Sel gal 14-1 CN

CT

12/12 x 27 TR

5. THE ALLOCATION OF EFF~~ECT~~ORT

A

5.1. Introduction

Consider the following problems:

1. ~~(i)~~ ~~One~~ a person has a certain sum of money, and a number of projects in which to invest it. The return from each project depends on the amount invested, and the problem is to split the money among projects so as to maximize the total return.

2. ~~(ii)~~ The return from an activity depends on the amount of time devoted to it. Split the 24 hours of a day among activities so as to maximize the total return.

3. ~~(iii)~~ The expected return from oil exploration depends on the region explored and the intensity of search effort there ~~in that region~~. Allocate a given total searching effort so as to maximize the expected total return.

4. ~~(iv)~~ The crime rate in an urban district depends on the district and the number of policemen patrolling it. Distribute the police force over Space so as to minimize total crimes.

5. ~~(v)~~ Again there is a range of possible activities. Some are productive, earning money but with a disutility attached to participating in them; some are consumptive, yielding utility for the spending of money. The problem is to maximize total net utility, subject to total spending being equal to total earning.

All ~~of~~ these problems have the following formal structure:

Maximize

$$\underline{f}_1(\underline{x}_1) + \dots + \underline{f}_n(\underline{x}_n) \qquad \text{(5.1.1)} \ \cancel{(1)}$$

subject to

$$\underline{x}_1 + \dots + \underline{x}_n = \underline{X} . \qquad \text{(5.1.2)} \ \cancel{(2)}$$

Here $f_i(x)$ is the return from allocating an amount $\underline{x}$ to project $\underline{i}$. Equation (2) is the fundamental "budget" constraint, stating the total amount ~~one has~~ available to allocate. $\underline{X}$ can be time, money, effort, resources, etc. The "projects" $\underline{i} = 1, \dots, \underline{n}$ can be regions of Space, periods of Time, activities, ~~etc.~~ The individual $\underline{x}_i$ may be required to be non-negative, but not necessarily. In problem ~~(v)~~ 5, for example, ~~one~~ we could measure the amount devoted to activity $\underline{i}$ by the money spent; in the case of productive activities this would be negative. Total spending equal to total earning is then represented as: $\underline{X} = 0$. The return functions $\underline{f}_i$ can also be negative. In problem ~~(iv)~~ 4, ~~for example,~~ $\underline{f}_i(\underline{x})$ would be minus the number of crimes in district $\underline{i}$ with $\underline{x}$ policemen assigned there.

Problems of the form (1)–(2) are among the simplest and most ubiquitous of all problems. One popular definition of economics, ~~in fact,~~ takes it to be the study of the relationship "... between ends and scarce means which have alternative uses,"[1] which may be construed as the study of problems involving constraints of the form (2). Though this seems rather narrow, it indicates the pervasiveness of this condition.

Our aim in the first part of this chapter is to study problem (1)–(2) under conditions of extreme generality. More exactly, we ~~shall~~ assume that the set of possible projects forms a general measure space, ~~and~~ not just a finite set. Such generalization is clearly in order for many problems. The distribution of resources over *Space* or *Time* is over a continuum. The number of possible alternative investment opportunities will often be infinite.[2]

Our results will generalize existing work in several directions:

(*i*) very weak restrictions on the nature of the payoff functions *f*;

(*ii*) very flexible feasibility conditions, including the possibility of negative investments;

(*iii*) no restriction that the measure ~~space~~ be over *n*-space, or that it be non-atomic; no topological or metrical conditions imposed on it; and

(*iv*) pseudomeasure-valued utilities.

All ~~of~~ these generalizations are of interest for one application or another.

~~On the other hand~~ However, we use just one constraint, while other formulations allow several.

5.2. Formulating the Problem

We start with the following ingredients: a measure space $(\underline{A},\Sigma,\mu)$, where μ is ~~sigma~~ σ-finite, and a measurable function f: A × reals → reals. (Where the real numbers are concerned, measurability refers, as always, to the Borel field).

The problem is to find a measurable function δ: A → reals ~~which~~ That maximizes the utility function

$$U(\delta) = \int_A f(a, \delta(a))\, \mu(da) \tag{5.2.1}$$

~~(1)~~

over a certain feasible set of such functions δ.

The expression (1) is an indefinite integral yielding a signed measure over the space A, or, more generally, a pseudo-measure in case (1) is not well defined in the ordinary sense. Integral (1) is always well-defined as a pseudomeasure, because μ is ~~sigma~~ σ-finite and the integrand f(a, δ(a)) is finite and measurable. (Measurability follows from the fact that it is the composition of two measurable functions: a → (a, δ(a)), and f itself).

We use standard ordering of pseudomeasures to rank alternative functions δ. In the present case, by the standard integral theorem this means that δ_1 is at least as preferred as δ_2 iff

114 20 198 set 12 pt brackets

$$\int_A \left[f(a,\delta_1(a)) - f(a,\delta_2(a))\right] \mu(da) \tag{5.2.2}$$

~~(2)~~

is well-defined as an ordinary definite integral, and is ≥ 0. This is possible even if (1) is not well-defined in the ordinary sense for either δ_1 or δ_2. If (2) is not well-defined, ~~then~~ δ_1 and δ_2 are not comparable under standard order. The possibility of non-comparability makes it important to recall the distinction between a given feasible δ^* being best ($\delta^* \succsim \delta$, for all feasible

visi clear for ~ (tilda)

δ), and being merely unsurpassed (there is no feasible $\delta > \delta^*$).

If the ~~integrals~~ integral (1) ~~are~~ is finite for all feasible δ, ~~then~~ standard order reduces to the ordinary comparison of definite integrals, and there is no need to bring in pseudomeasures. (The reader who is troubled by pseudomeasures has the option of adding conditions insuring that (1) is always finite and can ~~He will~~ then obtain a special case of most of the following theorems. Here, as elsewhere, the use of pseudomeasures simplifies and generalizes by enabling us to drop superfluous conditions.)

The utility function (1.1) ~~(1) of the preceding section~~ is a special case of (1) ~~of this section~~: Let A consist of just n points, Σ = all subsets of A, and let μ have the value one on each point (counting ~~enumeration~~ measure). In this case δ is ~~just~~ an n-tuple $(\delta_1,\ldots,\delta_n)$, $f(a,\delta(a))$ for the ith point a may be written $f_i(\delta_i)$, and the integral (1) reduces to the finite sum (1.1) ~~(1) of section 1.~~

Symbols A, f, μ, and δ may be interpreted as follows: A is the set of alternative "projects" among which we are allocating, and may be a set of locations, times, activities, etc. For fixed $a \in A$, $f(a,\cdot)$ is a real-valued function of a real variable, which gives the "payoff density" yielded from the "investment density" $\delta(a)$ applied to point a. μ may be areal measure over Space, or time-measure over Time, or some other measure such that (1) gives the utility. If it has a positive value at a single point, then in general a nonzero payoff may be obtained from that point. Finally, δ gives the density of the distribution of the resource, money, time, effort, etc. over the alternatives of A. That is, the distribution of the resource being allocated is given by the indefinite integral over A:

$$\int \delta \, d\mu \, . \qquad (5.2.3) \; \cancel{(3)}$$

~~So far~~ We have said nothing about the feasibility conditions for δ, (except that it must be real-valued and measurable). We shall always require that (3) be a finite signed measure; ~~that is,~~ i.e., any feasible δ satisfies the condition:

$$\left| \int_{A} \delta \, d\mu \right| < \infty \, . \qquad (5.2.4) \; \cancel{(4)}$$

Now consider the apparently much more specialized condition:

$$\int_{A} \delta \, d\mu = 0 \, . \qquad (5.2.5) \; \cancel{(5)}$$

9.3 Condition (5) seems very narrow. Only ~~example (x) of the ones we have discussed~~ problem 5 (of Section 5.1) satisfies it (total earning = total spending, so total net resource endowment X = 0). The others have positive total resource endowments. However, we now show that any allocation problem satisfying feasibility condition (4) can be converted into another allocation problem satisfying (5).

Theorem: Given A, Σ, μ, f, M, where (A, Σ, μ) is a measure space, μ is ~~sigma~~ σ-finite, f: A $\times$ reals $\to$ reals is measurable, and M is a set of measurable functions δ: A $\to$ reals, all satisfying (4), the problem being to maximize (1) over $\delta \in$ M.

Then ~~there exist~~ A', Σ', μ', f', $\mathbb{M}'$ satisfying the same conditions, and for which (5) is satisfied for all $\delta' \in \mathbb{M}'$, such that there is a 1-1 correspondence between $\mathbb{M}$ and $\mathbb{M}'$ ~~which~~ that preserves preferability relations.

(That is, if $\delta_1, \delta_2 \in \mathbb{M}$ correspond to $\delta_1', \delta_2' \in \mathbb{M}'$, respectively, then $\delta_1 \succcurlyeq \delta_2$ iff $\delta_1' \succcurlyeq' \delta_2'$. Here the preference relation $\succcurlyeq$ comes from (2), while $\succcurlyeq'$ comes from (2) with f', μ' substituted for f, μ).

Proof: Let z_o be an artificial point not belonging to A. Define

(i) $A' = A \cup \{z_o\}$; $\Sigma' = \{G \mid G \subseteq A' \text{ and } G \backslash \{z_o\} \in \Sigma\}$;

(ii) $\mu'(G) = \mu(G)$ if $z_o \notin G$; for all $G \in \Sigma'$;
$\mu'(G) = \mu(G \backslash \{z_o\}) + 1$ if $z_o \in G$;

(iii) $f'(a,x) = f(a,x)$, all $a \in A$, all real x;
$f'(z_o,x) = 0$, all real x.

Let $\mathbb{M}'$ ~~is~~ be the set of all functions δ': $A' \to$ reals whose restriction to A belongs to $\mathbb{M}$, and which satisfy

$$\delta'(z_o) = -\int_A \delta' \, d\mu. \qquad (5.2.6)$$

Note that $\delta'(z_o)$ is finite, by condition (4).

Now μ' restricted to A coincides with μ, and $\mu'\{z_o\} = 1$. Hence, for $\delta' \in \mathbb{M}'$,

$$\int_{A'} \delta' \, d\mu' = \int_A \delta' \, d\mu + \delta'(z_o)\mu'\{z_o\} = 0,$$

from (6). Hence δ' satisfies (5).

The correspondence between δ' and its restriction to A is 1-1 between $\mathfrak{M}'$ and $\mathfrak{M}$. The definitions of f' and μ' imply that preferability is preserved. $\Box$

Briefly, the new problem is obtained from the old by adding a point z_o to A, making $\{z_o\}$ measurable, giving it measure one, setting $f = 0$ on it, and giving any δ a value on z_o that just cancels the surplus or deficit of $\int_A \delta_A d\mu$ on A. (A procedure similar to this is very common in finite problems, where it takes the form of adding "disposal activities," "slack variables," etc.)

This result is very useful, because condition (~~7~~ 5) is mathematically convenient. Our standard procedure will be as follows. The heavy mathematical work will be on problems with the special condition (5). Having obtained a result, we then go to the general problem. This is translated into a problem satisfying (5) by means of the recipe in the proof just given. The result is applied to the translated problem, and usually yields a more general theorem for the general problem.

~~Let us~~ We now turn to a more specific system of feasibility conditions. Let two measurable functions, b, c: $A \to$ extended reals, be given, as well as two numbers, L_o and L^o, which may also be infinite. In terms of these, the feasible set consists of those functions δ ~~which~~ that are measurable, real-valued, ~~which~~ satisfy (4), and ~~which~~ also satisfy the two conditions

$$b \le \delta \le c, \qquad \text{(5.2.7)} \ \cancel{(7)}$$

5a 1 4

~~and~~

$$\underline{L}_o \leq \int_{\underline{A}} \delta \, d\mu \leq \underline{L}^o , \qquad (5.2.8)\ \cancel{(8)}$$

That is, for all $\underline{a} \in \underline{A}$, $\delta(\underline{a})$ satisfies the double inequality $\underline{b}(\underline{a}) \leq \delta(\underline{a}) \leq \underline{c}(\underline{a})$, and in addition its integral satisfies the double inequality (8). Allowing $\underline{b}$, $\underline{c}$, $\underline{L}_o$ or $\underline{L}^o$ to be infinite is simply a device for removing some of these constraints; e.g., ~~For example,~~ setting $\underline{L}_o = -\infty$ is the same as simply removing the left~~-hand~~ inequality in (8).

This system of constraints is very flexible; e.g., ~~For example,~~ if the density δ must by its nature be non-negative, this may be indicated formally by setting $\underline{b}$ ~~to be~~ identically zero. ~~On the other hand,~~ Or, if there is no lower limit to δ at any point, set $\underline{b}$ identically equal to $-\infty$. A total resource constraint that must be satisfied with equality, (as in (1.2) ~~(2)~~ of section 1) ~~the preceding section)~~ is indicated by setting $\underline{L}_o = \underline{L}^o$, and both equal to total available net ~~nnt~~ resources. The function $\underline{c}$ is an investment capacity constraint, limiting the amount of resource that can be squeezed into the various subsets of $\underline{A}$. The function $\underline{b}$ is an investment requirement constraint, in that it places a lower bound (positive, negative, or zero) on investment over the various subsets of $\underline{A}$.

This is the feasibility system that ~~which~~ will occupy most of our attention. We now prove a result ~~which~~ specializing the theorem above, and showing that this problem can be transformed into one with a simpler system of constraints.

#

Theorem: Given measure space (A, Σ, μ), μ ~~sigma~~ σ-finite, measurable functions $f: A \times \text{reals} \to \text{reals}$, $b, c: A \to$ extended reals, and extended real numbers L_o, L^o, consider the problem of maximizing (1) over those measurable real-valued functions δ ~~which~~ that satisfy (4), (7), and (8).

Then ~~there~~ exist A', Σ', μ', f', b', c', with analogous properties, from which the following problem is formulated: Maximize (1') over measurable real-valued functions δ' satisfying (5') and (7'). (The primes indicate that f', b', c', etc., ~~is~~ are to be substituted for f, b, c, etc.) There is a 1-1 correspondence between the feasible sets of these two problems, which preserves preferability relations.

Before going on to the proof, note the effect of this theorem. The constraint (8) is eliminated and replaced by (5), so that, instead of being confined to the interval $[L_o, L^o]$, $\int_A \delta \, d\mu$ must be zero. This is very convenient mathematically.

Proof: Take an artificial point z_o not belonging to A, and define A', Σ', μ', and f' exactly as in the proof of the theorem above. Let b' and c' be the functions ~~which are~~ identical to b and c, respectively, when restricted to A, and for which

$$b'(z_o) = -L^o, \quad c'(z_o) = -L_o. \qquad (5.2.9)$$

We show that with these definitions, the feasible sets of the original and transformed problems are in 1-1 correspondence. With each $\delta: A \to$ reals, feasible for the original problem, associate the function $\delta': A' \to$ reals ~~which~~ that coincides with δ on A, and ~~which~~ satisfies (6) for $\delta'(z_o)$. Any such δ' is feasible for the transformed problem, ~~For,~~ since (5') follows from (6) just as in the preceding proof. Also, (7') is satisfied for all $a \in A$, and as for z_o, the condition

$$b'(z_o) \leq \delta'(z_o) \leq c'(z_o) \qquad (5.2.10)\ ~~(10)~~$$

is an immediate consequence of (6), (8), and (9). This proves δ' is feasible for the transformed problem.

D

Conversely, let δ' be transformed-feasible. Its restriction to A is then feasible for the original problem, since (8) follows from (9) (10), and (5'), and the other feasibility conditions are obviously satisfied. Furthermore, the function $\delta'': A' \to$ reals associated with the restriction of δ' is δ' itself; this follows again from (5'). We have proved that the original feasible set is mapped onto the transformed feasible set.

Finally, if two functions δ_1 and δ_2, are unequal, their extensions are obviously unequal. This proves we have indeed a 1-1 correspondence. That preferability relations are preserved follows from the way ~~in which~~ f' and μ' are defined. ||| □

We make one final preliminary point. Given a measure space (A, Σ, μ), recall that a condition is said to hold almost everywhere, or for almost all $a \in A$, iff there is a set $E \in \Sigma$ of measure zero (a null set) such that the condition holds for all $a \notin E$. (The case $E = \emptyset$ is not excluded; here the condition would hold for all $a \in A$).

Now let $\delta_1, \delta_2 : A \to$ reals both be feasible, and be equal almost everywhere: $\mu\{a \mid \delta_1(a) \neq \delta_2(a)\} = 0$. ~~One easily sees~~ We can see that

$$\int f(a, \delta_1(a))\,\mu(da) = \int f(a, \delta_2(a))\,\mu(da)$$

~~and~~

$$\int \delta_1\, d\mu = \int \delta_2\, d\mu, \qquad (5.2.11)\ \ \cancel{(10)}\ (11)$$

so that δ_1 and δ_2 yield the same utility function and ~~same~~ resource distribution. In effect, δ_1 and δ_2 are two different representations of the same solution, ~~(10)~~ (11). ~~One~~ We could systematically ignore exceptions to rules ~~which~~ occur(ring) within null sets only. ~~For example~~ e.g., the constraint (7) could be weakened to:

$$b(a) \leq \delta(a) \leq c(a)$$

for almost all $a \in A$, without altering the problem in any essential way. In any case, ~~one~~ we should be prepared to find the following discussion well-seasoned with the phrases "almost everywhere", "almost all".

5.3. Sufficient Conditions for Optimality

A feature that characterizes a very wide class of optimization problems is the rôle played by "multipliers" or "shadow prices". These are numbers associated with the constraints of the problem from which special conditions are formed, either necessary or sufficient for optimality, and which sometimes allow ~~one~~ us to transform the original problem into a new, and simpler, ~~problem~~ one.

These "prices" are especially useful in economic and social science problems, because they not only expedite the solution, but suggest institutional arrangements ~~which~~ that will lead the economy to carry out the solution in practice.

In the ~~resource~~ allocation -of-effort problem we would expect that a "price" could be associated with the total resource constraint, in such a way that someone taking account of the "cost" of the resources allocated to the various projects - (as well as the "payoff" from these projects) - would be led to the optimal solution. ~~In reality~~ a number of qualifications must be added, but this idea is a red thread ~~which~~ ~~runs~~ running through the following results.

We first give a very general condition ~~which~~ ~~guarantees~~ guaranteeing that a feasible solution is best. ~~Note that~~ We are dealing with a utility function ~~that is~~ partially ordered, so the conclusion that δ is <u>best</u> is much stronger than the conclusion that it ~~it~~ be merely <u>unsurpassed</u>.

Theorem: Let (A,Σ,μ) be a measure space, μ ~~sigma~~ σ-finite; let f: $A \times$ reals $\to$ reals be measurable, and let $\mathfrak{M}$ be a collection of measurable function $\delta : A \to$ reals such that

$$\int_A \delta \, d\mu = 0 \tag{1}$$

for all $\delta \in \mathcal{M}$. Let $\delta^o \in \mathcal{M}$ be a function, and p^o a real number, such that, for all $\delta \in \mathcal{M}$,

$$f(a,\delta^o(a)) - p^o\delta^o(a) \geq f(a,\delta(a)) - p^o\delta(a), \tag{2}$$

for almost all $a \in A$.

Then δ^o is **best** for the problem of maximizing

$$\int f(a,\delta(a))\,\mu(da) \tag{3}$$

over $\delta \in \mathcal{M}$. (Maximization refers to standard ordering of pseudomeasures, here and throughout this discussion).

Proof: We must show that, for any feasible δ,

$$\int_A \left[f(a,\delta^o(a)) - f(a,\delta(a))\right]\mu(da) \tag{4}$$

is well-defined as an ordinary definite integral, and is ≥ 0.

From (2) we have

$$f(a,\delta^o(a)) - f(a,\delta(a)) \geq p^o[\delta^o(a) - \delta(a)]$$

almost everywhere. By (1), the integral of $p^o[\delta^o(a) - \delta(a)]$ is zero, and this fact is all we need. |||

Here p^o is the shadow price, and (2) asserts that, for each $a \in A$ (except possibly on a null set), the investment density $\delta^o(a)$ is chosen to maximize $f(a,x) - p^o x$ over the feasible investment levels x. The first term gives the payoff density, and the second reflects the "cost" of using up the resource. (Note that either f or p^o, or both, can be negative).

We can ~~easily~~ derive a generalization of this result.

Theorem: Let A, Σ, μ, f and $\mathbb{M}$ be as above, except that the feasibility condition (1) is replaced by the weaker condition

$$\left|\int_A \delta \, d\mu\right| < \infty \qquad \text{(5.3.5)} \; \cancel{(5)}$$

for all $\delta \in \mathbb{M}$. Let $\delta^o \in \mathbb{M}$ be a function and p^o be a real number such that (2) holds, and also

$$p^o \cdot \int_A (\delta^o - \delta)\, d\mu \geq 0 \qquad \text{(5.3.6)} \; \cancel{(6)}$$

for all $\delta \in \mathbb{M}$. Then δ^o is best for the problem of maximizing (3) over $\delta \in \mathbb{M}$.

Proof: We take an artificial point z_o and transform this problem into its equivalent on $A' = A \cup \{z_o\}$ (see section 5.2). This translated problem is in proper form for the theorem just given, and we need merely verify that condition (2) holds almost everywhere on $A \cup \{z_o\}$. By assumption this is ture on A. For the point z_o we have that $f(z_o,\cdot)$ is identically zero, $\delta'(z_o) = -\int_A \delta \, d\mu$, and similarly for $\delta^{o\prime}(z_o)$. Condition (2) for the point z_o is then precisely ~~the~~ condition (6). $\square$

The extra condition (6) ~~that is~~ imposed is easy to interpret. If p^o is positive, ~~then~~ there is no feasible δ for which $\int_A \delta_1 \, d\mu > \int_A \delta^o_1 \, d\mu$. In common sense terms, this states that if the resource is valuable, as much ~~of it~~ as possible should be allocated. ~~Similarly,~~ if p^o is negative, the "resource" is illth rather than wealth, and as little ~~of it~~ as possible should be used. Finally, if $\int_A \delta^o_1 \, d\mu$ is neither the highest nor the lowest possible value attainable, then (6) implies $p^o = 0$. In this case the resource is a "free good", and for each $a \in A$ we simply choose $\delta^o(a)$ to maximize $f(a,x)$ over attainable investment levels x without worrying about resource cost.

We can also find a sufficient condition for δ^o to be the *unique* best solution. But ~~one has to~~ we must be careful in interpreting the concept of "uniqueness" here. According to previous discussion we may identify two functions δ_1 and δ_2 ~~which~~ that are equal almost everywhere. Let us say that δ_1 and δ_2 are *essentially distinct* iff $\mu\{a \mid \delta_1(a) \neq \delta_2(a)\} > 0$. Then, in line with our discussion, we say that δ^o is the *unique* best solution iff δ^o is best and there is no essentially distinct δ ~~which~~ that is also best.

There may obviously be more than one best solution. For example, if f is identically zero, ~~then~~ *any* feasible solution is best.

14-2

We now give the uniqueness condition. Going back to the first sufficiency theorem, suppose that all the premises hold, and in addition the following: for each $\delta \in M$ ~~which~~ that is *essentially distinct* from δ^o, there is a set E_δ of *positive* measure, such that (2) holds with strict inequality for all $a \in E_\delta$. Then δ^o is the unique best solution.

The proof is simple. Reasoning ~~just~~ as before, we find ~~that~~ the integral (4) is positive, not merely non-negative. Hence $\delta^o > \delta$ for all $\delta \in M$ essentially distinct from δ^o.

This uniqueness condition immediately generalizes to the case where (1) is replaced by (5). Let all the premises of the second sufficiency theorem hold, and, in addition, the following: For each $\delta \in M$ ~~which is~~ essentially distinct from δ^o, either there is a set E_δ as above, or (6) holds with strict inequality (or both). Then δ^o is the unique best solution.

The proof consists in translating this problem into the one in which (1) holds, then applying the E_δ condition to this translated problem. In doing so, note that the singleton set $\{z_o\}$ has positive measure: $\mu'\{z_o\} = 1$. This shows that strict inequality in (6) for all δ insures uniqueness.

Finally, we mention a much weaker sufficient condition for δ^o being best. In the theorem above p^o is chosen in advance and (2) must be satisfied for all δ. But it suffices that, for any feasible δ there exists a p^o (which may be different for different δ's) such that (2) is satisfied for this p^o and δ. This variable p^o lacks the appeal of the shadow price interpretation, however.

A

5.4. Necessary Conditions for Optimality

The sufficient conditions we have just stated are very convenient to use where they hold. Unfortunately, it is easy to find problems whose optimal solution does not satisfy any such condition; ~~that is~~ i.e., the stated conditions are not always necessary.

Here is a simple example. Let A consist of two points: $A = \{a_1, a_2\}$, Σ = all subsets, $\mu = 1$ on each point; let the payoff function be given by: $f(a_1, x) = x^2$, $f(a_2, x) = -2x^2$; the investment function is given by (x_1, x_2), with the constraint $x_1 + x_2 = 0$. Thus the problem is of the finite form ~~(1)~~ (1.1) – ~~(2)~~ (1.2) ~~of section 1:~~

Maximize

$$x_1^2 - 2x_2^2, \qquad (5.4.1)\ \cancel{(1)}$$

subject to

$$x_1 + x_2 = 0. \qquad (5.4.2)\ \cancel{(2)}$$

The unique best solution is obviously $x_1 = x_2 = 0$. Now the sufficient condition (3.2) ~~(2) of section 3~~ requires that there be a real number p° such that (for point a_1)

$$0 \geq x^2 - p^{\circ}x$$

for all real x. Obviously, there is no such p°. (~~On the other hand~~ However, this example does not violate the weaker condition mentioned at the end of ~~the preceding~~ 5.3 section.)

We ~~shall~~ now investigate necessary conditions for optimality. Broadly speaking, these have been found more useful than sufficient conditions, because it is hard to find simple sufficient conditions ~~which~~ that cover a very wide range of problems. ~~Also,~~ establishing necessary conditions is~~, by and large,~~ much more difficult than establishing sufficient conditions. (Compare the length of proofs in this section with that preceding).

The classical example of a necessary condition is that a function maximized in the interior of a domain have a derivative of zero at the maximizing point, if it is differentiable there. Necessary conditions in general are used just as this one is; viz., ~~Namely, one~~ we ~~narrows~~ narrow the search for an optimal solution to those (hopefully few) points ~~which~~ that satisfy the necessary condition, and then ~~tries~~ try by other means to test ~~these~~ them directly for optimality.

14.8 We ~~shall~~ concentrate on the special class of allocation problems discussed above, where the feasible solutions are those ~~which lie~~ lying between two functions, b, c: A → extended reals, and ~~which integrate~~ integrating to zero. Afterward, some of the key results will be generalized to the problem where the constraint

Break after ≤

$$\underline{L}_o \le \int_A \delta \, d\mu \le \underline{L}^o$$

replaces the condition that the integral be zero.

A number of preliminary concepts and lemmas will be needed before we can get down to serious business.

Recall that

~~The~~ The **supremum** of a set of extended real numbers, ~~it will be recalled,~~ is the smallest extended real number not less than any of the numbers in the set (it is the "least upper bound" of the set). The supremum of a function, $g: A \to$ extended reals, is defined to be the supremum of its range, $\{g(a) | a \in A\}$.

We ~~Let us~~ use the notation $g|E$ to represent the restriction of g to the subdomain $E \subseteq A$. The supremum of $g|E$ is less than or equal to the supremum of g itself. ~~Now~~ Suppose the domain A of g is the universe set of a measure space (A,Σ,μ). We consider all possible restrictions $g|E$ such that $\mu(A\setminus E) = 0$, and take the supremum of each one. The infimum of the resulting set of extended real numbers is called the **essential supremum** of g. To put it another way, the essential supremum of g is the largest extended real number x such that $x \leq \sup \{g(a) | a \in E\}$ for all $E \in \Sigma$ such that $\mu(A\setminus E) = 0$.

One special case may be noted. If μ is the identically zero measure, A itself is a null set, and we may take $E = \emptyset$. The range of $g|E$ is then the empty set $\emptyset$. Applying the definition of supremum literally yields $\sup \emptyset = -\infty$. Thus the essential supremum of any function g is, in this case, $-\infty$.

The **essential infimum** of a function $g: A \to$ extended reals, with respect to (A,Σ,μ), is defined analogously. Just switch the words "supremum" and "infimum", and the words "greatest" and "least", in the preceding discussion. Or, equivalently, we could define it by the rule:

Display

$$\text{essential infimum of } g = -\text{essential supremum } (-g).$$

ll left

We ~~shall~~ use the standard abbreviations "ess sup", and "ess inf" for these concepts.

The following result, whose proof is omitted, wi~~th~~ll be needed later.

#

Lemma: Given measure space (A,Σ,μ) and function $g: A \to$ extended reals, let $\{A_0, A_1, \ldots\}$ be a countable measurable partition (or even just a covering) of A. Then

D

$$\text{ess sup } g = \sup\left\{\text{ess sup}(g|A_n) \,\middle|\, n = 0, 1, \ldots\right\} \qquad (5.4.3)\ \cancel{(3)}$$

-12pt braces

~~and~~

$$\text{ess inf } g = \inf\left\{\text{ess inf}(g|A_n) \,\middle|\, n = 0, 1, \ldots\right\}, \qquad (5.4.4)\ \cancel{(4)}$$

#

Here ess sup$(g|A_n)$ refers to the function $g|A_n$ and the measure space (A_n,Σ_n,μ_n), which is the restriction of (A,Σ,μ) to A_n, and similarly for ess inf$(g|A_n)$.

~~It will be~~ Note~~d~~ that measurability of g is nowhere mentioned. Indeed, these concepts are perfectly well-defined, and the lemma is correct, for any function g, measurable or not. This is important, because the functions g and h, which we define below, are not necessarily measurable.

Next, let f be a real-valued function whose domain is an interval of real numbers, $[b,c]$. (The endpoints of the interval may or may not be included, and we may have $b = c$.) Function f is **continuous** at the point $x_0 \in [b,c]$ iff, for any sequence $x_1, x_2, \ldots$ of points of $[b,c]$ whose limit is x_0, and any number $\varepsilon > 0$, there is an integer N such that, for all $n > N$,

set epsilon

$$-\varepsilon < f(x_o) - f(x_n) < \varepsilon. \qquad (5.4.5)\ \cancel{(5)}$$

Now (5) is a double inequality. If we drop the left~~-hand~~ inequality in (5), keeping everything else the same, we get a weaker concept. f is then said to be lower semi-continuous at the point x_o. Intuitively, lower semi-continuity at x_o allows f to take a "sudden" jump downward, but not upward, at x_o. (Continuity prohibits "sudden" jumps in either direction).

Function f is continuous, or lower semi-continuous, iff it is continuous, or lower semi-continuous, at every point of its domain, respectively.

14-9 There is another way to characterize lower semi-continuity ~~which~~ that is more useful (though less intuitive) then the definition just given. An open interval of real numbers is an interval not containing its endpoints (this is the same as an open disc on the real line). An open set on $[b,c]$ is the intersection of $[b,c]$ with any union of open intervals. Then $f: [b,c] \to$ reals is lower semi-continuous iff $\{x|f(x) > y\}$ is an open set on $[b,c]$, for all real numbers y. We omit the proof that the two lower semi-continuity concepts are the same.

We ~~shall~~ need the following result.

D **Lemma**: Let $f: [b,c] \to$ reals be lower semi-continuous, with $b < c$. Then

$$\sup f = \sup(f|E), \qquad (5.4.6)\ \cancel{(6)}$$

where E is the set of rational numbers in $[b,c]$.

Proof: Obviously, sup $f \geq \sup(f|E)$. Conversely, for any number $y < \sup f$, the set $\{x | f(x) > y\}$ is open and non-empty; hence there is a rational number x_o belonging to it: $f(x_o) > y$. |||

Next, we need several concepts related to, but more general than, the concept of derivative of a function. Let $f: [b,c] \to$ reals again have a real interval domain which may or may not include its endpoints, and let x be a point of the domain $\neq c$.

Definition: The lower right derivate of f at the point x is the limit, as ε goes to zero from above, of set

$$\inf\left\{[f(x+y) - f(x)]/y \;\middle|\; 0 < y < \varepsilon\right\}. \qquad (5.4.7) \quad \cancel{(7)}$$

That is, given $\varepsilon > 0$, we find for each point y in the open interval $(0,\varepsilon)$[3] the value of $[f(x+y) - f(x)]/y$ — which is the average slope of f from x to $x + y$ — and take the infimum of this set of values. Having done this for each $\varepsilon > 0$, we take the limit as $\varepsilon \to 0$.

This concept is well-defined for any real-valued function, and any domain point except the right endpoint, c, but it may take on an infinite value. For, ~~first of all~~ the infimum of any set of real numbers is some extended real number, so (7) is well-defined for fixed $\varepsilon > 0$. Furthermore, one sees that (7) is non-decreasing as $\varepsilon \to 0$, and hence has a limit in the extended real numbers. This of course contrasts with differentiability, which is not a universal property.

We ~~shall~~ use the notation $D_+f(x)$ for the lower right derivate of f at x. If the right endpoint, c, is part of the domain, we ~~shall~~ make the convention that $D_+f(c) = -\infty$.

Similarly, for all domain points, x, other than the left endpoint, b, we have the following.

Definition: The upper left derivate of f at the point x is the limit, as ε goes to zero from above, of

$$\sup\left\{[f(x) - f(x-y)]/y \;\middle|\; 0 < y < \varepsilon\right\}. \tag{8}$$

- 12 pt faces (5.4.8)

An argument similar to that ~~the one~~ just given shows ~~that~~ this is always well-defined in the extended real numbers. We ~~shall~~ use the notation $D^-f(x)$ for the upper left derivate of f at x. If the left endpoint, b, is in the domain, we make the convention that $D^-f(b) = +\infty$.

(One may also define the upper right derivate: replace "inf" by "sup" in (7); and the lower left derivate: replace "sup" by "inf" in (8), but we do not need these concepts. A function is differentiable at interior point x iff all four of these quantities are equal; their common value is then the derivative of f at x).[4]

Next, we need the following result concerning atomic measures.

14-10

Lemma: Let (A,Σ,μ) be a measure space, with μ atomic; let $(f_1, f_2,\ldots)$ be a sequence of measurable functions on A, taking values in the extended real numbers. Then there is <u>exactly one</u> sequence of extended real numbers $(x_1, x_2, \ldots)$ satisfying

$$\mu\left\{a \,\middle|\, f_n(a) = x_n \text{ for all } n = 1, 2, \ldots\right\} \neq 0. \qquad \text{(5.4.9)} \; \cancel{(9)}$$

12 pt braces + vertical bar

Proof: By definition of "atomic", exactly one of the two numbers $\mu(E)$, $\mu(A \setminus E)$ is zero for any choice of $E \in \Sigma$. ~~It is~~ clearly ~~that~~ there can be at <u>most</u> one sequence of numbers satisfying (9), hence we must show there is at <u>least</u> one.

Take any measurable function f and consider the supremum x_o of the set of numbers x satisfying

$$\mu\{a \mid f(a) < x\} = 0 \qquad \text{(5.4.10)} \; \cancel{(10)}$$

If $x_o > -\infty$, take a sequence (x_n) rising to x_o. Since (10) holds for each x_n, it holds for x_o itself. It also obviously holds if $x_o = -\infty$. Now consider the condition

$$\mu\{a \mid f(a) > x\} = 0 \qquad \text{(5.4.11)} \; \cancel{(11)}$$

This obviously holds for x_o if $x_o = \infty$. If $x_o < \infty$, take a sequence (x_n) decreasing to x_o; then (11) holds for each such $x_n > x_o$ (since the complement of the set in (11) has positive measure). Hence (11) holds again for x_o. We have thus shown that

$$\mu\{a \mid f(a) \neq x_o\} = 0 \qquad \text{(5.4.12)} \; \cancel{(12)}$$

Now for each f_n let x_n be the corresponding number satisfying (12). The set $\{a|f_n(a) \neq x_n$ for at least one $n = 1, 2,\ldots\}$ is the union of the countable collection of null sets $(\{a|f_n(a) \neq x_n\})$, $n = 1, 2,\ldots$, hence ~~it~~ itself has measure zero. Therefore its complement is not a null set. But this is exactly the statement (9). $|||$ □

Let us now get down to business. The allocation problem will be determined by the measure space (A,Σ,μ), the payoff function $f: A \times \text{reals} \rightarrow \text{reals}$, and the lower and upper capacity functions, b, $c: A \rightarrow$ extended reals.

We make the following convention. If a specific point $a \in A$ is chosen, $f(a,\cdot)$ is a function of a real variable. In referring to it, we ~~shall~~ always take this function to be restricted to the interval $[b(a), c(a)]$. (The endpoint $b(a)$ is to be included iff it is finite, and similarly for $c(a)$). Thus the statement that $f(a,\cdot)$ is lower semi-continuous refers to this function with the domain $[b(a), c(a)]$, and similarly for derivates of this function; in particular we have at the end-points that $D^{-}f(a, b(a)) = +\infty$ and $D_{+}f(a, c(a)) = -\infty$.

To explain the formulation of the following result, recall that, since μ is ~~sigma~~ σ-finite, there is a countable measurable partition $\mathcal{P}$ $\{A_\Theta, A_1,\ldots\}$ of A, such that μ restricted to A_Θ is non-atomic, while each A_n, $n = 1, 2,\ldots$, is an atom (~~that is~~ i.e., μ restricted to A_n is atomic). Set A_Θ is the non-atomic part of A, and $A\backslash A_\Theta$ is the atomic part. This partition is unique up to null sets. (We do not assume that atoms are simply concentrated. Doing so would simplify some of the following proofs.)

15°

We postulate that the given feasible density δ^o is unsurpassed in the allocation problem. ~~Note that~~ this is a weaker assumption, and therefore it yields a stronger result, than if we postulated that δ^o were best.

The following is called a lemma rather than a theorem because it lacks immediate intuitive appeal. The result is quite powerful, however, and implies all the other results we obtain in this section.

D

Lemma: Let (A,Σ,μ) be a measure space, with μ ~~sigma~~ σ-finite; let $f\colon A \times \text{reals} \to \text{reals}$, and let $b, c\colon A \to \text{extended reals}$, be measurable functions. Let feasible δ^o be unsurpassed for the problem of maximizing

14-11

$$\int f(a,\delta(a))\,\mu(da) \qquad \text{(5.4.13)} \ \cancel{(13)}$$

subject to $\delta\colon A \to$ reals being measurable, and satisfying

$$b \leq \delta \leq c \qquad \text{(5.4.14)} \ \cancel{(14)}$$

and

184 17 61

$$\int_A \delta\, d\mu = 0. \qquad \text{(5.4.15)} \ \cancel{(15)}$$

Let A_0, $A\backslash A_0$ be the non-atomic and atomic parts of A, respectively. Let $f(a,\cdot)$ be lower semi-continuous for all $a \in A_0$.

Define two functions g, $h\colon A \to$ extended reals as follows: For $a \in A_0$,

$$g(a) = \sup_y \Big\{ \big[f(a,\delta^o(a) + y) - f(a,\delta^o(a)) \big] / y \Big\}, \qquad \text{(5.4.16)} \ \cancel{(16)}$$

12 pt braces, brackets & parens

12 pt brackets + braces

the supremum being taken over all y in the open interval $(0, c(a) - \delta^{o}(a))$;

$$h(a) = \inf\left\{\left[f(a,\delta^{o}(a)) - f(a,\delta^{o}(a) - y)\right]/y\right\}, \qquad \text{(5.4.17)}\ \cancel{(17)}$$

the infimum being taken over all y in the open interval $(0, \delta^{o}(a) - b(a))$.

For $a \in A \setminus A_0$,

$$g(a) = D_{+}f(a,\delta^{o}(a)); \qquad \text{(5.4.18)}\ \cancel{(18)}$$

$$h(a) = D^{-}f(a,\delta^{o}(a)). \qquad \text{(5.4.19)}\ \cancel{(19)}$$

Then, for any pair of disjoint sets $G, H \in \Sigma$, we have

$$\text{ess sup}(g|G) \leq \text{ess inf}(h|H). \qquad \text{(5.4.20)}\ \cancel{(20)}$$

Proof: It suffices to prove (20) for the special case where $G \subseteq A_m$, $H \subseteq A_n$, for ~~some~~ some m, $n = 0, 1, 2, \ldots$. For, suppose this has been proved, for each m, n; and let G, H be any disjoint measurable sets. By assumption, we have

$$\text{ess sup}\left[g|(G \cap A_m)\right] \leq \text{ess inf}\left[h|(H \cap A_n)\right]$$

for all m, $n = 0, 1, 2, \ldots$. Then, by (3) and (4),

104 12

146 $$\text{ess sup}(g|G) = \sup_m \left[\text{ess sup}(g|(G \cap A_m))\right]$$

17 242

$$\leq \inf_n \left[\text{ess inf}(h|(H \cap A_n))\right] = \text{ess inf}(h|H),$$

fl left

Since $(G \cap A_m)$, $m = 0, 1, \ldots$, partitions G, and $(H \cap A_n)$, $n = 0, 1, \ldots$, partitions H. Thus (20) would be proved in general.

There are four cases to consider:

(i) $G \subseteq A_0$, $H \subseteq A_0$;

(ii) $G \subseteq A_m$, $H \subseteq A_n$ for $m, n \neq 0$;

(iii) $G \subseteq A_m$, $H \subseteq A_0$ for $m \neq 0$;

(iv) $G \subseteq A_0$, $H \subseteq A_n$ for $n \neq 0$.

In each case we ~~shall~~ assume that (20) is false, and show ~~that~~ there is a feasible δ ~~which~~ that surpasses δ^o, giving a contradiction.

Case (i) $G \subseteq A_0$, $H \subseteq A_0$:

For each positive real number y define the functions $g_y, h_y : A \to$ extended reals by

$$g_y(a) = \left[f\left(a, \delta^o(a) + y\right) - f\left(a, \delta^o(a)\right)\right]/y, \quad \text{if } c(a) - \delta^o(a) > y; \ = -\infty \text{ otherwise}; \tag{5.4.21}$$

$$h_y(a) = \left[f\left(a, \delta^o(a)\right) - f\left(a, \delta^o(a) - y\right)\right]/y, \quad \text{if } \delta^o(a) - b(a) > y; \ = +\infty \text{ otherwise}. \tag{5.4.22}$$

These are all measurable functions, since f, b, c, and δ^o are all measurable. Furthermore, when restricted to A_0,

$$g = \sup g_y, \quad h = \inf h_y, \tag{5.4.23}$$

the sup and inf being taken over all positive real y.

We now show that~~, when restricted to A_0,~~ (23) remains true even if y merely ranges over the positive rational numbers. Take a point $a \in A_0$, and consider $g_y(a)$ as a function of y, with domain the open interval $(0, c(a) - \delta^\circ(a))$. Since f is lower semi-continuous, it follows easily that the y-function $g_y(a)$ is lower semi-continuous. (Use the definition given by the right half of (5)). If $c(a) > \delta^\circ(a)$, it follows from (6) that (23) is true for g when the supremum is taken over the positive rational values of y. If $c(a) = \delta^\circ(a)$, this is still true, the two suprema both being $-\infty$. This proves the contention for g.

As for $h(a)$, we consider $h_y(a)$ as a function of y, with domain $(0, \delta^\circ(a) - b(a))$. ~~One~~ We then ~~verifies~~ verify that minus $h_y(a)$ is lower semi-continuous. ~~One~~ We then ~~verifies~~ verify as above that $\sup(-h_y(a))$ is the same whether taken over positive real y or positive rational y. But $\sup(-h_y(a)) = -\inf(h_y(a)) = -h(a)$, which proves the contention for h.

Now, assuming that (20) is false, choose a real number x satisfying

$$\text{ess sup}(g|G) > x > \text{ess inf}(h|H), \qquad \text{(5.4.24)} \quad \cancel{(24)}$$

and let

187 123

$$G' = G \cap \{a|g(a) > x\}, \quad H' = H \cap \{a|h(a) < x\},$$

~~and~~

$$G_y = G \cap \{a|g_y(a) > x\}, \quad H_y = H \cap \{a|h_y(a) < x\}, \qquad \text{(5.4.25)} \quad \cancel{(25)}$$

for all positive real y.

align both sets of =

Since we are operating within A_θ, (23) is true for y ranging over the positive rationals. Hence

$$G' = \cup\, G_y \text{ and } H' = \cup\, H_y,$$

the union taken over the positive rationals. G' and H', as unions of a countable number of measurable sets, are themselves measurable. Also $\mu(G') > 0$, $\mu(H') > 0$; for if not, (24) would be false. It follows that there must be positive (rational) numbers, y_1 and y_2, such that

$$\mu(G_{y_1}) > 0,\quad \mu(H_{y_2}) > 0, \tag{26}$$

since a countable union of null sets is a null set.

Now μ, being σ-finite and nonatomic on G_{y_1}, takes on all values between 0 and $\mu(G_{y_1})$ on this set, and similarly for H_{y_2}. Hence, from (26), we can find measurable subsets $G'' \subseteq G_{y_1}$ and $H'' \subseteq H_{y_2}$ such that

$$y_1\mu(G'') = y_2\mu(H''), \tag{27}$$

with this common value being positive and finite.

We are now ready to construct another feasible density, $\delta^{\circ\circ}$, which will surpass δ°. Let

$$\begin{aligned}
\delta^{\circ\circ}(a) &= \delta^{\circ}(a) + y_1 \quad \text{if } a \in G''; \\
\delta^{\circ\circ}(a) &= \delta^{\circ}(a) - y_2 \quad \text{if } a \in H''; \\
\delta^{\circ\circ}(a) &= \delta^{\circ}(a) \quad \text{if } a \notin G'' \cup H''.
\end{aligned} \tag{28}$$

~~Note that~~ this is well-defined, since G and H (hence G'' and H'') are disjoint. First, we verify that $\delta^{\circ\circ}$ is feasible. It is real-valued and measurable. Also,

$$\int_A (\delta^{\circ\circ} - \delta^{\circ})\,d\mu = y_1\mu(G'') - y_2\mu(H'') = 0,$$

10/12 X 27 TR Fl=2 ems

from (28) and (27). Thus it satisfies condition (15).

Next, if $a \in G''$, then $g_{y_1}(a) > x > -\infty$, by (25). Hence $y_1 < c(a) - \delta^{\circ}(a)$, by (21). This means that adding y_1 to $\delta^{\circ}(a)$ on G'' does not violate the feasibility condition $\delta(a) \leq c(a)$. Similarly, if $a \in H''$, then $h_{y_2}(a) < x < \infty$, by (25). Hence $y_2 < \delta^{\circ}(a) - b(a)$, by (22), so that subtracting y_2 on H'' does not violate the feasibility condition $\delta(a) \geq b(a)$. Thus (14) remains satisfied, and $\delta^{\circ\circ}$ is feasible.

Comparing utility from $\delta^{\circ\circ}$ and δ°, we have

12 pt brackets

$$\int_A \left[f(a,\delta^{\circ\circ}(a)) - f(a,\delta^{\circ}(a))\right]\mu(da) = \int_{G''} y_1 g_{y_1}\,d\mu - \int_{H''} y_2 h_{y_2}\,d\mu,$$

from (28), (21), and (22),

$$> \int_{G''} y_1 x\,d\mu - \int_{H''} y_2 x\,d\mu$$

since $G'' \subseteq G_{y_1}$ and $H'' \subseteq H_{y_1}$ and $\mu(G'')$, $\mu(H'')$ are positive,

$$= x\,y_1\mu(G'') - x\,y_2\mu(H'') = 0,$$

from (27). Thus $\delta^{\circ\circ}$ is preferred to δ° under standard ordering of pseudomeasures. The denial of (20) thus contradicts the premise that δ^{o} is unsurpassed. This establishes (20) for $G \subseteq A_\theta$, $H \subseteq A_\theta$.

18 Case (ii). $G \subseteq A_m$, $H \subseteq A_n$, where $m, n \neq 0$:

As above, we assume (20) is false, so that

$$\text{ess sup}\,(g|G) > \text{ess inf}\,(h|H). \qquad (5.4.29)$$

For this to hold, G and H must have positive measure, so that they are atoms. Applying (9) to the function $c-\delta^{o}$ restricted to G, there must be a unique constant, ε_θ, such that

$$\mu\Big[G \cap \{a|c(a) - \delta^{o}(a) = \varepsilon_\theta\}\Big] > 0. \qquad (5.4.30)$$

12 pt brackets

Since $c-\delta^{o}$ is non-negative, $\varepsilon_\theta \geq 0$. In fact, $\varepsilon_\theta > 0$. For if $c(a) = \delta^{o}(a)$, then $g(a) = D_+f(a,\delta^{o}(a)) = -\infty$; hence $\varepsilon_\theta = 0$ would imply $\text{ess sup}\,(g|G) = -\infty$, contrary to (29). A similar argument shows ~~that~~ there is a unique positive constant, η_θ, such that

12 pt brackets

$$\mu\Big[H \cap \{a|\delta^{o}(a) - b(a) = \eta_\theta\}\Big] > 0. \qquad (5.4.31)$$

Label the sets in (30) and (31) by G', H', respectively. These are subatoms of G and H.

A measure ~~which~~ that is atomic and σ-finite is bounded, so that

$$\infty > \mu(G') > 0, \quad \infty > \mu(H') > 0.$$

$\varepsilon_0 \eta_0$ zeros not subscripted

Now take two sequences of positive numbers, beginning with ε_0, η_0, respectively, and ~~and~~ decreasing to zero:

$$\varepsilon_0 > \varepsilon_1 > \varepsilon_2 > \ldots, \ \underline{\lim}\ \varepsilon_n = 0, \quad \eta_0 > \eta_1 > \eta_2 > \ldots, \ \underline{\lim}\ \eta_n = 0;$$

and chosen so that

$$\varepsilon_n \mu(G') = \eta_n \mu(H') \qquad (5.4.32)\ \ \cancel{(32)}$$

for all $n > 0$.

g 15.2 Consider the sequence of functions g_{ε_n}, $n = 1, 2, \ldots$, given by (21), all restricted to the domain G'. The condition $\varepsilon_n < \varepsilon_0$ guarantees ~~that~~ these are all finite. Similarly, consider the sequence h_{η_n}, $n = 1, 2, \ldots$, given by (22), all restricted to the domain H'. These are also all finite, since $\eta_n < \eta_0$.

Applying (9) once again, we find there must be two unique sequences of constants, say $c_1, c_2, \ldots,$ and $b_1, b_2, \ldots,$ such that

$$\mu\left[G' \cap \{a \mid g_{\varepsilon_n}(a) = c_n \ \text{ for all } n = 1, 2, \ldots\}\right] > 0, \qquad (5.4.33)\ \ \cancel{(33)}$$

~~and~~

$$\mu\left[H' \cap \{a \mid h_{\eta_n}(a) = b_n \ \text{ for all } n = 1, 2, \ldots\}\right] > 0. \qquad (5.4.34)\ \ \cancel{(34)}$$

These constants must be finite.

¶ Label the sets in (33) and (34) by G'', H'', respectively. Also let

$$c_o = \lim\inf c_n = \lim\inf g_{\varepsilon_n}(a), \quad \text{~~and~~} \qquad \text{(5.4.35)} \; \cancel{(35)}$$

and

$$b_o = \lim\sup b_n = \lim\sup h_{\eta_n}(a). \qquad \text{(5.4.36)} \; \cancel{(36)}$$

(Equation (35) is valid for all $a \in G''$, (36) for all $a \in H''$. We also have

$$\lim\inf g_{\varepsilon_n}(a) \geq D_+ f(a, \delta^o(a)), \quad \text{~~and~~} \qquad \text{(5.4.37)} \; \cancel{(37)}$$

~~and~~

$$\lim\sup h_{\eta_n}(a) \leq D^- f(a, \delta^o(a)), \qquad \text{(5.4.38)} \; \cancel{(38)}$$

for $a \in G''$, $a \in H''$, respectively. (Inequality (37) results from the fact that D_+ is the limit of infima taken over entire intervals, while the left side of (37) is the limit of infima taken over subsets of these intervals (~~namely~~ viz., the points ε_n). A similar argument establishes (38).

Now the right-hand sides of (37) and (38) are nothing but g and h, respectively. Hence

124

235]

$$c_o \geq \text{ess sup}\,(g|G'') = \text{ess sup}\,(g|G) > \text{ess inf}\,(h|H) = \text{ess inf}\,(h|H'') \geq b_o, \qquad \text{(5.4.39)} \; \cancel{(39)}$$

on one line

from (35), (37), (29), (36), and (38).

Since $c_o > b_o$, ~~there~~ must exist an n' such that $c_{n'} > b_{n'}$. We are now ready to construct a feasible density δ^{oo} ~~which~~ that surpasses δ^o. Let

$$\begin{aligned} \delta^{oo}(a) &= \delta^{o}(a) + \varepsilon_{n'} && \text{if } a \in G''; \\ \delta^{oo}(a) &= \delta^{o}(a) - \eta_{n'} && \text{if } a \in H''; \\ \delta^{oo}(a) &= \delta^{o}(a), && \text{if } a \notin G'' \cup H''. \end{aligned} \qquad \cancel{(40)}\ (5.4.40)$$

align it

This is feasible. First ~~of all~~,

$$\int_A (\delta^{oo} - \delta^{o})\, d\mu = \varepsilon_{n'}\mu(G'') - \eta_{n'}\mu(H'')$$

$$= \varepsilon_{n'}\mu(G') - \eta_{n'}\mu(H') = 0,$$

from (32). Second~~ly~~, for $a \in G''$, $\delta^o(a) + \varepsilon_{n'} < \delta^o(a) + \varepsilon_\theta = c(a)$, so the upper bound constraint $\delta \leq c$ remains satisfied. Similarly, $\delta^o(a) - \eta_{n'} > \delta^o(a) - \eta_\theta = b(a)$, for $a \in H''$, so the lower bound constraint $\delta \geq b$ remains satisfied. This proves feasibility.

Comparing utilities, we obtain

$$\int_A \left[f(a, \delta^{oo}(a)) - f(a, \delta^{o}(a)) \right] \mu(da) = \int_{G''} \varepsilon_{n'} g_{\varepsilon_{n'}}\, d\mu - \int_{H''} \eta_{n'} h_{\eta_{n'}}\, d\mu$$

12 pt brackets

from (40), (21), and (22),

$$= \varepsilon_{n'} c_{n'} \mu(G'') - \eta_{n'} b_{n'} \mu(H'')$$

from (33) and (34),

$$> b_{n'} [\varepsilon_{n'} \mu(G'') - \eta_{n'} \mu(H'')] = 0,$$

since $c_{n'} > b_{n'}$, $\mu(G'') = \mu(G')$, $\mu(H'') = \mu(H')$, and all terms are positive. Thus $\delta^{\circ\circ}$ is preferred to δ° under standard ordering of pseudomeasures. The denial of (20) again leads to contradiction.

g.d. 15.3 # 18

Case (iii). $G \subseteq A_m$, $H \subseteq A_0$, $m \neq 0$:

A0 zero

The proof for this case combines the techniques of the two preceding ~~parts~~ cases, and we ~~shall just~~ merely outline the procedure. As before, we assume (20) is false. Let x be a real number satisfying

$$\text{ess sup } (g|G) > x > \text{ess inf } (h|H). \qquad \text{(5.4.41)} \; \cancel{(41)}$$

Reasoning as in ~~part~~ case (i), we can find a positive (rational) number y_2 such that $\mu(H_{y_2}) > 0$, where $H_{y_2} = H \cap \{a | h_{y_2}(a) < x\}$, and h_{y_2} is given by (22).

ε0 zero

Next, we find the number ε_0 and the set G' as in ~~part~~ case (ii), (30), and again take a positive sequence decreasing to zero:

$$\varepsilon_0 > \varepsilon_1 > \varepsilon_2 > \ldots; \lim \varepsilon_n = 0.$$

As in case (ii), we find a unique sequence of real numbers, $c_1, c_2, \ldots,$ such that $\mu(G'') > 0$, where

$$G'' = G' \cap \{a | g_{\varepsilon_n}(a) = c_n \text{ for all } n = 1, 2, \ldots\}.$$

(cf. (33)). Continuing, we find that (cf. (39))

$$\liminf c_n \geq \operatorname{ess\,sup}\,(g|G). \qquad \text{(5.4.42)}\ \cancel{(42)}$$

Now choose n' so large that

$$\varepsilon_n < y_2\mu(H_{y_2})/\mu(G''), \qquad \text{(5.4.43)}\ \cancel{(43)}$$

for all $n > n'$. This can always be done, since $\lim \varepsilon_n = 0$, and the right side of (43) is positive: ($\mu(G'')$ is finite, since μ is ~~sigma~~ σ-finite and atomic on G''). Inequalities (42) and (41) imply that there is an $n'' > n'$ for which

$$c_{n''} > x. \qquad \text{(5.4.44)}\ \cancel{(44)}$$

Cont gal. 153

400 to go!

Since μ is ~~sigma~~ σ-finite, and non-atomic on H_{y_2}, it takes on all values between 0 and $\mu(H_{y_2})$. Hence there is a measurable subset $H'' \subseteq H_{y_2}$ such that

$$\varepsilon_{n''}\mu(G'') = y_2\mu(H''), \qquad \text{(5.4.45)}\ \cancel{(45)}$$

since $\varepsilon_{n''}$ satisfies inequality (43). The common value in (45) is positive and finite.

We now construct a new feasible density δ^{oo} as follows:

align ¶

$$\begin{aligned}
\delta^{oo}(a) &= \delta^{o}(a) + \varepsilon_{n''} && \text{if } a \in G''; \\
\delta^{oo}(a) &= \delta^{o}(a) - y_2 && \text{if } a \in H''; \\
\delta^{oo}(a) &= \delta^{o}(a) && \text{if } a \notin G'' \cup H''.
\end{aligned}$$

To show feasibility, we have, first, ~~of all,~~

$$\int_A (\delta^{\circ\circ} - \delta^{\circ})\, d\mu = \varepsilon_{n''}\mu(G'') - y_2\mu(H'') = 0,$$

from (45). Arguments already given in ~~parts~~ cases (i) and (ii) show that the bounding constraints $b \leq \delta \leq c$ remain valid for $\delta^{\circ\circ}$. Hence, it is feasible. Comparing utilities:

$$\int_A \left[f(a,\delta^{\circ\circ}(a)) - f(a,\delta^{\circ}(a))\right]\mu(da) = \int_{G''} \varepsilon_{n''} g_{\varepsilon_{n''}}\, d\mu - \int_{H''} y_2 h_{y_2}\, d\mu$$

$$> \varepsilon_{n''} c_{n''}\mu(G'') - y_2\, x\, \mu(H'')$$

$$> x\,[\varepsilon_{n''}\mu(G'') - y_2\mu(H'')] = 0,$$

from (21), (22), (25), (44), and (45). Hence $\delta^{\circ\circ}$ is preferred to δ° under standard ordering of pseudomeasures. This contradiction establishes (20) for ~~part~~ case (iii).

Case (iv) $G \subseteq A_{\theta}$, $H \subseteq A_n$, $n \neq 0$:

This is completely symmetric with ~~part~~ case (iii): One finds a set G_{y_1} as in ~~part~~ case (i), a sequence $b_1, b_2, \ldots$ as in ~~part~~ case (ii), etc. Details are left as an exercise.

Thus (20) has been verified in all four special cases. By the argument ~~which~~ that begins this proof, (20) is now established in general. ~~The proof is complete.~~

This has been a long and tedious proof, and the result itself does not look prepossessing. A few intuitive remarks may be in order, then, to indicate what the lemma says, and why it "should" be true.

The function g represents, roughly, the return per unit increase in investment at various points of A. The function h represents, again roughly, the loss in return per unit dis-investment. If (20) is violated for a pair of disjoint sets G, H, this means there are subsets G', H', of positive measure such that g on G' is higher than h on H'. Then, if we transfer some mass from H' to G', the net gain on the latter set outweighs the net loss on the former, resulting in a new feasible density ~~which~~ that surpasses δ^o. The main burden of the proof just given, ~~in fact,~~ is to find the appropriate subsets, and the appropriate mass to transfer.

We stress the generality of this result. The only special conditions imposed on A, Σ, μ, b, c, or f, aside from measurability, are that μ be ~~sigma~~ σ-finite, f real, and $f(a,\cdot)$ lower semi-continuous for $a \in A_0$. These are very weak conditions, and will nearly always be satisfied in practice.

Our aim now is to use this result to derive the existence of a number that behaves somewhat like a shadow price.

Lemma: Let (A,Σ,μ), f, b, c, and δ^o satisfy the conditions of the preceding lemma, with δ^o unsurpassed for the problem of maximizing (13) subject to (14) and (15). Let the functions g, h: $A \to$ extended reals be defined, as above, by (16)–(19).

Then there exists an extended real number, p°, and a set $E \in \Sigma$, such that

~~(i)~~ E is either an atom or a null set, and

~~(ii)~~
$$g(a) \leq p^{\circ} \leq h(a), \qquad (5.4.46)\ ~~(46)~~$$

for all $a \in A \backslash E$.

(This may be expressed by the statement: Except for at most one atom, (46) is true almost everywhere).

Proof: Let $\{A_{\theta}, A_1, \ldots\}$ be the decomposition of A into non-atomic part A_{θ} and atoms $A_1, A_2, \ldots$. Let $g^n = \text{ess sup } g|A_n$, $h^n = \text{ess inf } h|A_n$, for $n = 0, 1, \ldots$. The preceding lemma then states that

$$g^m \leq h^n \qquad (5.4.47)\ ~~(47)~~$$

for all m, $n = 0, 1, \ldots$ such that $m \neq n$. It gives no information if $m = n$, since then there is no disjointness of A_m, A_n.

First ~~let us~~ we suppose that

$$\sup g^m \leq \inf h^n . \qquad (5.4.48)\ ~~(48)~~$$

Choose any number p° between these bounds. Then (46) is true for all $a \in A_n$ except possibly for a null set $E_n \subseteq A_n$, ~~all~~ $n = 0, 1, \ldots$. Hence (46) is true everywhere except for the set $E = \cup E_n$. Since E is a null set, the lemma is established if (48) is true.

Now let (48) be false. There must be a pair of indices, m' and n' for which $g^{m'} > h^{n'}$. By (47), these indices must be equal: $m' = n'$. But then, for all $m \neq n'$, $n \neq n'$, we have

$$g^m \leq h^{n'} < g^{n'} \leq h^n, \qquad \text{(5.4.49)} \quad \cancel{(49)}$$

by (47) again. Hence, if we exclude the set $A_{n'}$, (48) will be re-established for the rest of A. Thus (46) will be true almost everywhere on $A\setminus A_{n'}$.

0 = zero where inked

It remains to prove only that the anomalous $A_{n'}$ (prime) must be an atom, i.e. ~~that is~~, that $n' \neq 0$. Suppose on the contrary that $g^0 > h^0$. Choose two real numbers, x, y, so that

$$g^0 > x > y > h^0,$$

and let

$$G = A_0 \cap \{a|g(a) > x\}, \quad H = A_0 \cap \{a|h(a) < y\}.$$

As in ~~part~~ (i) of the preceding proof we conclude that G and H are measurable sets. If $\mu(G \cap H) = 0$, then

$$\text{ess sup}\,[g|(G\setminus H)] = g^0 > h^0 = \text{ess inf}\,[h|(H\setminus G)]. \qquad \text{(5.4.50)} \quad \cancel{(50)}$$

The first equality in (50) is obtained by noting that the essential supremum of g on A_0 is the same as on G, since $g^0 > x$. This in turn is the same as on $G\setminus H$, since this just removes a null set, and similarly for the last equality in (50). But (50) contradicts the preceding lemma, since $G\setminus H$ and $H\setminus G$ are disjoint.

If $\mu(G \cap H) > 0$, we can split $G \cap H$ into two pieces of positive measure, F_1 and F_2, since μ is non-atomic on A_0. But then

$$\text{ess sup } (g|F_1) \geq x > y \geq \text{ess inf } (h|F_2).$$

This again contradicts the preceding lemma, since F_1 and F_2 are disjoint. Hence $n' \neq 0$, and the offending set $A_{n'}$ must be an atom. |||

The exceptional case, where (46) fails for some $A_{n'}$, will be referred to as the case of the anomalous atom. Here is a trivial problem in which it arises. Let A consist of just one point: $A = \{a_o\}$, of measure one; let the payoff function for this point be $f(x) = |x|$; the bounds satisfy $b < 0 < c$. Because of the integral constraint (15) there is just one feasible — hence optimal — solution, ~~namely~~ viz. $\delta(a_o) = 0$. The space consists of one atom, and ~~one~~ we ~~verifies~~ verify ~~easily~~ that $g(a_o) = +1$, $h(a_o) = -1$, so (46) cannot be true for a_o. $\delta(a_0)$

The anomalous atom situation is related to, but not identical with, the well-known case in which there can be "increasing marginal returns" on (at most) one alternative project at the optimal allocation.[5] This occurs in the example beginning this section, (1)–(2), in which the second derivative is positive on point a_1. But this latter situation involves "second-order" conditions, while the anomalous atom involves "first-order" conditions.

The number p of this lemma will turn out to behave much like a shadow price. In this connection it is desirable that it be finite. This is not guaranteed, and it is not always possible to find a finite $\underline{p}$ satisfying (46), but there are several simple conditions ~~which~~ That imply ~~that~~ there is such a real number.

One such sufficient condition is that the anomalous atom case occur. For then the inequalities (49) let in daylight between sup $\underline{g}^{\underline{n}}$ and inf $\underline{h}^{\underline{n}}$ (both sup and inf over $\underline{n} \neq \underline{n}'$), and allow us to pick a real number between them.

A second condition ~~which~~ that insures that the $\underline{p}$ in (46) will ~~be~~ be finite is that

0 = 3 where inked

$$\mu\left[\underline{A}_{\Theta} \cap \{\underline{a} \mid \delta^{o}(\underline{a}) < \underline{c}(\underline{a})\}\right] > 0, \qquad \text{(5.4.51)} \; ~~(51)~~$$

~~and~~

12 pt brackets

$$\mu\left[\underline{A}_{\Theta} \cap \{\underline{a} \mid \delta^{o}(\underline{a}) > \underline{b}(\underline{a})\}\right] > 0. \qquad \text{(5.4.52)} \; ~~(52)~~$$

That is, there is a non-atomic set of positive measure on which δ^{o} is strictly below its upper bound, and there is a similar set on which δ^{o} is strictly above its lower bound. For, from their definitions, $\underline{g} > -\infty$ on the set in (51), and h $< \infty$ on the set in (52). Hence

$$-\infty < \underline{g}^{\Theta} \leq \underline{p} \leq \underline{h}^{\Theta} < \infty ,$$

and $\underline{p}$ is finite.

~~Let us~~ We now introduce a differentiability condition. We make the following conventions. First, let $\underline{b}(\underline{a}) < \delta^{o}(\underline{a}) < \underline{c}(\underline{a})$ for a certain point $\underline{a} \in \underline{A}$.

revised ©

$\underline{f}(\underline{a},\cdot)$ is said to be differentiable at the point $\delta^{o}(\underline{a})$ iff

12 pt parens + brackets

$$\lim_{y \to 0} \left[f\left(a, \delta^o(a) + y\right) - f\left(a, \delta^o(a)\right) \right] / y \qquad (5.4.53) \quad \cancel{(53)}$$

exists. (This is the usual definition, except that the values $\pm\infty$ are possible. These occur with a vertical tangent at $\delta^o(a)$, $+\infty$ if $f(a,\cdot)$ is increasing at $\delta^o(a)$, $-\infty$ if it is decreasing.) Next, if $\delta^o(a) = c(a)$, we say that $f(a,\cdot)$ is differentiable at $\delta^o(a)$ iff the limit (53) exists when $y \to 0$ through negative values, and this limit $= -\infty$. Finally, if $\delta^o(a) = b(a)$, we say that $f(a,\cdot)$ is differentiable at $\delta^o(a)$ iff the limit exists when $y \to 0$ through positive values, and this limit $= +\infty$. When f is differentiable, the value of the limit in (53) is called the derivative, and denoted $Df(a, \delta^o(a))$.

The next result seems sufficiently interesting to be labeled a theorem.

gal 4

D

Theorem: Given measure space (A, Σ, μ), μ ~~sigma~~ σ-finite, and measurable functions $f: A \times \text{reals} \to \text{reals}$, $b, c: A \to$ extended reals, with $f(a,\cdot)$ lower semicontinuous on the nonatomic part of A; let δ^o be unsurpassed for the problem of maximizing

$$\int_A f(a, \delta(a))\, \mu(da)$$

over measurable functions $\delta: A \to$ reals ~~which~~ that satisfy

$$b \le \delta \le c$$

and

$$\int_A \delta \, d\mu = 0.$$

Let g, h: $A \to$ extended reals be defined as usual by (16)–~~(18)~~ (19).

In addition, let there be a set E of positive measure such that $f(a,\cdot)$ is differentiable at $\delta^{o}(a)$, for all $a \in E$. Then

~~(i)~~ $Df(a,\delta^{o}(a))$ is equal to a constant, p^{o}, almost everywhere on E, and

~~(ii)~~ p^{o} is the unique number ~~which~~ that satisfies

$$g(a) \leq p \leq h(a) \qquad (5.4.54)\ \text{~~(54)~~}$$

almost everywhere.

Proof: Let $\{A_0, A_1, A_2, \ldots\}$ be a decomposition of A, with μ non-atomic on A_0, and atomic on A_n, $n = 1, 2, \ldots$. Since $\mu(E) > 0$, $\mu(E \cap A_n) > 0$ for at least one value of $n = 0, 1, \ldots$.

First, take the case when $\mu(E \cap A_{n'}) > 0$ for some $n' \neq 0$. Then, for any $a \in A_{n'}$, for which $Df(a,\delta^{o}(a))$ exists, we have

$$g(a) = D_{+}f(a,\delta^{o}(a)) = Df(a,\delta^{o}(a)) = D^{-}f(a,\delta^{o}(a)) = h(a), \qquad (5.4.55)\ \text{~~(55)~~}$$

from the definitions of these functions. (If $\delta^{o}(a) = c(a)$, the common value in (55) is $-\infty$; if $\delta^{o}(a) = b(a)$, the common value is $+\infty$).

Now $Df(a,\delta^{o}(a))$, as a function of a with domain E, is measurable, since it is the limit of a sequence of measurable functions g_{y_i} (given by (21)) (or h_{y_i} given by (22)) as $y_i \to 0$ through positive values. Also $A_{n'}$, hence $E \cap A_{n'}$, is an atom. Invoking the lemma (9), we find that $Df(a,\delta^{o}(a))$ — hence g and h — is equal to a constant, p^{o}, almost everywhere on $E \cap A_{n'}$.

It follows that

$$\underline{\text{ess}}\ \underline{\text{sup}}\ (g|A_{n'}) = \underline{\text{ess}}\ \underline{\text{inf}}\ (h|A_{n'}) = \underline{p}^{\circ}.$$

Using the abbreviations $g^n = \text{ess sup}\,(g|A_n)$, $h^n = \text{ess inf}\,(h|A_n)$, it follows from (20) that, for all $n_1 \neq n'$, $n_2 \neq n'$,

$$g^{n_1} \leq h^{n'} = p^{\circ} = g^{n'} \leq h^{n_2}, \qquad (5.4.56)\ \cancel{(56)}$$

so that

$$\sup g^n = p^{\circ} = \inf h^n, \qquad (5.4.57)\ \cancel{(57)}$$

and (54) is true almost everywhere.

Next take the case when $\mu(E \cap A_{\theta}) > 0$. For any $a \in A_{\theta}$ for which $Df(a, \delta^{\circ}(a))$ exists we have

$$g(a) \geq Df(a, \delta^{\circ}(a)) \geq h(a). \qquad (5.4.58)\ \cancel{(58)}$$

For, on A_{θ}, g is the supremum of the functions g_y while Df is their limit as $y \to 0$; and h is the infimum of the functions h_y while Df is again their limit as $y \to 0$. It follows from (58) that $g^{\theta} \geq h^{\theta}$. But from the preceding lemma (46), the opposite inequality is also true, so $g^{\theta} = h^{\theta}$ (= p°, say). It then follows from (58) that $Df(a, \delta^{\circ}(a)) = p^{\circ}$ almost everywhere on A_{θ}.

Also, the same argument (56) establishes (57) again. Hence (57) is established in all cases for a unique number p°, and (54) is true for this number. The arguments just given show that $Df(a, \delta^{\circ}(a))$ is equal to this number p° almost everywhere on $E \cap A_n$ for all $n = 0, 1, \ldots$, hence almost everywhere on $\underline{E}$.

Note that the anomalous atom case cannot arise here. It is precluded by the condition of differentiability on a set of positive measure. This theorem yields a third simple sufficient condition for p^o to be finite, ~~namely,~~ viz. that there be a set of positive measure on which $Df(a,\delta^o(a))$ exists and is finite. For $Df(a,\delta^o(a)) = p^o$ almost everywhere, and the condition just given insures that p^o is finite.

Finally, we want to show that p^o acts as a shadow price, ~~that is~~ i.e., that $\delta^o(a)$ maximizes

$$f(a,x) - p^o x \qquad (5.4.59)\ ~~(59)~~$$

over all real x satisfying $b(a) \leq x \leq c(a)$, for almost all $a \in A$. Here (59) may be interpreted as giving the "payoff density" $f(a,x)$ minus the "resource cost", $p^o x$.

We make a special convention as to the meaning of "maximization" in case p^o is infinite in (59). Namely, if $p^o = -\infty$, then "maximizing" ~~is to~~ mean taking x as large as possible, so that $x = c(a)$ is the maximizer of (59) (if $c(a)$ is finite). And if $p^o = +\infty$, then x must be taken as small as possible, so that, if $b(a)$ is finite, ~~then~~ $x = b(a)$ is the maximizer.

For finite p^o, the condition that $\delta^o(a)$ maximize (59) will be recognized as the sufficient condition for optimality given by (3.2) ~~(2) of section 3~~, specialized to the particular feasibility conditions of the problem we are studying in this section.[6]

This suggests the question: D ~~does~~ (3.2) ~~(2) of section 3~~ still suffice for "bestness" if the $\underline{p}^{o}$ appearing in it is infinite, and we interpret "maximization" according to the convention just mentioned? The answer is yes, for any arbitrary feasibility conditions that include

183 20 61

$$\int_{\underline{A}} \delta \, d\mu = 0. \qquad (5.4.60) \; \cancel{(60)}$$

(# M) We ~~Let us~~ demonstrate this. The feasible set, M, consists of measurable functions $\delta: \underline{A} \to$ reals, all of which satisfy (60). The sufficient condition is that, for all δ,

$$\underline{f}(\underline{a},\delta^{o}(\underline{a})) - \underline{p}^{o}\delta^{o}(\underline{a}) \geq \underline{f}(\underline{a},\delta(\underline{a})) - \underline{p}^{o}\delta(\underline{a}) \qquad (5.4.61) \; \cancel{(61)}$$

for almost all $\underline{a} \in \underline{A}$. Using our convention, if $\underline{p}^{o} = +\infty$, (61) reduces to

$$\delta^{o}(\underline{a}) \leq \delta(\underline{a}). \qquad (5.4.62) \; \cancel{(62)}$$

But (60) implies that

161 20 104

$$\int_{\underline{A}} (\delta - \delta^{o}) \, d\mu = 0, \qquad (5.4.63) \; \cancel{(63)}$$

and this together with (62) means that $\delta = \delta^{o}$ almost everywhere. Hence δ^{o} is best because the feasible set is trivial: ~~Apart from null sets,~~ δ^{o} is the only feasible solution. Similarly, if $\underline{p}^{o} = -\infty$, (61) reduces to (62) reversed, and the same argument applies. This completes the demonstration.

\# 18 We are thus on the verge of establishing a necessary and sufficient condition for optimality. One consequence of this will be that "unsurpassed" and "best" solutions coincide for this problem. For, starting from an unsurpassed δ^o, we derive a condition ~~which~~ that suffices for δ^o to be best.

Now to establish the shadow price condition (59). For this we need an extra condition on the atoms of μ. The trouble is that g and h are defined on the atoms in such a way that they depend only on the immediate neighborhood of $\delta^o(a)$, whereas (59) asserts something about the entire range $[b(a), c(a)]$. The assumption of concavity will bridge the gap.

We define this abstractly. Let f be a real-valued function whose domain is a real interval $[b,c]$. (Endpoints may or may not be included, and $b = c$ is possible); f is said to be concave iff, for any numbers x, y in its domain, and any number t in the interval $[0,1]$, we have

$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y).$$

A concave function may be shown to be continuous, except possibly at the endpoints of its domain, where a "sudden" downward jump is possible. Thus a concave function (defined on an interval) is always lower semi-continuous. We also state without proof the following well-known facts about concave functions:

$$\frac{f(x+y_1) - f(x)}{y_1} \leq D_+f(x) \leq D^-f(x) \leq \frac{f(x) - f(x-y_2)}{y_2}, \qquad (5.4.24)\ \cancel{(64)}$$

for any positive real y_1, y_2 and real x such that $x - y_2$, x, and $x + y_1$ are in the domain of f. (Our special conventions concerning D_+ and D^- insure that (64) holds even if x is an endpoint of the domain.)

Let us return to the problem in hand. In stating that $f(a,\cdot)$ is concave, we follow our standing convention of taking the domain of this function to be restricted to $[b(a), c(a)]$, ($b(a)$ to be included iff it is finite, and similarly for $c(a)$).

The following theorem may be taken to be the main result of this section.

Theorem: Let (A,Σ,μ) be a measure space, with μ ~~sigma~~ σ-finite; let $f\colon A \times \text{reals} \to \text{reals}$, and $b, c\colon A \to \text{extended reals}$ be measurable; let $f(a,\cdot)$ be lower semi-continuous on the non-atomic part of A, and concave on the rest of A; let δ^o be feasible for the problem of maximizing

$$\int f(a,\delta(a))\mu(da) \qquad \text{(5.4.65)} \quad (65)$$

over measurable functions $\delta\colon A \to$ reals, which satisfy

$$b \leq \delta \leq c \qquad \text{(5.4.66)} \quad (66)$$

~~and~~

$$\int_A \delta \, d\mu = 0. \qquad \text{(5.4.67)} \quad (67)$$

Then the following conditions are logically equivalent:

(i) δ^o is unsurpassed for this problem;

(ii) δ^o is best for this problem;

(D)

(iii) there is an extended real number, p^o, and a null set E, such that $\delta^o(a)$ maximizes

$$f(a,x) - p^o x \qquad \text{(5.4.68)} \; \cancel{(68)}$$

over all real x in the closed interval $[b(a), c(a)]$, for all $a \in A \backslash E$.

Proof: (iii) implies (ii) is already contained in the sufficiency theorem ~~(1)-(3) of section 3~~ (3.1)–(3.3), plus the argument of (61)–(63). Condition (ii) implies (i) by definition. It remains to show that (i) implies (iii).

Let $\{A_0, A_1, \ldots\}$ be a decomposition, so that μ is nonatomic and f lower semicontinuous on A_0 (zero), while μ is atomic and f concave on each A_n, $n \neq 0$.

~~First~~ We show that concavity precludes the occurrence of an anomalous atom in (46). Choose any atom A_n. ~~First of all,~~ $g|A_n$ is measurable (without) ~~(g is given by (16), (18))~~. This follows from the fact that, by concavity, $g_y(a)$ given by (21) is nonincreasing in y, for fixed $a \in A_n$. Hence the lower right derivate $D_+f(a,\delta^o(a))$, which is $g(a)$, is the limit of any sequence $g_{y_i}(a)$, y_i going to zero through positive values. $g|A_n$ is thus the limit of a sequence of measurable functions $g_{y_i}|A_n$, and so is itself measurable. A similar argument establishes the measurability of $h|A_n$ (with h ~~is~~ given by (17) and (19)).

The lemma (9) then implies that g and h are constant almost everywhere on A_n. These constants must equal g^n, the essential supremum of $g|A_n$, and h^n, the essential infimum of $h|A_n$.

There is a point $a \in A_n$ such that $g(a) = g^n$ and $h(a) = h^n$, since these relations hold almost everywhere on A_n, and $\mu(A_n) > 0$. But $g(a)$ and $h(a)$ equal $D_+f(a,\delta^o(a))$ and $D^-f(a,\delta^o(a))$, respectively. (64), middle, then implies that $g^n \le h^n$.

This is true for every $n = 1, 2, \ldots$. We also have $g^\theta \le h^\theta$ from (46), and $g^{n_1} \le h^{n_2}$ for all $n_1 \ne n_2$, from (20). Hence $\sup g^m \le \inf h^n$ (both taken over all $m, n = 0, 1 \ldots$), and (46) is true almost everywhere. There is no anomalous atom.

Thus there is an extended real number, p^o, such that

$$g(a) \le p^o \le h(a) \qquad (5.4.69)\ \cancel{(69)}$$

for almost all $a \in A$. We now show that this p^o satisfies the condition (iii) almost everywhere.

First, suppose $p^o = -\infty$. Then $g = -\infty$ almost everywhere, from (69). Let $g(a) = -\infty$ for some $a \in A\backslash A_\theta$. $f(a,\cdot)$ is then concave, and it follows that there cannot be any positive y_1 satisfying (64), left. That is, $\delta^o(a)$ is at the upper limit: $\delta^o(a) = c(a)$. If $g(a) = -\infty$ for some a in the non-atomic part A_θ, then it follows from the definition of g, (16), that $\delta^o(a)$ is again at the upper limit. Thus $\delta^o = c$ almost everywhere. But, by our convention for p^o infinite, this is precisely the condition that δ^o maximize (68) for $p^o = -\infty$, almost everywhere. Hence (iii) is established in this case.

Next, suppose $p^o = +\infty$. Then $h = +\infty$ almost everywhere. An argument similar to ~~the~~ that one just given shows that $\delta^o = b$ almost everywhere, which is the condition for (iii) to be satisfied when $p^o = +\infty$.

It remains to establish (iii) when p^o is finite. In this case, the condition that $\delta^o(a)$ maximize (68) is equivalent to the following double inequality:

18

$$\frac{f(a,\delta^o(a) + y_1) - f(a,\delta^o(a))}{y_1} \le p^o \le \frac{f(a,\delta^o(a)) - f(a,\delta^o(a) - y_2)}{y_2}, \quad (70) \quad (5.4.70)$$

which must hold for all positive y_1 such that $\delta^o(a) + y_1 \le c(a)$, and for all positive y_2 such that $\delta^o(a) - y_2 \ge b(a)$. We now show that (69) implies (70). If $a \in A \setminus A_0$, then $f(a,\cdot)$ is concave, and the implication follows at once from (64).

Finally, let $a \in A_0$. $g(a)$ is defined by (16) as the supremum of the left ~~terms in~~ side of (70), as y_1 varies over the open interval $(0, c(a) - \delta^o(a))$. Similarly, $h(a)$ is, by (17), the infimum of the right ~~terms in~~ side of (70), as y_2 ranges over $(0,\delta^o(a) - b(a))$. (Inequality) (69) then implies that (70) holds for all interior points of $[b(a), c(a)]$. But it must then hold for the endpoints as well, because $f(a,\cdot)$, being lower semi-continuous, makes no sudden upward jump at $b(a)$ or $c(a)$. This establishes the implication in general.

Since (69) holds for almost all $a \in A$, so does (70). This completes the proof that condition (i) implies condition (ii).

Thus, under the special assumptions made concerning f, a necessary and sufficient condition that feasible δ° be unsurpassed, or best, for the problem of maximizing (65) subject to (66) and (67), is that ~~there exist~~ a "shadow price", p°, exists under which (except for a null set), for each "project" $a \in A$ separately, $\delta^{\circ}(a)$ is chosen to maximize the "payoff" $f(a,x)$, net of the "resource cost" $p^{\circ}x$.

This result is important. First ~~of all~~, it suggests an efficient method for finding an optimal solution (if there is one). Namely, choose an arbitrary number p, and, for each $a \in A$, choose $\delta^{\circ}(a)$ to maximize $f(a,x) - px$ over the feasible interval [b(a), c(a)], disregarding the total resource constraint (67). If, by chance, (67) is satisfied by this process, we have found an optimal solution. If not, adjust p to a new value, p', and try again.

How should p be adjusted? It is easily seen from (68) that the maximizing value of x is a non-increasing function of p. (We are implicitly assuming, for simplicity, that there is a unique maximizer of (68) for each p and $a \in A$). Hence, if total resource availability is exceeded by the trial solution, raise the tentative shadow price, p, and lower it in the opposite case. This simple monotonic relation between p and total resource demand makes it easy to "zero in" on the proper p (again, assuming that there *is* an optimal solution). The necessity of the shadow price condition also guarantees that we will not overlook the optimal solution by this procedure.

Furthermore, the shadow price condition suggests institutional arrangements for arriving at an optimal solution. For example, if one overall organization is responsible for this allocation, separate divisions might be responsible for separate subsets of projects $a \in A$. The "head office" might dictate the tentative shadow price to the divisions, note the consequent resource demand, adjust the price accordingly, etc.[7] Going a step further, the free market itself is an institutional mechanism for carrying out the price-adjustment process discussed above.

Two special cases in which the results of the preceding theorem are valid may be noted. The first is when μ is non-atomic (as well as sigma-finite), and $f(a,\cdot)$ is lower semi-continuous for all $a \in A$. The second is when $f(a,\cdot)$ is concave for all $a \in A$ (with no assumptions on μ other than sigma-finiteness). The validity of the first case is obvious: since μ has no atomic part, no concavity assumption is needed. The validity of the second follows from the fact that a concave function is lower semi-continuous, so that if $f(a,\cdot)$ is concave everywhere, the premises of the theorem are certainly fulfilled.

Finally, we recall that the condition (67) is much less narrow than it appears. We can ~~in fact~~ formulate an apparently much more general theorem, which falls out as an immediate corollary.

Theorem: Let all the premises of the preceding theorem be fulfilled, except that the feasibility condition (67) is replaced by

$$L_o \leq \int_A \delta \, d\mu \leq L^o .$$

(Here L_o and L^o are two extended real numbers. $\int_A \delta \, d\mu$ is still required to be finite, however). Then the following are equivalent:

(i) δ^o is unsurpassed for this problem;

(ii) δ^o is best for this problem;

(iii) there is an extended real number, p^o, and a null set E, such that $\delta^o(a)$ maximizes

$$f(a,x) - p^o x \qquad (5.4.71)$$

over all real x in the closed interval $[b(a), c(a)]$, for all $a \in A \backslash E$. Furthermore, if $\int_A \delta^o d\mu > L_o$, then $p^o \geq 0$; and if $\int_A \delta^o d\mu < L^o$, then $p^o \leq 0$.

Proof: (ii) implies (i) by definition, and (iii) implies (ii) by a sufficiency theorem already given (cf. (3.5)–(3.6)). To show that (i) implies (iii) we add a point z_o to A and transform this problem into one of the preceding type by the recipe given in section 2.

The singleton set $\{z_o\}$ is an atom, since $\mu'\{z_o\} = 1 > 0$; $f(z_o,\cdot)$ is identically zero, and this is a concave function, so the premises are fulfilled on $A' = A \cup \{z_o\}$. By the preceding theorem, ~~there~~ exists a number, p^o, such that $\delta^{o\prime}(a')$ maximizes (71), for almost all $a' \in A'$. In particular, it must maximize (71) for $a' = z_o$, since $\mu\{z_o\} > 0$. Since $f(z_o,\cdot)$ is identically zero, (71) reduces to:

$$-p^o x. \qquad \text{(5.4.72)} \; \cancel{(72)}$$

~~Now,~~ if $p^o > 0$, then the maximizer of (72) is as small as possible: $\delta^{o\prime}(z_o) = b'(z_o)$. But, recalling the translation recipe, (2.9) ~~(9) of section 2~~, this is ~~simply~~:

$$-\int_A \delta^o d\mu = -L^o. \qquad \text{(5.4.73)} \; \cancel{(73)}$$

Similarly, if $p^o < 0$, then we must have $\delta^o(z_o) = c'(z_o)$, which is to say, by (2.9) again ~~(9) of section (2) again~~,

$$-\int_A \delta^o d\mu = -L_o. \qquad \text{(5.4.74)} \; \cancel{(74)}$$

Equations (73) and (74) give the two extra conditions on p^o.

Thus, if δ^o is optimal for this more general problem, we get a shadow price condition of the same type as above, with an extra sign condition on p^o, depending on where $\int_A \delta^o\, d\mu$ is located in the interval $[L_o, L^o]$. The economic interpretation of these sign conditions is the same as in section 3. 5.

5.5. Existence of Feasible Solutions

~~Up to~~ Until now we have been examining conditions ~~which~~ that imply, or are implied by, the fact that a given feasible solution δ^o is optimal. We now take up a different task, that of proving ~~that~~ an optimal solution exists for a given problem. First, we start with the simpler task (i.e., simpler~~, that is,~~ for a given problem) of proving ~~that~~ at least one feasible solution exists. Even this is by no means trivial for the problem with which we have been dealing, characterized by constraints ~~(66)~~ (4.66) ~~and~~ (4.67) ~~(67) of the preceding section.~~

Theorem: Let $(\underline{A},\Sigma,\mu)$ be a measure space, with μ ~~sigma~~ σ-finite; let $\underline{b}$, $\underline{c}:\underline{A} \to$ extended reals be measurable functions. The following conditions are logically equivalent:

(i) ~~There~~ exists a measurable function $\delta:\underline{A} \to$ reals such that

$$\underline{b} \leq \delta \leq \underline{c}, \qquad \text{(5.5.1)}$$

~~and~~

$$\int_{\underline{A}} \delta \, d\mu = 0; \qquad \text{(5.5.2)}$$

(ii) $\underline{b} \leq \underline{c}$, $\underline{b} < \infty$, $\underline{c} > -\infty$, and

$$\int_{\underline{A}} \underline{b} \, d\mu \leq 0 \leq \int_{\underline{A}} \underline{c} \, d\mu. \qquad \text{(5.5.3)}$$

Proof: That (i) implies (ii) is obvious. Conversely, assume (ii). First, because μ is ~~sigma~~ σ-finite, ~~there~~ exists a **positive** measurable function $\underline{k}:\underline{A} \to$ reals such that

$$\infty > \int_A k \, d\mu \geq 0. \qquad (5.5.4)\ \text{~~(4)~~}$$

If $\mu = 0$, take $k = 1$ everywhere. Otherwise, let $\{G_1, G_2, \ldots\}$ be a countable measurable partition with $\infty > \mu(G_n) > 0$ for all n. On G_n let k be equal to the constant $1/[2^n \mu(G_n)]$. This function fulfills the stated conditions.

Next define the *median function* m as the one ~~which~~ that picks the middle in size order of three extended real numbers; thus $m(3, -2, \infty) = 3$, $m(-\infty, \infty, 17) = 17$, etc.

~~Next,~~ Then for each real number x define the function $m_x: A \to$ extended reals by

$$m_x(a) = m(b(a), c(a), xk(a)).$$

Because $b \leq c$, ~~one finds~~ we find that

$$m_x = \max[b, \min(c, xk)] = \min[\max(b, xk), c]. \qquad (5.5.5)\ \text{~~(5)~~}$$

m_x is thus measurable and also ~~real-valued,~~ finite, since $b(a)$ and $c(a)$ are never both infinite of the same sign. We now show that if

$$\int_A b \, d\mu < 0 < \int_A c \, d\mu \qquad (5.5.6)\ \text{~~(6)~~}$$

then m_x will be a feasible solution for (1) and (2), for some real number x.

(If (6) is false, then either $\int_A b \, d\mu = 0$ or $\int_A c \, d\mu = 0$. In either case we get an immediate feasible solution. In the first case, ~~for example~~ e.g., set $\delta(\underline{a}) = \underline{b}(\underline{a})$ whenever $\underline{b}(\underline{a}) > -\infty$; and, on the null set where $\underline{b} = -\infty$, choose $\delta = \underline{\min}[\underline{c}, 0]$. Hence finding a feasible solution in case (6) will prove the theorem).

First of all, by (5),

$$\underline{b} \leq \underline{m}_{\underline{x}} \leq \underline{c}.$$

Next we show that $\int_{\underline{A}} \underline{m}_{\underline{x}}\, d\mu$ exists and is finite for all real $\underline{x}$. We have

$$-(\underline{c}^{-} + |\underline{xk}|) \leq \underline{\min}(\underline{c}, \underline{xk}) \leq \underline{m}_{\underline{x}} \leq \underline{\max}(\underline{b}, \underline{xk}) \leq \underline{b}^{+} + |\underline{xk}|.$$

Also, $-\int_{\underline{A}} (\underline{c}^{-} + |\underline{xk}|)\, d\mu > -\infty$, and $\int_{\underline{A}} (\underline{b}^{+} + |\underline{xk}|)\, d\mu < \infty$, because of (3) and (4), which shows that $\int_{\underline{A}} \underline{m}_{\underline{x}}\, d\mu$ is finite.

Next we show that $\int_{\underline{A}} \underline{m}_{\underline{x}}\, d\mu$, as a function of $\underline{x}$, is continuous. Let $\underline{x}_1, \underline{x}_2, \ldots$ be a sequence, either increasing or decreasing, whose limit is the real number $\underline{x}$. By the monotone convergence theorem we have

$$\underline{\lim_{n}} \int_{\underline{A}} \underline{m}_{\underline{x}_{\underline{n}}}\, d\mu = \int_{\underline{A}} \underline{\lim_{n}}\, \underline{m}_{\underline{x}_{\underline{n}}}\, d\mu = \int_{\underline{A}} \underline{m}_{\underline{x}}\, d\mu,$$

proving continuity.

If $\underline{x}_1, \underline{x}_2, \ldots$ is a sequence increasing to $+\infty$, then $\underline{m}_{\underline{x}}$ increases to $\underline{c}$, so

$$\lim_{n} \int_A m_{x_n}\,d\mu = \int_A c\,d\mu > 0.$$

If $x_1, x_2, \ldots$ is a sequence decreasing to $-\infty$, then m_x decreases to b, so

$$\lim_{n} \int_A m_{x_n}\,d\mu = \int_A b\,d\mu < 0.$$

(Both these results are again by monotone convergence. The inequalities are from (6).)

Hence for sufficiently large real x, $\int_A m_x\,d\mu$ is positive; and for sufficiently small real x, it is negative. Since it is continuous, there must then be an x-value for which $\int_A m_x\,d\mu = 0$, and this m_x is feasible. Hence (ii) implies (i). □

As usual, this theorem has an immediate generalization:

Theorem: Let (A, Σ, μ) be a measure space, with μ ~~sigma~~ σ-finite, let $b, c: A \to$ extended reals be measurable, and let L_o, L^o be two extended real numbers. The following are equivalent:

(i) ~~There exists~~ a measurable function $\delta: A \to$ reals exist such that

$$b \leq \delta \leq c,$$

$\int_A \delta\,d\mu$ is finite, and

$$L_o \leq \int_A \delta\,d\mu \leq L^o$$

(ii) $\infty > b \leq c > -\infty$, $\infty > L_o \leq L^o > -\infty$,

$$\int_A b\,d\mu \leq L^o, \text{ and } \int_A c\,d\mu \geq L_o.$$

This is an easy corollary of the preceding theorem, by the now familiar procedure of transforming this problem into one for which (2) holds. We leave the details as an exercise.

5.6. Existence of Optimal Solutions

We now come to the much more difficult problem of proving the existence of optimal solutions. A number of assumptions will be made ~~which~~ that are more restrictive than those made ~~up to now~~ previously. In particular, the bounding functions, b and c, will have **finite** integrals, and the payoff functions f(a,·) will be **continuous** (as usual, this refers to the interval [b(a), c(a)]). Even so, much work is involved.

We ~~shall~~ first prove existence under the assumption that μ is non-atomic. The basic procedure is to find functions satisfying the sufficient condition for optimality, and then show that one ~~of these~~ is feasible. Next we go to the opposite case where μ is ~~sigma~~ σ-atomic, using an entirely different procedure. Finally, we combine these results to prove existence under general (~~sigma~~ σ-finite) μ.

Standard ordering of pseudomeasures is still used for ordering utilities, and existence is proved for **best** solutions.

Theorem: Let (A,Σ,μ) be a measure space, with μ ~~sigma~~ σ-finite and non-atomic; let b,c:A → reals, and f:A × reals → reals, be measurable; let f(a,·) be continuous for all a ∈ A. Assume

$b \leq c$, and let L be a real number such that

$$-\infty < \int_A b \, d\mu \leq L \leq \int_A c \, d\mu < \infty. \tag{5.6.1}$$

Then the problem: Maximize

$$\int f(a, \delta(a)) \mu(da)$$

over measurable functions $\delta : A \to$ reals, subject to

$$b \leq \delta \leq c, \tag{5.6.2}$$

~~and to~~

$$\int_A \delta \, d\mu = L, \tag{5.6.3}$$

has a *best* solution.

Proof: For each $a \in A$, and each extended real number p, let $E_{a,p}$ be the set of real numbers x ~~which~~ that maximize the expression

$$f(a,x) - px \tag{5.6.4}$$

over the closed interval $[b(a), c(a)]$. By our convention concerning infinite p-values, we have $E_{a,\infty} = \{b(a)\}$, and $E_{a,-\infty} = \{c(a)\}$. Since (4) is continuous in x (for finite p), and the maximization is over a closed bounded interval, the sets $E_{a,p}$ are in all cases nonempty, closed, and bounded. Hence they themselves have a minimum value and a maximum value for all (a,p). Define the functions $\beta, \gamma : A \times$ extended reals $\to$ reals by

$$\beta(a,p) = \min E_{a,p},$$

$$\gamma(a,p) = \max E_{a,p}. \qquad (5.6.5)\ ~~(5)~~$$

Let $-\infty < p_1 < p_2 < \infty$. We have the following chain of relations:

$$c(a) = \gamma(a,-\infty) = \beta(a,-\infty) \geq \gamma(a,p_1) \geq \beta(a,p_1)$$

$$\geq \gamma(a,p_2) \geq \beta(a,p_2) \geq \gamma(a,\infty) = \beta(a,\infty) = b(a). \qquad (5.6.6)\ ~~(6)~~$$

The equalities in (6) follow from $E_{a,\infty} = \{b(a)\}$, $E_{a,-\infty} = \{c(a)\}$. The middle inequality in (6) is the only one that needs proving. ~~In fact,~~ From the definition of $E_{a,p}$ we have

$$f(a,\beta(a,p_1)) - p_1\beta(a,p_1) \geq f(a,\gamma(a,p_2)) - p_1\gamma(a,p_2),$$

~~and~~

$$f(a,\gamma(a,p_2)) - p_2\gamma(a,p_2) \geq f(a,\beta(a,p_1)) - p_2\beta(a,p_1).$$

Adding these two inequalities and simplifying, we get $\beta(a,p_1) \geq \gamma(a,p_2)$. This establishes (6). Thus, for fixed $a \in A$, $\beta(a,\cdot)$ and $\gamma(a,\cdot)$ are non-increasing functions.

Next, we show that $\gamma(a,\cdot)$ is continuous from the left; ~~that is~~ i.e., if $p_1, p_2, \ldots$ is an increasing sequence whose limit is p_o (possibly $+\infty$), then the limit of $\gamma(a,p_n)$ is $\gamma(a,p_o)$. First let p_o be finite. The sequence $y_n = \gamma(a,p_n)$ is non-increasing, hence it has a limit $y_o \geq \gamma(a,p_o)$. We must show ~~that~~ this is an equality; ~~and~~ to do this it suffices to prove that y_o maximizes (4), since $\gamma(a,p_o)$ is the largest number

~~which~~ that does so. ~~No~~

Now (4) is jointly continuous in x and p ($a \in A$ is fixed), hence

$$\lim_{n\to\infty} [f(a,y_n) - p_n y_n] = f(a,y_o) - p_o y_o. \qquad (5.6.7)\ \cancel{(7)}$$

For any $x \in [b(a), c(a)]$ we have $f(a,y_n) - p_n y_n \geq f(a,x) - p_n x$, since y_n maximizes (4). Hence, by (7),

$$f(a,y_o) - p_o y_o \geq \lim_{n\to\infty} [f(a,x) - p_n x] = f(a,x) - p_o x.$$

Thus y_o maximizes (4) for $p = p_o$, so $\gamma(a,p)$ is continuous from the left for any finite p.

Next let $p_o = +\infty$. Since $f(a,\cdot)$ is continuous on the closed bounded interval $[b(a), c(a)]$, it has a finite upper bound, N. For any $\varepsilon > 0$, choose p finite, and greater than $[N - f(a,b(a))]/\varepsilon$. Then for any $c(a) \geq x \geq b(a) + \varepsilon$, we have

$$p[x - b(a)] \geq p\varepsilon > N - f(a,b(a)) \geq f(a,x) - f(a,b(a)),$$

so that

$$f(a,b(a)) - pb(a) > f(a,x) - px. \qquad (5.6.8)\ \cancel{(8)}$$

Inequality (8) shows that no such x can maximize (4), hence $\gamma(a,p) < b(a) + \varepsilon$. Thus if $p_1, p_2, \ldots$ increases without bound, $\gamma(a,p_n)$ approaches $b(a) = \gamma(a,\infty)$. This proves that $\gamma(a,\cdot)$ is continuous from the left. A similar argument shows that $\beta(a,\cdot)$ is <u>continuous from the right</u>.

18 Next, we show that, for fixed p, $\gamma(a,p)$ is a measurable function of a. We split A into two measurable pieces and consider each separately. On the set $\{a \mid b(a) = c(a)\}$, $\gamma(\cdot,p) = b = c$ for any p, and so is measurable.

~~Now~~ consider the complementary set $E = \{a \mid b(a) < c(a)\}$. To show that $\gamma(\cdot,p)$ is measurable, it suffices to show that the sets $E \cap \{a \mid \gamma(a,p) < y\}$ are all measurable as y ranges over the rational numbers. Now, for fixed a, p, (4) is a continuous, hence lower semi-continuous, function of x. Hence its supremum on any interval $[b,c]$ (with or without the endpoints, and $b < c$) equals its supremum over the rational numbers on that interval ~~(see (4.6))~~ ((6) of section 4). One then verifies that, for any rational y,

$$E \cap \{a \mid \gamma(a,p) < y\} =$$

$$E \cap \left\{ a \,\middle|\, \begin{array}{l} \sup\{f(a,x) - px \mid x \text{ rational}, x < y\} \\ \qquad > \sup\{f(a,x) - px \mid x \text{ rational}, x \geq y\} \end{array} \right\}. \qquad (9)$$

In (9) a and p are held fixed, and the two sups are taken over x as indicated. For this formula only, we define $f(a,x)$ to be $-\infty$ if x is not in the closed interval $[b(a), c(a)]$, and we note that, for fixed x, the function $f(\cdot,x)$ thus defined is measurable. In verifying (9) there are five cases to consider, depending on whether y is in the interval $[b(a), c(a)]$, at an endpoint, or on either side of it. ~~Note that~~ this interval is non-degenerate for $a \in E$, hence it always contains a rational point. We omit details.

Since the rational numbers are countable, the two sups in (9) are over a countable number of measurable functions, hence they are themselves measurable functions of $\underline{a}$ on $\underline{E}$ ($\underline{p},\underline{y}$ fixed). Hence the right side of (9) is a measurable set. This proves $\gamma(\cdot,\underline{p})$ is measurable. A similar argument proves that $\beta(\cdot,\underline{p})$ is measurable.

Next, consider the integral $\int_A \gamma(\underline{a},\underline{p})\mu(d\underline{a})$ as a function of $\underline{p}$. For any $-\infty \le \underline{p} \le \infty$ this is well-defined, and in fact finite, by (6) and (1). It is also non-increasing in $\underline{p}$, since $\gamma(\underline{a},\underline{p})$ is non-increasing for each $\underline{a} \in A$. Using the monotone convergence theorem, it follows from the left-continuity of $\gamma(\underline{a},\cdot)$ that $\int_A \gamma(\underline{a},\cdot)\mu(d\underline{a})$ is also left-continuous. (Take a sequence $\underline{p}_1, \underline{p}_2, \ldots$ increasing ~~tp~~ to $\underline{p}_o$; then $\gamma(\underline{a},\underline{p}_n) \to \gamma(\underline{a},\underline{p}_o)$ by left-continuity; monotone convergence then yields $\int_A \gamma(\underline{a},\underline{p}_n)\mu(d\underline{a}) \to \int_A \gamma(\underline{a},\underline{p}_o)\mu(d\underline{a})$.)

A similar argument shows that $\int_A \beta(\underline{a},\cdot)\mu(d\underline{a})$ is finite, non-increasing, and right-continuous.

We are now ready to construct the optimal solution. Let

$$\underline{p}^\circ = \underline{\sup}\left\{\underline{p} \,\middle|\, \int_A \gamma(\underline{a},\underline{p})\mu(d\underline{a}) \ge \underline{L}\right\}. \qquad \text{(5.6.10)} \; \cancel{(10)}$$

Since $\int_A \gamma(\underline{a},\cdot)\mu(d\underline{a})$ is continuous from the left, it follows that

$$\int_A \gamma(\underline{a},\underline{p}^\circ)\mu(d\underline{a}) \ge \underline{L}, \qquad \text{(5.6.11)} \; \cancel{(11)}$$

if $p^o > -\infty$. (Inequality (11) is also true if $p^o = -\infty$, by (6) and (1). Next, we show that

$$\int_A \beta(a,p^o)\mu(da) \le L. \qquad (5.6.12)\ \cancel{(12)}$$

If $p^o = +\infty$, then (12) follows from (6) and (1). If $p^o < \infty$, then for every $p > p^o$ we have

$$\int_A \beta(a,p)\mu(da) \le \int_A \gamma(a,p)\mu(da) < L,$$

from (10). (12) then follows from the right-continuity of $\int_A \beta(a,\cdot)\mu(da)$.

This p^o turns out to be the shadow price of the optimal solution.

Now consider the indefinite integral

$$\int [\gamma(a,p^o) - \beta(a,p^o)]\mu(da). \qquad (5.6.13)\ \cancel{(13)}$$

This is finite. Also, since μ is non-atomic, (13) is a non-atomic measure. Hence it takes on every value between 0 and its value on A, inclusive of these bounds. Now

$$L - \int_A \beta(a,p^o)\mu(da) \qquad (5.6.14)\ \cancel{(14)}$$

lies between these bounds, from (11) and (12). Hence there is a measurable set F such that (14) equals the value of (13) at F. This yields

$$\int_{F} \gamma(a,p^{\circ})\mu(da) + \int_{A\setminus F} \beta(a,p^{\circ})\mu(da) = L. \qquad \cancel{(15)}\ (5.6.15)$$

We now claim that the function $\delta^{\circ}: A \to$ reals ~~which~~ coincides with $\gamma(\cdot,p^{\circ})$ on F, and with $\beta(\cdot, p^{\circ})$ on $A \setminus F$, is best. For, it is measurable and satisfies (2), since both $\gamma(\cdot,p^{\circ})$ and $\beta(\cdot, p^{\circ})$ satisfy these conditions. Also it satisfies (3), by (15). Hence δ° is feasible. Also it satisfies the sufficient condition for "bestness", since both $\gamma(a,p^{\circ})$ and $\beta(a,p^{\circ})$ maximize (4) for $p = p^{\circ}$. ▯

We now remove the condition that μ is non-atomic. In its place, however, we are obliged to add a further condition on f, ~~namely~~ viz., that $|f(a,x)| \leq \theta(a)$, where θ is some function with a finite integral. One consequence of this new condition may be noted: It guarantees that the utility function is a finite signed measure for all feasible δ. Hence standard ordering of pseudomeasures reduces to the ordinary comparison of definite integrals, and the distinction between "best" and "unsurpassed" disappears. To emphasize this point, we ~~shall~~ write the utility functions in the following theorem and proof in the form of *definite* integrals.

Theorem: Let (A,Σ,μ) be a measure space, with μ ~~sigma~~ σ-finite; let $b,c: A \to$ reals, and $f: A \times$ reals $\to$ reals, be measurable; let $f(a,\cdot)$ be continuous for all $a \in A$. Assume $b \leq c$, and

$$-\infty < \int_A b\, d\mu \le 0 \le \int_A c\, d\mu < \infty. \qquad \text{(5.6.16)}\ \cancel{(16)}$$

Also assume ~~there is~~ a measurable function $\theta: A \to$ reals exists such that $\int_A \theta\, d\mu$ is finite, and

$$|f(a,x)| \le \theta(a), \qquad \text{(5.6.17)}\ \cancel{(17)}$$

for all $x \in [b(a), c(a)]$, $a \in A$.

Then the problem: Maximize

$$\int_A f(a, \delta(a))\, \mu(da) \qquad \text{(5.6.18)}\ \cancel{(18)}$$

over measurable functions $\delta: A \to$ reals, subject to

$$b \le \delta \le c, \qquad \text{(5.6.19)}\ \cancel{(19)}$$

~~and to~~

$$\int_A \delta\, d\mu = 0, \qquad \text{(5.6.20)}\ \cancel{(20)}$$

has a best solution.

Proof: This proof is divided into two parts. In the first, we assume ~~that~~ μ is ~~sigma~~ σ-atomic; i.e. ~~that is, there is~~ a countable measurable partition $\{A_1, A_2, \ldots\}$ exists such that μ restricted to each A_n is atomic. Since μ is also ~~sigma~~ σ-finite, it must be bounded on each A_n.

First, we show that total utility depends only on how mass is distributed among the atoms, and, given this, is independent of how mass is distributed within the atoms. That is, suppose δ_1 and δ_2 are two densities such that

$$\int_{A_n} \delta_1 \, d\mu = \int_{A_n} \delta_2 \, d\mu = \lambda_n, \qquad (5.6.21) \ \cancel{(21)}$$

say, for all $n = 1, 2, \ldots$. Then

$$\int_A f(a, \delta_1(a)) \, \mu(da) = \int_A f(a, \delta_2(a)) \, \mu(da). \qquad (5.6.22) \ \cancel{(22)}$$

To show this, we invoke the lemma, ~~(9)~~ (4.9), of ~~the preceding section,~~ section 4, stating that δ_1 and δ_2 must each be equal to constants almost everywhere on any atom: say $\delta_1 = d_{1n}$, $\delta_2 = d_{2n}$ on A_n, almost everywhere. From (21) we obtain $d_{1n}\mu(A_n) = d_{2n}\mu(A_n) = \lambda_n$, which means that $d_{1n} = d_{2n}$ for all n. Hence $\delta_1 = \delta_2$ almost everywhere, so that (22) is ~~of course~~ correct. Thus utility depends only on the sequence $(\lambda_1, \lambda_2, \ldots)$, and is in fact given by

$$g_1(\lambda_1) + g_2(\lambda_2) + \ldots, \qquad (5.6.23) \ \cancel{(23)}$$

where

$$g_n(\lambda_n) = \int_{A_n} f\big(a, \lambda_n/\mu(A_n)\big) \, \mu(da). \qquad (5.6.24) \ \cancel{(24)}$$

12pt brackets

g16-4

This reduces the problem to two simpler issues. First, for what sequences $(\lambda_1, \lambda_2, \ldots)$ are there feasible densities δ such that (21) is satisfied for all $n = 1, 2, \ldots$? Second, among these feasible sequences is there one that maximizes (23)?

The first question is easily answered. Integrating the constraint (19) over A_n, we obtain

$$\underline{b}_n \leq \lambda_n \leq \underline{c}_n \qquad \text{(5.6.25)} \; \text{~~(25)~~}$$

first

127 18 69 18 215

for all $\underline{n}$, where $\underline{b}_n = \int_{A_n} b\,d\mu$, $\underline{c}_n = \int_{A_n} c\,d\mu$. Furthermore, the constraint (20) implies that

start

$$\lambda_1 + \lambda_2 + \ldots = 0. \qquad \text{(5.6.26)} \; \text{~~(26)~~}$$

Then (25) and (26) give necessary conditions for any feasible sequence $(\lambda_1, \lambda_2, \ldots)$. Conversely, they are also sufficient for ~~there~~ to exist a feasible δ yielding this sequence. For if (25) is satisfied, one easily sees that some weighted average $\underline{t}_n b + (1-\underline{t}_n)c = \delta$ will satisfy (21) for $\underline{n}$. The δ thus constructed automatically satisfies (19), and satisfies (20) because of (26).

We have thus reduced the problem to one with a countable number of unknowns: Maximize (23) over real sequences $(\lambda_1, \lambda_2, \ldots)$ satisfying (25) and (26).

Let $\lambda^{\underline{k}} = (\lambda_1^{\underline{k}}, \lambda_2^{\underline{k}}, \ldots)$, $\underline{k} = 1, 2, \ldots$, be a sequence of these feasible sequences, such that the value of (23) approaches its supremum as $\underline{k} \to \infty$. We first give a standard argument to show ~~that~~ there is a subsequence $\lambda^{k_1}, \lambda^{k_2}, \ldots$, such that, for all $\underline{n} = 1, 2, \ldots$, the sequence $\lambda_{\underline{n}}^{k_1}, \lambda_{\underline{n}}^{k_2}, \ldots$ has a limit $\lambda_{\underline{n}}^{o}$.

First, the sequence $\lambda_1^1, \lambda_1^2, \ldots$ is all contained in the closed, bounded interval $[\underline{b}_1, \underline{c}_1]$. Hence there is a convergent subsequence. By the same argument, there is a subsequence of this subsequence such that the λ_2 values converge.

Continuing, we get a sequence of sequences, each a subsequence of its predec~~ison~~essor, the n-th subsequence having convergent λ_n values. Finally, one takes the "diagonal": viz., the q-th term of the q-th subsequence. This yields a subsequence converging for all $n = 1, 2, \ldots$ ⊙

We now show that the resulting limiting sequence $(\lambda_1^o, \lambda_2^o, \ldots)$ is optimal. First we prove feasibility. Inequality (25) is satisfied by λ_n^o, since it is the limit of a sequence in that interval $[b_n, c_n]$. ~~Next,~~ we must show that

$$\lambda_1^o + \lambda_2^o + \ldots = 0. \qquad \text{(5.6.27)} \; \cancel{(27)}$$

Let $\lambda^{k_1}, \lambda^{k_2}, \ldots$ be the subsequence converging to $\lambda^o = (\lambda_1^o, \lambda_2^o, \ldots)$. Think of

$$\lambda_1^{k_i} + \lambda_2^{k_i} + \ldots \qquad \text{(5.6.28)} \; \cancel{(28)}$$

as the integral of a function $\lambda^{k_i}: \{1, 2, \ldots\} \to$ reals, all subsets of the positive integers being measurable, and having the ~~enumeration~~ counting measure: $\nu\{n\} = 1$ for all $n = 1, 2, \ldots$. We also have, for all i and all n,

$$\left|\lambda_n^{k_i}\right| \le |b_n| + |c_n|,$$

break at $\le$ +

by (25); and the sum over all n of $|b_n| + |c_n|$ is finite, by (16). Hence we may invoke the dominated convergence theorem, and assert that the limit of the sums in (28) as $k_i \to \infty$ is the sum of the limits. But all sums in (28) equal zero, by (26).

Hence (27) is true. This proves that λ^{o} is feasible.

It remains to show that $(\lambda_1^{o}, \lambda_2^{o}, \ldots)$ maximizes (23) over the set of feasible sequences, (25)–(26). First, we show that the function g_n given by (24), over the domain (25), is continuous. Let $\lambda^1, \lambda^2, \ldots$ be a sequence of numbers with limit λ, all satisfying (25). Since $f(a, \cdot)$ is continuous,

$$\lim_{k\to\infty} f(a, \lambda^k/\mu(A_n)) = f(a, \lambda/\mu(A_n)),$$

for all $a \in A_n$. Also, $|f(a, \lambda^k/\mu(A_n))| \leq \theta(a)$, and

[164]

$$\int_{A_n} \theta \, d\mu = \theta_n < \infty \qquad \text{(5.6.29)} \; \cancel{(29)}$$

by (17). Hence we may invoke the dominated convergence theorem again, and assert that

16-7

$$\lim_{k\to\infty} g_n(\lambda^k) = g_n(\lambda).$$

Hence g_n is continuous for all $n = 1, 2, \ldots$. Next, for each i think of the sum

$$g_1(\lambda_1^{k_i}) + g_2(\lambda_2^{k_i}) + \ldots \qquad \text{(5.6.30)} \; \cancel{(30)}$$

as the integral of a function with domain $\{1, 2, \ldots\}$, the measure on this space being counting ~~the enumeration~~ measure, as in (28). We have, for all i and all n,

362

$$\left| g_n(\lambda_n^{k_i}) \right| \leq \theta_n,$$

262

from (24) and (29). Also,

13 72

$$\theta_1 + \theta_2 + \ldots = \int_A \theta_A \, d\mu < \infty .$$

Hence we may invoke dominated convergence yet a third time, and assert that the limit of the sums in (30) as $k_i \to \infty$ is the sum of the limits. Now $\lambda_n^{k_i} \to \lambda_n^{o}$ for all n, as $k_i \to \infty$. By the continuity of g_n it follows that

159 26 103

$$\lim_{k_i \to \infty} g_n(\lambda_n^{k_i}) = g_n(\lambda_n^{o}),$$

for all $n = 1, 2, \ldots$. Hence the limit of the sums in (30) is

$$g_1(\lambda_1^{o}) + g_2(\lambda_2^{o}) + \ldots . \qquad (31)$$

(5.6.31)

But the limit of the sums in (30) is also the supremum of (23) over all feasible $(\lambda_1, \lambda_2, \ldots)$, by the construction of the original sequence of sequences, $\lambda^1, \lambda^2, \ldots$. Hence (31) is the maximum of (23), and $(\lambda_1^{o}, \lambda_2^{o}, \ldots)$ is optimal. Any feasible δ yielding this sequence via (21) is then a best solution.

#

This completes the first half of the proof.

18 We now drop the restriction that μ be ~~sigma~~ σ-atomic. Since μ is ~~sigma~~ σ-finite, there is a countable measurable partition $\{A_\Theta, A_1, \ldots\}$ such that μ is non-atomic on A_Θ, and atomic on each A_n, $n = 1, 2, \ldots$.

0 = zero unless marked

86 18 95 18 205

Let $b_\Theta = \int_{A_\Theta} b \, d\mu$ and $c_\Theta = \int_{A_\Theta} c \, d\mu$, and let λ_Θ be a number with

0 = zero where inked

$$b_\theta \le \lambda_\theta \le c_\theta \qquad \text{(5.6.32)}\ \cancel{(32)}$$

Consider the problem of maximizing (18), with A_θ in place of A, over measurable functions $\delta : A_\theta \to$ reals satisfying (19) and

[179] 21 67

$$\int_{A_\theta} \delta\, d\mu = \lambda_\theta . \qquad \text{(5.6.33)}\ \cancel{(33)}$$

The preceding theorem ~~(page)~~ states that ~~there exists~~ a best solution, δ^o, exists to this problem, since μ restricted to A_θ is non-atomic. For this best solution the utility function has the value

[160] 21 105

$$\int_{A_\theta} f(a, \delta^o(a))\, \mu(da) . \qquad \text{(5.6.34)}\ \cancel{(34)}$$

Now δ^o, hence (34), depends on λ_θ. ~~Let us~~ We write $g_\theta(\lambda_\theta)$ for the value (34) as a function of λ_θ. The domain of g_θ is $[b_\theta, c_\theta]$. ~~Also,~~ because of the special assumption (17), we interpret (34) as a (finite) definite integral; hence g_θ is real-valued, rather than pseudomeasure-valued as in the general case.

Consider now the problem of maximizing

$$g_\theta(\lambda_\theta) + g_1(\lambda_1) + \ldots \qquad \text{(5.6.35)}\ \cancel{(35)}$$

over all sequences $(\lambda_\theta, \lambda_1, \ldots)$ satisfying

$$b_n \le \lambda_n \le c_n \qquad \text{(5.6.36)}\ \cancel{(36)}$$

stress zero

for all $\underline{n} = \underline{0}$, 1, 2,..., and

$$\lambda_\theta + \lambda_1 + \ldots = 0. \qquad \text{(5,6,37)}\ \cancel{(37)}$$

Here λ_n, b_n, c_n, and g_n ~~for~~ $n = 1, 2, \ldots$ have exactly the same meanings as before, while λ_θ, b_θ, c_θ, and g_θ have just been defined.

16-8

If $(\lambda^o_\theta, \lambda^o_1, \ldots)$ is an optimal solution to (35)-(37), we can construct an optimal solution, δ^o, to the original problem (18)-(20) as follows. On A_θ, let δ^o coincide with the optimal solution to the non-atomic problem (32)-(33), with parameter $\lambda_\theta = \lambda^o_\theta$. On A_n, for $n > 0$, choose any feasible δ satisfying (21) for λ^o_n. The resulting function δ^o is clearly feasible. It is also optimal, since the utility function (18) for any feasible δ does not exceed (35), where the λ_n's are determined from δ by (21); for $\delta = \delta^o$, the utility function is equal to (35), which is the maximum of its possible values.

It suffices, then, to show that (35)-(37) has an optimal solution. ~~Now~~ this is of exactly the same form as the problem of maximizing (23) subject to (25) and (26), with one possible exception: We do not know whether the function g_θ is continuous. If this could be shown, then the first half of this proof demonstrates the existence of an optimal $(\lambda^o_\theta, \lambda^o_1, \ldots)$, and we would be finished.

We now show that g_θ is continuous. As a first step we show it is concave[8]. Let L_1, L_2, L_3 satisfy

$$\int_{A_\theta} b \, d\mu \leq L_1 < L_2 < L_3 \leq \int_{A_\theta} c \, d\mu, \tag{38}$$

and let $\delta_i^\circ : A_\theta \to$ reals be the optimal solution for the parameter L_i $(i = 1, 2, 3)$. The proof of the preceding theorem shows that these optimal solutions have shadow prices. Hence for δ_2° ~~there exists~~ an extended real number, p°, exists such that $\delta_2^\circ(a)$ maximizes

$$f(a,x) - p^\circ x$$

over $x \in [b(a), c(a)]$, for almost all $a \in A_\theta$. This p° must be finite; for if $p^\circ = +\infty$, then $\delta_2^\circ = b$ almost everywhere on A_θ, which contradicts (38); similarly, $\delta_2^\circ = c$ almost everywhere on A_θ if $p^\circ = -\infty$, again contradicting (38). It follows that

$$f(a, \delta_2^\circ(a)) - p^\circ \delta_2^\circ(a) \geq f(a, \delta_i^\circ(a)) - p^\circ \delta_i^\circ(a)$$

almost everywhere on A_θ $(i = 1,3)$. Integration over A_θ yields

$$g_\theta(L_2) - g_\theta(L_i) \geq p^\circ(L_2 - L_i)$$

$(i = 1, 3)$, so that

$$\frac{g_\theta(L_2) - g_\theta(L_1)}{L_2 - L_1} \geq p^\circ \geq \frac{g_\theta(L_3) - g_\theta(L_2)}{L_3 - L_2}$$

for all $L_1 < L_2 < L_3$, a condition equivalent to concavity.

Since a concave function is continuous at all interior points, the only thing left to prove is that g_θ is continuous at the endpoints, b_θ and c_θ. Because g_θ is concave, to establish continuity at b_θ it suffices to prove that

$$g_\theta(b_\theta) = \lim_{k\to\infty} g_\theta(L_k) \qquad \text{(5.6.39)} \ ~~(39)~~$$

for any one sequence $L_1, L_2, \ldots$ converging to b_θ.

~~Now~~ we re-introduce the function $\gamma(a,p)$ given by (5). In the preceding proof it was established that, for fixed $a \in A_\theta$, $\gamma(a,p)$ has the limit $b(a)$ as $p\to\infty$. Also, for fixed p, $\gamma(\cdot,p)$ is measurable, and

$$\int_{A_\theta} \gamma(a,p)\,\mu(da),$$

as a function of p, approaches

$$\int_{A_\theta} \gamma(a,\infty)\,\mu(da) = \int_{A_\theta} b\,d\mu = b_\theta$$

as $p\to\infty$.

~~Now~~ let $p_1, p_2, \ldots$ be a sequence increasing without bound, and define $L_1, L_2, \ldots$ by

$$L_k = \int_{A_\theta} \gamma(a,p_k)\,\mu(da). \qquad \text{(5.6.40)} \ ~~(40)~~$$

The sequence $L_1, L_2, \ldots$ then converges to b_θ. Also,

$$g_\theta(L_k) = \int_{A_\theta} f(a,\gamma(a,p_k))\,\mu(da).$$

This follows from the fact that the function $\delta_k^o(a) = \gamma(a,p_k)$ has the shadow price p_k, and satisfies the resource constraint (40), hence it is optimal for L_k. Now for each $a \in A_\theta$,

$$\lim_{k\to\infty} f(a,\ \gamma(a,p_k)) = f(a,\ b(a)),$$

by the continuity of $f(a,\cdot)$. Hence, by (17), we may apply the dominated convergence theorem, and conclude that

$$\lim_{k\to\infty} g_\theta(L_k) = \int_{A_\theta} f(a,\ b(a))\mu(da) = g_\theta(b_\theta).$$

This proves (39) for the sequence $L_1, L_2, \ldots$, and establishes the continuity of g_θ at b_θ. Continuity of g_θ at c_θ is proved by a similar argument, with $\beta(a,p)$ in place of $\gamma(a,p)$, and $p \to +\infty$.

This supplies the missing link in the proof, and we conclude that a best solution δ^o exists in the general case. $\square$

As usual, there is an immediate generalization. If (20) in the preceding theorem (or (3) in the one before that) is replaced by the condition

$$L_o \le \int_A \delta\, d\mu \le L^o, \qquad \text{(5.6.41)}\ \cancel{(41)}$$

with $\int_A \delta\, d\mu$ required to be finite, and (16) in the preceding theorem (or (1) in the one before that) is replaced by the condition

Stet

117] $$-\infty < \int_A b\, d\mu - L^o \leq 0 \leq \int_A c\, d\mu - L_o < \infty,$$

then ~~there~~ still exists a best solution in these respective cases.

The proof, which consists as always in transforming this problem into an equivalent one in which (20) and (16), or (3) and (1) hold, is left as an exercise.

We now give an example of a problem not having an optimal solution. Let $A = \{1, 2, 3, \ldots\}$, Σ = all subsets, μ the counting ~~enumeration~~ measure; $b(n) = 0$ and $c(n) = 1$ for all $n = 1, 2, \ldots$. The payoff function is: $f(n,x) = -x/n^2$. The density function, in addition to satisfying $0 \leq \delta(n) \leq 1$ for all n, must satisfy

$$\delta(1) + \delta(2) + \ldots = 1.$$

We can ~~One~~ easily verify ~~verifies~~ that all ~~of~~ the premises of the preceding theorem are satisfied (take $\theta(n) = 1/n^2$ in (17); $L_o = L^o = 1$ in (41)), with one exception: $\int_A c\, d\mu = c(1) + c(2) + \ldots = \infty$.

There is no optimal solution to this problem, since any given feasible solution can be improved. To see this, let δ be feasible, and choose any n for which $\delta(n) > 0$. Alter δ by replacing $\delta(n)$ by 0, and $\delta(n+1)$ by $\delta(n+1) + \delta(n)$, everything else the same. This remains feasible, and the change in the

$\delta(n)[n^{-2}-(n+1)^{-2}]$

utility function is ~~$\delta(n)$~~ $\boxed{\cancel{\frac{1}{n^2} - \frac{1}{(n+1)^2}}}$ > 0, so δ is non-optimal.

This example gives a certain insight into the rôle of the finiteness condition on $\int_A b \, d\mu$ and $\int_A c \, d\mu$.

5.7. Uniqueness of Optimal Solutions

By "uniqueness" we mean the property that there is at most one optimal solution. (The ordinary word "uniqueness" sometimes carries the connotation of "exactly one"; but we are not here concerned with existence, only with "non-duplication" of solutions).

As ~~in our previous discussion~~ before, we identify any two densities ~~which~~ that are unequal only on a null set. Thus, to say ~~that~~ there is at most one optimal solution is to say: If δ_1 and δ_2 are both optimal solutions, ~~then~~ $\mu\{a \mid \delta_1(a) \neq \delta_2(a)\} = 0$.

We need two new concepts for the following result.

6-10

Definition: f:reals → reals is strictly concave iff, for any two distinct real numbers, x, y, and any $0 < t < 1$,

$$f(tx + (1-t)y) > tf(x) + (1-t)f(y). \qquad (5.7.1)\ (1)$$

This is a bit stronger than concavity *per se*, because of the strict inequality in (1). A linear function is concave but not strictly concave.

Definition: Let $\mathbb{M}$ be a set of real-valued functions, all with domain A. (Set) $\mathbb{M}$ is convex iff, for any δ_1, $\delta_2 \in \mathbb{M}$, and any $0 < t < 1$, the function $t\delta_1 + (1-t)\delta_2$ belongs to $\mathbb{M}$.

For example, the feasible sets we have been dealing with throughout sections 4, 5, and 6 are all convex.

Utility ordering is still that of standard ordering of pseudomeasures. The distinction between best and unsurpassed must again be stressed, because it is critical in the following result.

Theorem: Let (A,Σ,μ) be a measure space, with μ ~~sigma~~ σ-finite. Let $f: A \times \text{reals} \to \text{reals}$ be measurable, and such that, for all $a \in A$, $f(a,\cdot)$ is strictly concave. Let $\mathbb{M}$ be a convex set of real-valued measurable functions.

Then the problem: Maximize

$$\int f(a, \delta(a))\mu(da) \tag{5.7.2}$$

over $\delta \in \mathbb{M}$, has at most one best solution.

Proof: Let δ_1 and δ_2 both be best. Then, for any $\delta \in \mathbb{M}$ we have

$$\int_A \Big[f(a,\delta_i(a)) - f(a,\delta(a))\Big]\mu(da) \geq 0,$$

$(i = 1, 2)$. Adding these two inequalities, we get

$$\int_A \Big[f(a,\delta_1(a)) + f(a,\delta_2(a)) - 2f(a,\delta(a))\Big]\mu(da) \geq 0. \tag{5.7.3}$$

set close per use proper fraction

Now consider the function $\delta = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2$. This $\delta \in \mathbb{M}$, by convexity. For this δ, the integrand in (3) is never positive, and is in fact negative on the set $\{\underline{a} | \delta_1(\underline{a}) \neq \delta_2(\underline{a})\}$, by strict concavity. Hence this set has measure zero, which establishes uniqueness. $\square$

It is not true that there must be at most one unsurpassed solution, as the following counterexample shows.

Let $(\underline{A}, \Sigma, \mu)$ be Lebesgue measure on the real line. Let $\underline{f}(\underline{a},\underline{x}) = -\underline{x}^2 + 2\underline{x}$ if $\underline{a} \geq 0$; let $\underline{f}(\underline{a},\underline{x}) = -\underline{x}^2 - 2\underline{x}$ if $\underline{a} < 0$; and let $\mathbb{M}$ be the set of constant functions whose single value lies in the closed interval $[-1,1]$. Set $\mathbb{M}$ is obviously convex, and one sees ~~easily verifies~~ that $\underline{f}(\underline{a},\cdot)$ is strictly concave for all $\underline{a} \in \underline{A}$.

raised dot

Now let $\underline{x}_1$, $\underline{x}_2$ be two numbers in $[-1,1]$, with $\underline{x}_1 > \underline{x}_2$, and let ψ_1, ψ_2 be the pseudomeasures obtained by substituting the corresponding functions in (2). Since $\underline{f}(\underline{a},\cdot)$ is increasing for $\underline{a} \geq 0$ and decreasing for $\underline{a} < 0$, ~~hence~~ $(\psi_1 - \psi_2)^+$ is a multiple of Lebesgue measure truncated to the positive half-line, while $(\psi_1 - \psi_2)^-$ is a multiple of Lebesgue measure truncated to the negative half-line. It follows that

for ψ

$$(\psi_1 - \psi_2)^+(\underline{A}) = (\psi_1 - \psi_2)^-(\underline{A}) = \infty, \qquad (5.7.4)\ \cancel{(4)}$$

Hence all feasible solutions are unsurpassed, because, by (4), no two of them are comparable.

9.17.1 (A) 1.6 type

5.8. Police, Criminals, and Victims

We ~~shall~~ now apply the preceding theory to the problem of the spatial distribution of crime. ~~Actually, the present~~ This section goes beyond the simple optimization framework of the rest of ~~this~~ the chapter, in that several different populations, with diverse motives, are interacting. Thus we are in a "game" situation, and what is optimal for one agent may not be optimal for another.

10/12 x 27 TR
Q = 2 ems

There is a population of potential victims, a population of potential criminals, who commit crimes upon the victims when the opportunity presents itself, and a population of policemen, who try to prevent criminals from perpetrating their misdeeds.

use Bask italic for vel'p throughout

The three populations inh~~i~~abit the measure space (S, Σ, α), α being ideal area over Space, S. If v, c, and p are the densities of the three respective populations (density with respect to α), the density of crimes at location s is given by a function

$$f(v(s), c(s), p(s)), \qquad \text{(5.8.1)} \; ~~(1)~~$$

and total crime is then given by $\int_S f\, d\alpha$. We would expect f to be an increasing function of v and c, and a decreasing function of p.

Consider, for illustrative purposes, the crime function

$$f(v,c,p) = vce^{-p}. \qquad \text{(5.8.2)} \; ~~(2)~~$$

fl left

A semi-plausible rationalization for (2) might run as follows. For $p = 0$, a crime occurs if there is an "encounter" between a potential victim and criminal, and, with "random" movements, the frequency of encounters should be proportional to the product of the densities, vc.[9] Next, suppose ~~that~~ the commission of a crime is inhibited if a policeman is present within a certain "surveillance radius". If policemen are randomly distributed, the probability of no policeman being present within the critical radius declines exponentially with police density, and this gives (2). (Units of measurement for v, c, and p may be chosen to avoid multiplicative constants, as in (2)).

Whatever one thinks of such arguments, it is still illuminating to discuss the consequences of (2) – or, more generally, (1) – under various behavioral assumptions. We assume that victims and police distribute themselves over Space so as to reduce crimes, while criminals distribute themselves so as to increase crimes.

One further objection to this set-up should be mentioned. Shouldn't these population distributions be integer-valued, or finitely concentrated? – If so, ~~in which~~ case they are unlikely to have density functions. The answer is that c, v and p are best thought of, not as densities for cross-sectional distributions, but for the time-averages arising from the random perambulations of the populations. Thus a measure μ, where $\mu(E)$ is the *expected* number of people of a certain type in region E, can very well have an areal density.

Before launching into details, ~~let us~~ we briefly consider some specific interpretations of this general model. "Crime" is a rather heterogeneous category, and not all types of crimes would be well-represented by the model. ~~In the first place,~~ there are numerous "crimes without victims"[10]: gambling, traffic in drugs, prostitution, etc. In some of these cases the frequency might be described by (1) and (2) (e.g., ~~- for example,~~ "random" encounters with streetwalkers) ~~-~~ but ~~one~~ we would not expect the "victims" to be motivated to reduce the incidence of such "crimes". Secondly, there are crimes ~~which~~ that do not require a specific "encounter" for their commission (e.g., counterfeiting, or anti-trust law violation~~, for instance.~~)

Burglary, larceny, robbery, and rape are examples of types of crime ~~which~~ that do not have these disqualifying features, and their incidence might be approximately represented by a function of the form (1). One might want to re-interpret v, c, or p in some of these cases. For example, in burglary the spatial distribution of (movable) wealth would seem more relevant than the spatial distribution of people, so v should perhaps be taken as wealth density rather than population density.

Certain non-crime situations may also be represented by this model. Consider military attacks against targets (installations, opposing forces, civilians~~, etc.~~). Letting v

9.17.2

be the density of targets, c the density of, say, bombing, and p the density of "defense equipment", the above model might then predict the volume of destruction in terms of these three distributions. The controllers of p and v are motivated to reduce destruction, the controllers of c to increase it. Hence we might expect to find the spatial distributions here similar to those ~~which~~ resulting from crime incidence.

Again, consider the following "imitation-snob" situation. There is a "high-prestige" and a "low-prestige" population. The high-prestigers want to avoid contacts with the low-prestigers, while the latter want to increase contacts with the former. Interpreting "criminals" and "victims" to be the low- and high-prestige populations, respectively, and "crimes" to be contacts between the two populations, we get something like the model above. (The police might enter as harassers of the low-prestige population in its attempts to increase contacts).

~~As throughout this book,~~ our aim here is to develop and explore theoretical models, not to tailor them closely to any particular real-world situation. (For crimes such an attempt would in any case be difficult, because of the spotty quality of most crime data).[11]

We now return to the formal model, which has not yet been completely specified. For simplicity we assume that the three populations are mutually exclusive, and that no transformation from one to the other is possible. (Thus we ignore the possibility that victims themselves can inhibit crimes by surveillance, the possibility that some criminals can be victimized by other criminals, etc.).

Two cases will be explored. In the first there are no police (the anarchistic, or "Wild West" case) and the two populations, victims and criminals, are freely mobile over space. In the second case, the distribution of victims is *fixed*, and the remaining two populations are freely mobile. (This might occur, for example, if crime is of minor importance so that it exerts no locational pull on the population at large. Another interpretation is that the population distribution of victims adjusts very slowly compared to the other two populations, so that it may be considered fixed in the short run.)

A given population tries to reduce — or increase — crime. What does this mean? There are (at least) two interpretations: the *individualistic*, and the *collusive*. If criminals act collusively, ~~for example,~~ the entire body of criminals distributes itself so as to maximize total crime; if they act individualistically, then they will move from places where the density of crimes *per criminal* is low to where it is high. Similarly, if victims collude they will distribute themselves to minimize total crime; if they are individualists, they move from places where the density of crimes *per victim* is high to where it is low. For some crime functions *f* the resulting distributions are the same under either assumption, but in general they will differ.

For the police the most plausible assumption is collusion: They are distributed by central headquarters to minimize total crime. For victims the individualistic assumption is more plausible: Each potential victim moves to reduce the incidence of crime on himself. For criminals, both possibilities are plausible, depending on whether crime is "petty", or "organized" by some criminal mastermind.

We ~~shall~~ analyze just three of the many possible combinations: (i) no police, both victims and criminals are individualists; (ii) no police, both victims and criminals collude; (iii) victim distribution fixed, both police and criminals collude.

Case (i): With no police, we have a crime density function f:reals^2 $\rightarrow$ reals, ~~namely~~ viz. $f(v,c)$ is the crime density at a location, if victim density there is v and criminal density c. All functions are assumed to be measurable, ~~real-valued~~ finite, and nonnegative. We also assume that

$$f(0,c) = f(v,0) = 0,$$

and that the right-hand partial derivatives $D_1f(0,c)$ and $D_2f(v,0)$ exist for all c, $v \geq 0$. (We use the notation

$$D_1f(0,c) = \lim_{v\to 0+} [f(v,c) - f(0,c)]/v,$$

~~and similarly for $D_2f(v,0)$.)~~

$$D_2f(v,0) = \lim_{c\to 0+} [f(v,c) - f(v,0)]/c.$$

9.17.3

To avoid trivialities, we assume that the total population of victims, V, and criminals, C, are fixed positive real numbers, as is the total available area, $\alpha(S)$. The constraints on the density functions, v and c, ~~is~~ are that they be non-negative real measurable, and satisfy

$$\int_S v \, d\alpha = V, \quad \int_S c \, d\alpha = C. \qquad \cancel{(3)}\ (5.8.3)$$

Definition: The pair of feasible densities $v^o, c^o : S \to$ reals is an <u>individualistic equilibrium pair</u> iff there is a null set $E \in \Sigma$, and two real numbers, $k_v, k_c \geq 0$, such that, for all $s \in S \backslash E$,

$$\frac{f(v^o(s), c^o(s))}{v^o(s)} = k_v, \text{ if } v^o(s) > 0, \qquad \cancel{(4)}\ (5.8.4)$$

$$\frac{f(v^o(s), c^o(s))}{c^o(s)} = k_c, \text{ if } c^o(s) > 0, \qquad \cancel{(5)}\ (5.8.5)$$

$$D_1 f(0, c^o(s)) \geq k_v, \text{ if } v^o(s) = 0, \qquad \cancel{(6)}\ (5.8.6)$$

$$D_2 f(v^o(s), 0) \leq k_c, \text{ if } c^o(s) = 0. \qquad \cancel{(7)}\ (5.8.7)$$

Editor end of def?

The intuitive meaning of (4)–(7) is as follows. We interpret the "incidence of crime" on any victim at location s to be the <u>crimes per victim</u> at that point, which is the left side of (4), if $v(s) > 0$. If $v(s) = 0$, the natural interpretation is $D_1 f(0, c(s))$, which is what crimes per victim would be for a "low-density" migration there. (4) and (6) are then precisely the conditions under which no potential victim can, by moving, reduce the incidence of crime on himself. Similarly, we take the "gain from crime" for any criminal at location $\underline{s}$ to be the crimes per criminal at that point, which is the left side of (5) or (7), for $c(s) > 0$, $c(s) = 0$, respectively. (5) and (7) are then the conditions that no criminal can gain from moving. As usual, we allow exceptions on a set of measure zero. Two density functions, v_1 and v_2, which differ only on a null set, are taken to be identical, and similarly for c_1 and c_2.

24 Case (ii). Now ~~let us~~ we go to collusive criminal-victim interaction. The set of feasible densities is again given by (3).

D **Definition:** The pair of feasible densities, v°, c°, is a collusive equilibrium pair iff

66 339

$\int f(v^\circ, c^\circ)\,d\alpha$ is unsurpassed in the set of pseudomeasures,

$\int f(v^\circ, c)\,d\alpha$, c ranging over the feasible criminal densities; ~~and~~

and 332

$-\int f(v^\circ, c^\circ)\,d\alpha$ is unsurpassed in the set of pseudomeasures,

$-\int f(v, c^\circ)\,d\alpha$, v ranging over the feasible victim densities.

display on page

That is, given the distribution v°, criminals arrange themselves over Space so that no other arrangement of criminals leads to a distribution of total crimes surpassing the one resulting from c°; conversely, given c°, v° is chosen so that minus the distribution of total crimes is not surpassed by that resulting from any other victim distribution. The reason for the "minus" is of course that victims are trying to reduce total crime, which is equivalent to trying to increase minus total crime.

~~One has to~~ We must invoke pseudomeasures only if the total crime integral can be unbounded. For the present application we may safely assume that, for any f arising in practice, all integrals are finite. Nonetheless we give the more general definition above because the results obtained are valid for it, and no extra work is involved.

If all integrals are finite, the above definition may be restated in simpler form: The feasible pair (v^o, c^o) is a collusive equilibrium pair iff

$$\int_S f(v^o,c)\,d\alpha \leq \int_S f(v^o,c^o)\,d\alpha \leq \int_S f(v,c^o)\,d\alpha \qquad \text{(5.8.8)}\ \cancel{(8)}$$

for all feasible v, c. The left inequality in (8) states that, given v^o, c^o is chosen to maximize total crime; the right inequality states that, given c^o, v^o is chosen to minimize total crime.

9.17a4

(Inequality) (8) is precisely the saddlepoint condition ~~which~~ That constitutes an equilibrium in ~~two~~ 2-person zero-sum games. Since both sides are colluding we have, in effect, just two decision makers; and the whole problem may be thought of as a game between a maximizing player, ~~Crim,~~ Crim, and a minimizing player, Vic, the payoff to Crim under the strategy pair (v, c) being

$$\int_S f(v,c)d\alpha.$$

do not break

The difference between (4)–(7) on the one hand, and (8) on the other, is that in the former the individual victim or criminal does not take account of the effects of his moves on the gains or losses of his "colleagues". In this respect, the difference is vaguely similar to that of a competitive vs. monopolistic industry, respectively. To put the matter in a slightly different, and slightly inaccurate, way, under individualism <u>average</u> gains or losses are equated over Space, while under collusion <u>marginal</u> gains or losses are equated.

Consider the very simple crime function

$$f(v,c) = vc \qquad (5.8.9)\ \ ~~(9)~~$$

(which is (2) with p set equal to zero). One easily verifies that the pair of <u>uniform</u> <u>distributions</u>

$$v^{\circ}(s) = V/\alpha(S), \quad c^{\circ}(s) = C/\alpha(S) \qquad (5.8.10)\ \ ~~(10)~~$$

(almost everywhere) is an equilibrium pair under <u>both</u> definitions. ($k_v = C/\alpha(S)$, $k_c = V/\alpha(S)$, (6) and (7) are satisfied with equality, and (8) is satisfied with equality for <u>any</u> feasible v, c, total crime being $VC/\alpha(S)$).

One also suspects that (10) is the <u>only</u> equilibrium pair under <u>either</u> definition. A non-rigorous argument would go as follows. Suppose~~, say,~~ that v were not constant almost everywhere. Then the criminals would all crowd into the region where v was densest. But then victims would ~~nof~~ not be in equilibrium, since they could move into the region vacated by criminals. A similar argument applies if c were not constant almost everywhere.[12]

Our aim now is to generalize this argument and make it rigorous. We prove this separately for the individualistic and collusive cases. Each of these cases has its own appropriate class of functions for which the statement is proved, and (9) belongs to both classes.

Theorem: Given measure space (S,Σ,α), with $\infty > \alpha(S) > 0$, positive real numbers, V and C, and a measurable function f: non-negative reals^2 $\rightarrow$ non-negative reals ~~which~~ that satisfies:

(i) $f(v,c) = 0$ iff either $v = 0$ or $c = 0$ (or both);

(ii) the right-hand partial derivatives $D_1f(0,c)$ and $D_2f(v,0)$ exist for all real $v, c \geq 0$, and, if $v > 0$, then $D_2f(v,0) > 0$;

(iii) $f(Vx, Cx)/x$ is a strictly increasing function of x $(x > 0)$.

Then (10) is the unique individualistic equilibrium pair.

Proof: ~~One~~ We ~~verifies~~ verify at once that (10) satisfies (4)–(7) almost everywhere: (4)–(5) follows from the constancy of v^o and c^o, while (6)–(7) are trivial because v^o and c^o are positive. Hence it remains only to show the uniqueness of this solution.

Let $(\underline{v}^o, \underline{c}^o)$ be an individualistic equilibrium pair, and let $\underline{E} = \{s | \underline{v}^o(s) > 0\}$, $\underline{F} = \{s | \underline{c}^o(s) > 0\}$. Suppose first that $\underline{k}_c = 0$. It follows that $\alpha(\underline{E} \cap \underline{F}) = 0$, for otherwise (5) would be violated, since $\underline{f}(v,c) > 0$ if $\underline{v} > 0$ and $\underline{c} > 0$. Also, we must have $\alpha(\underline{E} \setminus \underline{F}) = 0$; for otherwise (7) would be violated, since $D_2 f(\underline{v}, 0) > 0$ on this set. But this means that $\alpha(\underline{E}) = 0$, which implies the feasibility condition (3) is violated for $\underline{v}^o$, since $\underline{V} > 0$. We have a contradiction, and it follows that $\underline{k}_c > 0$.

Note!

This implies that $\alpha(\underline{F} \setminus \underline{E}) = 0$; for otherwise (5) would be violated, since $\underline{f}(0,c) = 0$. Hence $\alpha(\underline{E} \cap \underline{F}) > 0$; for otherwise $\alpha(F) = 0$, violating (3) for $\underline{c}^o$. Now for almost all points, s, of $\underline{E} \cap \underline{F}$ we have

START p. 17.5

Back. Grali 4c vers see pg. 499

$$k_v \underline{v}^o(s) = \underline{f}(\underline{v}^o(s), \underline{c}^o(s)) = k_c \underline{c}^o(s), \qquad \cancel{(11)}\ (5.8.11)$$

from (4) and (5). Since $\alpha(\underline{E} \cap \underline{F}) > 0$, there exists $\underline{s}$ satisfying (11). Hence $\boxed{k_v} > 0$. It follows that $\alpha(\underline{E} \setminus \underline{F}) = 0$; for otherwise (4) would be violated, since $\underline{f}(\underline{v}, 0) = 0$. Also $\alpha(\underline{A} \setminus (\underline{E} \cup \underline{F})) = 0$; for otherwise (6) would be violated, since $D_1 f(0,0) = 0$.

We have now shown that $\underline{v}^o > 0$ and $\underline{c}^o > 0$ almost everywhere, hence (11) is valid almost everywhere. Integrating (11) over $\underline{S}$, we find, from (3), that $Vk_v = \underline{C}k_c$. Hence

$$\underline{v}^o/V = \underline{c}^o/C \qquad \cancel{(12)}\ (5.8.12)$$

almost everywhere. Letting $x(s)$ be the common ratio in (12) at the point $\underline{s}$, we find, from (11) that

$$Vk_v = \frac{f(Vx(s),\ Cx(s))}{x(s)} = Ck_c, \qquad \text{(5.8.13)}\ ~~(13)~~$$

almost everywhere. The middle term in (13) is strictly increasing in x, hence there is just one solution x^o: v^o and c^o must be constant almost everywhere, which yields (10). ///

CONT. gal. 17-5 500 to go!

We now give the corresponding result for collusive equilibrium. The uniqueness, rather than the existence, of equilibrium is the more interesting condition, and that is ~~what is~~ established in the following theorem.

Theorem: Let (S,Σ,α) be a measure space, with $\infty > \alpha(S) > 0$, and α non-atomic; let V and C be positive real numbers; let measurable function f:non-negative reals^2 → reals satisfy:

(i) for any fixed real $c \geq 0$, $f(v,c)$ is continuous in v, and differentiable with respect to v for $v > 0$ (notation: $D_1f(v,c)$);

(ii) for any fixed real $v \geq 0$, $f(v,c)$ is continuous in c, and differentiable with respect to c for $c > 0$ (notation: $D_2f(v,c)$);

(iii) for any fixed real $c > 0$, $D_2f(v,c)$ is strictly increasing in v;

(iv) for any fixed real $v > 0$, $D_1f(v,c)$ is strictly increasing in c.

Then (10) is the only possible collusive equilibrium pair.

Proof: Let v^o, c^o be a collusive equilibrium pair. c^o is then unsurpassed for the problem of maximizing

$$\int f(v^o(s), c^o(s))\,\alpha(ds) \qquad \text{(5.8.14)} \; \cancel{(14)}$$

329

over the non-negative densities $c{:}S \to$ reals satisfying $\int_S c \, d\alpha = C$. Since α is non-atomic and $f(v,\cdot)$ is continuous, all $v \geq 0$, we have as a necessary condition for this that there exist an extended real number, k, and a null set $E \in \Sigma$, such that, for all $s \in S\backslash E$, $c^o(s)$ maximizes

$$f(v^o(s), x) - kx \qquad \text{(5.8.15)} \; \cancel{(15)}$$

over non-negative real x (section 5.4 ~~above~~).

The "shadow-price" k must, in fact, be ~~real.~~ finite. For if $k = -\infty$, (15) has no maximizer, while if $k = +\infty$, c^o would be zero almost everywhere, which violates $\int_S c^o \, d\alpha = C > 0$.

We now show that, if $s_1, s_2 \in S\backslash E$, and $v^o(s_1) > v^o(s_2)$, then

$$\text{either } c^o(s_1) > c^o(s_2), \text{ or } c^o(s_1) = c^o(s_2) = 0. \qquad \text{(5.8.16)} \; \cancel{(16)}$$

Condition (16) will be demonstrated by eliminating two possibilities.

Possibility ~~(1)~~ A: $c^o(s_1) = c^o(s_2) > 0$. Since (15) is differentiable in x, the derivative must be zero at these respective points, ~~That is,~~ i.e.,

$$D_2 f(v^o(s_i), c^o(s_i)) = k \qquad \text{(5.8.17)} \; (17)$$

($i = 1,2$). But $D_2 f(\cdot, c')$ is strictly increasing, (c' being the common positive value of $c^{o}(s_1) = c^{o}(s_2)$). Hence at least one of the two equations in (17) must be invalid.

possibility ~~(ii)~~ B: $c^{o}(s_2) > c^{o}(s_1)$. Consider

$$f(v^{o}(s_1),\, x) - f(v^{o}(s_2),\, x) \qquad \text{(5.8.18)} \ \cancel{(18)}$$

as a function of the non-negative real variable x. Function (18) is continuous, and has a positive derivative for all $x > 0$. Hence, using the mean value theorem, (18) is strictly increasing in x. It follows that

12pt brackets

$$f\left\{v^{o}(s_1),\, c^{o}(s_2)\right\} - f\left\{v^{o}(s_2),\, c^{o}(s_2)\right\}$$
$$> f\left\{v^{o}(s_1),\, c^{o}(s_1)\right\} - f\left\{v^{o}(s_2),\, c^{o}(s_1)\right\}, \qquad \text{(5.8.19)} \ \cancel{(19)}$$

Set on one line if possible

Also, since $c^{o}(s_i)$ maximizes (15) for $s = s_i$, ($i = 1, 2$), we get

12pt brackets

$$f\left\{v^{o}(s_1),\, c^{o}(s_1)\right\} - kc^{o}(s_1) \geq f\left\{v^{o}(s_1),\, c^{o}(s_2)\right\} - kc^{o}(s_2), \qquad \text{(5.8.20)} \ \cancel{(20)}$$

and also (20) with subscripts 1 and 2 interchanged. Adding these three inequalities (19)–(20) and simplifying, we get the contradiction $0 > 0$. This eliminates possibility ~~(ii)~~ B and establishes (16).

This entire argument may now be repeated with the rôles of v^{o} and c^{o} interchanged, the only difference being that v^{o} is maximizing the negative of (14). Since $D_1[-f(v,c)]$ is strictly decreasing in c, we obtain the analog of (16) again, but with one inequality sign reversed:

If $s_1, s_2 \in S \setminus E'$ (E' being a certain null set) and $c^o(s_1) > c^o(s_2)$, then

$$\text{either } v^o(s_1) < v^o(s_2), \text{ or } v^o(s_1) = v^o(s_2) = 0. \qquad \text{(5.8.21)} ~~(21)~~$$

Finally, suppose ~~there are~~ two points $s_1, s_2 \in S \setminus (E \cup E')$, exist such that $v^o(s_1) > v^o(s_2)$. We cannot also have $c^o(s_1) > c^o(s_2)$, for then (21) would lead to a contradiction. Hence $c^o(s_1) = c^o(s_2) = 0$, by (16). For any other point $s \in S \setminus (E \cup E')$, choose s_1 or s_2, depending on which s_i satisfies $v^o(s_i) \neq v^o(s)$. The argument just given then shows that $c^o(s) = 0$. Hence, if v^o is not constant on $S \setminus (E \cup E')$, c^o is identically zero on this set. Since $E \cup E'$ is a null set, this gives the contradiction

$$0 = \int_S c^o \, d\alpha = C > 0.$$

Hence v^o is constant almost everywhere. A similar argument shows that c^o is constant almost everywhere. ~~This completes the proof.~~

This does not prove that the pair of uniform densities (10) is a collusive equilibrium pair: There is the possibility that no such pair exists. To ~~test whether, in fact,~~ (10) is a collusive equilibrium pair is not difficult. stet ~~Note that,~~ under the premises of this theorem, the shadow price conditions are both necessary and sufficient for equilibrium ~~(page~~ 000).

Hence (10) is such a pair iff there are numbers k_1 and k_2 such that $c/\alpha(S)$ maximizes

$$f(V/\alpha(S),\ x) - k_1 x \qquad \text{(5.8.22)} \quad \cancel{(22)}$$

over $x \geq 0$, and $V/\alpha(S)$ maximizes

$$-f(x,\ C/\alpha(S)) - k_2 x \qquad \text{(5.8.23)} \quad \cancel{(23)}$$

over $x \geq 0$. If ~~these~~ conditions (22) and (23) are added to the premises, we may assert that the pair of uniform distributions is the unique collusive equilibrium pair.

gul 1

~~There are~~ Some real-world situations ~~which~~ illustrate these results, at least approximately. The policy of dispersing population to reduce losses from air attack is an example. If carried out to the limit, both targets and attackers would be uniformly distributed. In a uniform environment, a predator and a prey species would tend to become uniformly distributed. The distribution of Christians and lions in the Roman arena must have been roughly uniform.

24

Case (iii): Now we introduce police. The distribution of victims is given and fixed, while police and criminals are freely mobile. The former try to reduce, the latter to increase, total crimes. As the crime density function we take

$$f(v,\ c,\ p) = g(v)\,ce^{-p} \qquad \text{(5.8.24)} \quad \cancel{(24)}$$

where g is strictly increasing, nonnegative and finite. ~~real-valued.~~

This is just a generalization of (2). One can obtain results for more general functions, but (24) leads to a very simple and elegant equilibrium with some provocative implications.

In the following result, premise (26) is introduced to avoid an uninteresting complication; it is quite weak. Also premise (25) could have been deduced from rather weak assumptions, but we take the simpler course of assuming it outright. All logarithms are to base e.

<u>Theorem</u>: Let (S,Σ,α) be a measure space, with $\infty > \alpha(S) > 0$; let C and P be positive real numbers; let $v{:}S \to$ non-negative reals be measurable, with $\int_S v\, d\alpha > 0$ (i.e. there are some victims); let g:non-negative reals $\to$ non-negative reals be strictly increasing. Assume:

(i) there is exactly one number L satisfying

$$\int_{\{s|g(v(s))>L\}} \log\left[\frac{g(v(s))}{L}\right]\alpha(ds) = P; \qquad (5.8.25)\ \cancel{(25)}$$

(ii) for this L, $\alpha\{s|g(v(s)) = L\} = 0.$ $\qquad$ (5.8.26) ~~(26)~~

Then there is <u>exactly one</u> collusive equilibrium pair (c^o, p^o) for the problem: Maximize over c, minimize over p:

$$\int g(v(s))c(s)e^{-p(s)}\alpha(ds).$$

Here $c{:}S \to$ reals and $p{:}S \to$ reals must be non-negative measurable, and must satisfy

$$\int_S c\, d\alpha = C, \quad \int_S p\, d\alpha = P. \qquad (5.8.27)\ ~~(27)~~$$

Apart from a null set, c^o and p^o have the following form: ~~If~~

$$\text{If}\quad g(v(s)) \le L, \text{ then } c^o(s) = p^o(s) = 0. \qquad (5.8.28)\ ~~(28)~~$$

If $g(v(s)) > L$, then

$$c^o(s) = C/\alpha\{s \mid g(v(s)) > L\} \quad \text{(a constant)}, \qquad (5.8.29)\ ~~(29)~~$$

and

$$p^o(s) = \log\left[g(v(s))/L\right]. \qquad (5.8.30)\ ~~(30)~~$$

Proof: Let $g(v(s)) = h(s)$. First we show that (28)–(30) is the only possible collusive equilibrium pair. Let c^o, p^o be such a pair. The conditions for the existence of shadow prices are satisfied, since $h(s)xe^{-p^o(s)}$ is concave in x, and $-h(s)c^o(s)e^{-x}$ is concave in x. Hence there are extended real numbers, k_c and k_p, such that (except on a null set), $c^o(s)$ maximizes

$$h(s)e^{-p^o(s)}x - k_c x \qquad (5.8.31)\ ~~(31)~~$$

over $x \ge 0$, and $p^o(s)$ maximizes

$$-h(s)c^o(s)e^{-x} - k_p x \qquad (5.8.32)\ ~~(32)~~$$

over $x \ge 0$.

We ~~shall~~ once and for all exclude the null set on which (31) or (32) is not maximized. Thus, "all s" means "all s in the complement of this set"; "there is a point" refers to the complement, etc.

First, $k_c < \infty$, for if not, then $c^\circ(s) = 0$, all s, violating (27). Similarly, $k_p < \infty$.

Next, $k_c > 0$. To see this, note that the assumptions on v and g imply that $\{s \mid h(s) > 0\}$ has positive measure. If $k_c \leq 0$, then on this set (31) would be strictly increasing in x, hence have no maximizer.

Next, $k_p > 0$. ~~For, first of all,~~ For if $h(s) = 0$, then $c^\circ(s) = 0$, since $k_c > 0$ in (31). From (27), there is a point s_1 for which $c^\circ(s_1) > 0$; hence also $h(s_1) > 0$. Now if $k_p \leq 0$, then for point s_1 (32) would be strictly increasing, hence have no maximizer.

If $c^\circ(s) = 0$, then $p^\circ(s) = 0$. This follows from $k_p > 0$ in (32). If $c^\circ(s) > 0$, then (32) is strictly concave in x and hence has (at most) one maximizer. This maximizer is zero iff the slope at $x = 0$ is non-positive. Thus we have

$$p^\circ(s) = 0 \text{ iff } h(s)c^\circ(s) \leq k_p. \qquad \text{(5.8.33)} \; \cancel{(33)}$$

If ~~(33) fails,~~ the slope at $x = 0$ is positive, the maximizer of (32) is obtained by setting the derivative equal to zero. ~~We find that~~ Thus, if $h(s)c^\circ(s) > k_p$, then

$$p^\circ(s) = \log\left[h(s)c^\circ(s)/k_p\right]. \qquad \text{(5.8.34)} \; \cancel{(34)}$$

Now the set $\{s \mid p^\circ(s) > 0\}$ has positive measure from (27). Hence $c^\circ(s)$ maximizes (31) on this set. Substituting from (34) into (31), we find that $c^\circ(s)$ maximizes

$[(k_p/c^\circ(s)) - k_c]x$

$$\left[\frac{k_p}{c^\circ(s)} - k_c\right]x \qquad (5.8.35)\ \cancel{(35)}$$

over $x \geq 0$, for any s such that $p^\circ(s) > 0$. Since $c^\circ(s) > 0$, the bracketed expression in (35) must be zero:

$$c^\circ(s) = k_p/k_c, \qquad (5.8.36)\ \cancel{(36)}$$

so that c° is constant on the set $\{s|p^\circ(s) > 0\}$.

Next, the two sets $\{s|p^\circ(s) > 0\}$ and $\{s|h(s) > k_p/c^\circ(s)\}$ are the same. Hence, from (34) and (27),

set subscript to integral

$$\int_{\{s|h(s)>k_p/c^\circ(s)\}} \log\,(h(s)c^\circ(s)/k_p)\,\alpha(ds) = P. \qquad (5.8.37)\ \cancel{(37)}$$

From (37) and (25) we obtain $k_p/c^\circ(s) = L$, if $p^\circ(s) > 0$, so that, from (36),

$$k_c = L. \qquad (5.8.38)\ \cancel{(38)}$$

Substituting (38) and (36) into (34), we obtain

$$p^\circ(s) = \log[h(s)/L] \qquad (5.8.39)\ \cancel{(39)}$$

wherever $p^\circ(s) > 0$ (so that $h(s) > L$ on this set).

Now consider the maximization of (31). If $c^\circ(s) = 0$, then $p^\circ(s) = 0$, and the fact that $x = 0$ maximizes (31) implies $h(s) \leq k_c$. Hence, from (38),

if $c^\circ(s) = 0$, then $h(s) \leq L$.

Next let $c^\circ(s) > 0$. The case where $p^\circ(s) > 0$ also holds has already been discussed, yielding (35) and (36). Suppose $p^\circ(s) = 0$. The fact that $x > 0$ maximizes (31) yields $h(s) = k_c = L$. But from (26) this occurs only on a set of measure zero.

We are now finished, ~~For~~ as the two sets, $E = \{s | p^\circ(s) > 0, c^\circ(s) > 0\}$ and $F = \{s | p^\circ(s) = 0, c^\circ(s) = 0\}$ together exhaust S, except for a null set. On ~~the latter~~ set F, $g(v(s)) = h(s) \leq L$, and on ~~the former~~ set E, $g(v(s)) > L$. Condition (30) is the same as (39). Also, c° is constant on the set $\{s | g(v(s)) > L\}$, and zero off it, so (29) follows from (27). place on one line

To show that (28)–(30) actually gives an equilibrium pair, ~~one~~ we need only verify the shadow price conditions (31) and (32) for some k_c, k_p, since these are sufficient for unsurpassedness (in fact, for bestness). The pair (c°, p°) given by (28)–(30) do indeed maximize (31), (32), respectively, for $k_c = L$, $k_p = LC/\alpha\{s | g(v(s)) > L\}$. Verification is left as an exercise. ∎

The equilibrium solution (28)–(30) may be characterized as follows. There are two radically different régimes, a high-density régime, I (characterized by victim densities satisfying $g(v) > L$), and a low-density régime, II. In régime II there are no police, no criminals, and no crimes. In régime I, while density of police rises with that of victims, the density of criminals is uniform; so is the density of crimes, as one verifies by substituting (29) and (30) into (24).

This leads to the surprising conclusion that the most crime-ridden victims are those living at intermediate densities. For, since crime is uniformly distributed in régime I, crimes per victim must be inversely proportional to the density of victims. Starting at the highest victim densities, crimes per victim rise as victim density falls, reaching a peak and then suddenly falling to zero as the critical density is passed and régime II is entered.

Here is another slightly paradoxical implication. Suppose there is an anti-crime drive, and the total police force, P, is expanded. Since the integrand in (25) is non-increasing in L, the new equilibrium L must be lower. Total crime — which equals CL — does indeed fall. But in the process the critical victim density falls, and régime I — which is $\{s|g(v(s)) > L\}$ — expands at the expense of régime II. People who were living at densities just below the old critical density will suddenly find themselves engulfed in a crime wave, crimes per victim jumping from zero to the highest level in the system. All this is the result of increased law enforcement!

The explanation, of course, is that the increased "heat" on criminals in the old régime I induces them to disperse into the "greener pastures" of régime II.

The "spillover" effect of law enforcement in one community on the crime rate in neighboring communities has been recognized. It is sometimes claimed that better law enforcement decreases crime in neighboring communities.[13] This may be true with respect to the apprehension of wanted criminals. But insofar as the police serve the function of deterring potential criminals from committing crimes, the argument just given indicates that better law enforcement may well increase crime in the environs.

This sort of two-régime equilibrium is by no means unusual in game theory.[14] But we would expect any such effect to be blurred when applied to the real world. In general ~~one~~ we ~~does~~ do not find the intermediate density peaking of crimes per victim as predicted by this model (bank robbery may conform to this pattern, if ~~one~~ we ~~defines~~ define "victim" density properly). Instead, the usual pattern is for crimes per victim to rise with size of place, and to be inversely related to distance from the central city. There are, however, many exceptions, and different types of crime have distinct patterns.[15]

~~There are, of course,~~ In any number of ways ~~in which~~ the preceding model could have gone wrong. The three populations are not fixed in size and are not homogeneous. The crime function may be misspecified. Movement costs have been ignored. Finally, the motivations of the three populations may have been misspecified. In particular, it is not at all clear that police are allocated to minimize (an index of) total crime.[16,17]

FOOTNOTES - CHAPTER 5

415 [1]L. Robbins, An Essay on the Nature and Significance of Economic Science (Macmillan, London, 2nd ed, 1952), page 16.

---

416 [2]Some work has been done in generalizing (1)-(2) to a continuum. The calculus of variations may be applied in some special cases. Work beyond this point was begun by Bernard Koopman in 1956. See J. DeGuenin, "Optimum Distribution of Effort: An Extension of the Koopman Basic Theory", Operations Research, 9:1-7, January-February, 1961. Another exposition may be found in S. Karlin, Mathematical Methods and Theory in Games, Programming, and Economics (Addison-Wesley, Reading, Mass., 1959), Volume 2, Chapter 8, where the connection with the Neyman-Pearson lemma is stressed. Also see M. E. Yaari, "On the Existence of an Optimal Plan in a Continuous-Time Allocation Process," Econometrica, 32:576-590, October, 1964. For some more recent work on the Neyman-Pearson problem, see R. L. Francis, "On Relationships Between the Neyman-Pearson Problem and Linear Programming", pages 259-279 of Optimizing Methods in Statistics, J. S. Rustagi, editor (Academic Press, New York, 1971).

---

436 [3]To be precise, we never choose ε larger than c-x. This insures that x + y remains in the domain of f, so f(x + y) is well-defined.

---

[4] On derivates see E. J. McShane and T. A. Botts, *Real Analysis* (Van Nostrand, Princeton, 1959), pages 110-111.

[5] G. J. Stigler, *The Theory of Price* (Macmillan, New York, rev. ed., 1952), pages 119-120.

[6] (59) is even a bit stronger than the sufficient condition (2). In (59), the null set for which $\delta^\circ(a)$ does not maximize is chosen once and for all. But the null set for which (2) fails depends on δ, and conceivably there is no null set E such that (2) holds for all δ and all $a \in A \setminus E$.

[7] cf. G. B. Dantzig and P. Wolfe, "The Decomposition Algorithm for Linear Programs," *Econometrica*, 29: 767-778, October 1961.

(491) [8]The concavity and continuity of g_{θ} are of some independent interest, yielding a "comparative statics" result for the preceding theorem, as the parameter L varies in (3). 2610

---

(500) [9]Two similar cases come to mind. The law of mass action in chemistry takes reaction rate to be proportional to the concentrations of the reagents. In the Lotka-Volterra theory of predator-prey interaction, encounter frequency is again proportional to the product of the species densities. A. J. Lotka, Elements of Mathematical Biology (Dover, New York, 1956), ~~pages 88 ff.~~ pp. 88 ff. 1956

---

(501) [10]See E. M. Schur, Crimes Without Victims (Prentice-Hall Englewood Cliffs, N. J., 1965).

---

(502) [11]U.S. President's Commission on Law Enforcement and Administration of Justice, The Challenge of Crime in a Free Society (Dutton, New York, 1968); M. E. Wolfgang, "Urban Crime", in ~~Chapter 8 of~~ The Metropolitan Enigma, Ch. 8, J. Q. Wilson, ed. (Harvard University Press, Cambridge, 1968), esp. ~~especially pages~~ pp. 253–263, 276–281.

---

(509) [12]~~Incidentally,~~ using this argument as a guide ~~one~~ we can ~~easily~~ develop a dynamic model of redistribution of the two populations from a nonequilibrium position. We ~~shall~~ do not go into this.

(For example,

521 [13] ~~e.g., page 423~~ of C. M. Tiebout, "A Pure Theory of Local Expenditures," Journal of Political Economy, 64: ~~416-424~~ 423, (October 1956).

522 [14] See, e.g., M. Dresher, Games of Strategy (Prentice-Hall, Englewood Cliffs, N.J., 1961), ~~pages~~ pp. 124-127.

522 [15] E. H. Sutherland and D. R. Cressey, Principles of Criminology, (Lippincott, Philadelphia, 7th ed, 1966), ~~pages~~ pp. 187-191.

522 [16] On alternative criteria for the distribution of police, see C. S. Shoup, "Standards for Distributing a Free ~~Governmental~~ Service: Crime Prevention," Public Finance, 19: 383-394, 2, No.4, ~~1964~~; and his Public Finance (Aldine, Chicago, 1969), ~~pages~~ pp. 115-118.

stet

no. 4 (1964)

A.M. [17] A simplified version of Section 5.8 is to be found in Faden, Essays in Spatial Economics, Ph.D. Diss., Dept. Econ., Columbia Univ., 1967.