directions of change in w that would Pareto improve U and V while respecting the incentive constraint. For this, first compute c that makes $\lambda DV(w,a) + (1-\lambda)DU(w,a) - c[-Df(w), 1]$ orthogonal to $[-Df(w), 1]$. Denoting by $\overline{c}$ the value of such c, $\overline{c} = \frac{1}{|[-Df(w),1]|^2}[-(\lambda DV_w(w,a) + (1-\lambda)DU_w(w,a))Df(w) + (\lambda D$ Next let $\overline{\lambda} \in \arg\min |\lambda DV(w,a) + (1-\lambda)DU(w,a) - \overline{c}(\lambda, w, a)[-Df(w), 1]|$. Then, a Pareto improving change in w respecting the incentive constraint of the agent is given by $dw = \overline{\lambda} D_w V(w,a) + (1-\overline{\lambda})DU_w(w,a) + \overline{c}(\lambda, w, a)Df(w)$. If $\overline{\lambda} \in (0,1)$, U and V improve at the same rate; $dU = dV$.

References

[1] Aubin, J-P, *Viability Theory*, Birkhäuser, 1991.

[2] Clarke, F., *Optimization and Nonsmooth Analysis*, Wiley-Interscience, 1983.

[3] Cornet, Bernard, "The Second Welfare Theorem in Nonconvex Economies," *CORE* discussion paper 8630, 1986.

[4] D'Aspremont, Claude and Henry Tulkens,"Commodity Exchanges as Gradient Processes," *Econometrica* **48**, no. 2, March 1980.

[5] Dixit, Avinash, "On Pareto-improving Redistributions of Aggregate Economic Gains," *Journal of Economic Theory* **41**, 133-153, 1987.

[6] Grossman, Sanford J. and Oliver Hart, "An Analysis of the Principal-Agent Problem," *Econometrica* **51**, 7-46, 1983.

[7] Guesnerie, R., "Pareto Optimality in Nonconvex Economies," *Econometrica* **43**, no. 1, January 1975.

[8] Guesnerie, R.,"On the Direction of Tax Reform," *Journal of Public Economics* **7**, 179-202, 1977.

[9] Harsanyi, J.C.,"Approaches to the Bargaining Problems Before and After the Theory of Games:A Critical Discussion of Zeuthen's, Hick's and Nash's Theories," *Econometrica* **24**, 114-157, 1956.

[10] Jewitt, Ian, "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica* **56**, no. 5, 1177-1190, September 1988.

[11] Luenberger, D., *Optimization by Vector Space Method*, John Wiley & Sons, 1968.

[12] Malinvaud, E., *Lectures on Microeconomic Theory*, North Holland 1972.

[13] Myerson, Roger, *Game Theory-analysis of conflict*, Harvard University Press 1991.

[14] Nash, J., "The Bargaining Problem," *Econometrica* **18**, 155-162, 1950.

[15] Rogerson, William P., "The First-Order Approach to Principal-Agent Problems," *Econometrica* **53**, 1357-1368, 1985.

[16] Shapley, L. S., "Utility Comparison and the Theory of Games," in *La Decision*, Editions du CNRS, Paris, 251-263, 1969.

[17] Stoer, Josef and Christoph Witzgall, *Convexity and Optimization in Finite Dimensions* **I**, Springer-Verlag, 1970.

[18] Tirole, J and R. Guesnerie, "Tax Reform from the Gradient Projection Method," ***J**ournal of Public Economics* **15**, 275-293, 1981.

[19] Tulkens, Henry and Shmuel Zamir, "Surplus-sharing Local Games in Dynamic Exchange Processes," *Review of Economic Studies* **46**, no. 2, 305-313, 1979.

[20] Weymark, J., "Undominated Directions of Tax Reform," *Journal of Public Economics* **16**, 343-369, 1981

[21] Turunen-Red, Arja H. and Alan D. Woodland, "Strict Pareto-Improving Multilateral Reforms of Tariffs," *Econometrica* **59** July 1991.

[22] Yun, Kwan Koo, "Pareto Optimality in Nonconvex Economies and Marginal Cost Pricing Equilibria," *Korean Economic Review* **1**, 1-13, December 1985.

[23] Yun, Kwan Koo, "Structure of Distortion Equilibria and Welfare in Multi-regional Economies," *Economic Theory* **5**, 491-512, 1995.

[24] Yun, Kwan Koo, "Dubovickii-Miljutin Lemma and Characterizations of Optimal Allocations in Nonsmooth Economies," *Journal of Mathematical Economics* **24,** 435-460, 1995.

[25] Yun, Kwan Koo, "Pareto Improving Changes and Approximate Pareto Optima," mimeo 1995.