

Utilitarian Marginalism (Nozick, Rawls, Justice, and Welfare)

Abba P. Lerner*

One purpose of this article is to show the relationship of Robert Nozick's "Entitlement" and John Rawls' "Justice" to the utilitarian marginalism that derives from the classical economic approach to the distribution of income.¹ Nozick and Rawls claim to derive guides for income distribution from general principles of justice. But general rules for the distribution of income cannot be derived from abstract principles of justice because "justice" has no objective meaning. For any individual, justice is what he considers fitting, or *just fitting* the situation. It has been used to defend the traditional, like slavery; the inevitable or natural, like death and genetic differences; the expected, like hanging for stealing a sheep; the unexpected, like catastrophes called "acts of God"; and the satisfaction of revenge, sometimes called just punishment or punishment to fit the crime. It is hard to think of any inhumanity—one is tempted to say injustice—that has not at some time been defended as Demanded by Justice. I shall try to show that Nozick's and Rawls' suggestions for the desirable distribution of income are in fact based fundamentally, if not correctly, on nothing but (utilitarian) claims that they would yield higher levels of satisfaction.

A second purpose of this article is to examine the interdependence between two aspects of social efficiency—*distributive efficiency* in the

division, or sharing, of the social income among the different members of society, and *productive efficiency* in the allocation of the different productive resources among their different uses and of the different consumer goods among the different consumers of the incomes so divided, and their relation to the capitalistic and socialistic principles of economic organization.

More logical terms would be "division" of the social income and "allocation" of the different goods and services, but the traditional separation of economics into "production" and "distribution" is far too well established. We shall occasionally, however, find it convenient to speak of "allocative" efficiency to refer both to the different factors of production and the different consumption goods and services.

I. "Entitlement"

Nozick claims that his theory of "Entitlement" is rooted in "the laws of nature." In his "state of nature" individuals have perfect freedom to order their actions and dispose of their possessions as they think fit within the bounds of the laws of nature. This freedom requires the state to be minimal, its function limited to the provision of a justice system for the enforcement of contracts and the protection of its citizens against violence, theft, and fraud.

The entitlement theory states that a person is justly entitled to any holdings bestowed on him by nature or justly acquired by inheritance, or by free exchange for justly acquired holdings, i.e., when their acquisition has not violated any

*Queens College, CUNY and FSU.

¹The references are to Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974) and John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard, 1971).

individual's natural rights. If any natural rights have been violated in the acquisition or the transfer of holdings, such injustice should be rectified. Thus, whether a distribution is just depends on how it came about.

Nozick advances the usual economic arguments about the advantages of the division of labor, specialization, comparative advantage, and free exchange. Cooperation comes about through contract and exchange among individuals. Since the resulting distribution of income and wealth depends on the exchange ratios or prices at which the exchanges are made, justice depends on the prices being "fair." The prices in free exchange are judged as "fair" because people freely choose to make exchanges, and to transfer entitlements, with no restrictions on the freedom to trade with any other party on mutually acceptable terms.

There can be no serious objection to Nozick calling free exchange, and its effects, "just." Nobody is harmed by a free exchange and so there is at least nothing "unjust" about it. On the other hand, redistribution of income or wealth by the government (in taxing some to provide income or services to others) is declared by Nozick to be coercive and therefore "unjust." Nozick's "justice" therefore requires that redistributive actions by the government be limited to what cannot be avoided in financing the justice system.

But although the word "justice" is not very meaningful to me (and I remember that for a large part, perhaps a majority, of the world's population free exchange is criminal rather than just) I am in agreement with the substance of Nozick's argument. The situation *after* a free exchange is preferable (on utilitarian grounds) to the situation *before* the exchange. There is a net gain. Some have benefited from the exchange and none have been harmed. Nozick, however, draws a quite different conclusion. His conclusion is that the set of holdings after the exchange is a just one. In his enthusiasm for the net gains from exchanges freely under-

taken, in that it can only benefit the parties to it so that the changes brought about may be called just, he goes on to attribute this virtue of the *change to the total situation* after the change—to the new distribution of income and wealth—and declares that too to be just.

This assumes that the distribution was a just distribution before the exchange, and therefore could not be improved upon (made more just?) but only worsened by any redistribution. The application of this argument must ultimately credit "nature" with providing the *initial* inheritances in some just manner. But we have no reason to presume that "nature" somehow knows what it is doing and therefore distributes initial endowments of property and of abilities according to some concept of justice. Professor Israel M. Kirzner (in his article in this volume "Entrepreneurship, Entitlement, and Economic Justice") is more consistent than Nozick in attributing the virtues of an *act* of free exchange to the distribution of income and wealth in a *regime* of free exchange. He seems to defend the distribution "brought about" by free exchange (i.e., in existence after some free exchange) as just, and therefore not to be disturbed by any redistribution, even if the previous distribution is known to have been unjust.

Nozick's attribution of justice to the distribution existing in a free exchange economy seems to be based on a misunderstanding of the significance of "marginal product" for economics. There were one or two economists who claimed that paying a factor its (his?) marginal product is "just" because that is "what he (it?) produces." But almost all economists reject this argument. Accepting it would indeed made them the "apologists for capitalism" that Marxists accuse them of being. In this context paying "what he produces" means paying a *person* the value of the marginal product, either of his own personal activity or of the land or capital goods which he happens to own. There is certainly not much to the argument that a

person *deserves* (is "entitled" to) what is the result of his good fortune in inheriting these either legally or genetically. The proposition is an empty one from which no moral conclusion can be drawn. Nozick's attempts to bolster the proposition by reference to natural law do not make it any less vacuous.

There is, indeed, a powerful social reason why the price paid for the use of a factor should be equal to the value of its marginal product. But it is based not on justice, and not even on any concern at all as to who is going to receive that payment, as we shall see below. Its purpose is to serve the productive (allocative) efficiency of the economy by getting the factor to go to that use where the value of the marginal product is greatest, and not to a less important use where the value of its marginal product is less.

The distribution of income and wealth is, of course, affected by the degree of freedom of exchange, but it also depends on the distribution *before* the exchange; and that is the result of history. History is also a part of nature. And history is full of much that can certainly not be called just. To take seriously Nozick's final principle that rectification is called for where some injustice has been perpetrated, we must look into all the bloodshed, the invasions, the massacres, the oppressions, the enslavements throughout the history and prehistory of mankind. We would find very little justice left in the distribution to be protected by the minimalist state. Nozick's recommended "rectification" would call for just that kind of governmental intervention and redistribution that he rejects as "coercion."

We could claim that free exchange will yield a just distribution only by following Kirzner in completely disregarding both history and nature and "let the dead past bury its dead." This would require us to believe that all the inequalities of inheritance of property as well as of human capacities would have been rendered insignificant by the current benefits from a

regime of free exchange and free enterprise for a generation or two.

II. Maximin

Rawls comes very much closer to an economist's treatment of the subject. In his "difference principle" he implicitly recognizes the emptiness of "justice" and, by the device of redefinition, replaces it by "fairness." His basic principle is "Justice as Fairness." Fairness is the attribute of judgments made with no special preference for benefits to the judge himself. It is the absence of bias. Rawls dramatizes this most strikingly by his use of the concept of the "original position." In the imagined "original position" one is asked to choose between possible worlds into which to be born without knowing one's status, wealth, abilities, preferences, or luck. There is no possibility of distinguishing between oneself and others, and consequently one can only choose in a completely impartial way. Other writers in this field have defined the word "moral" in the same way that Rawls defines "just," namely, as unbiased or impartial. A "moral preference" is thus a preference for what is good for people in general with no special consideration for what is good for oneself or one's friends.²

I find myself in complete sympathy with Rawls' "moral" or "fair" approach to the problem of the distribution of income, but in applying it he makes the opposite error from that of Nozick. Whereas Nozick claims *too much* for the benefits from the equalization of relative marginal utilities by free exchange, extending these virtues to whatever distribution happens to exist (giving the credit to "nature"),

²Adam Smith defined moral rules of behavior as what would be recommended by an "impartially sympathetic observer." Professor John C. Harsanyi argues that "a value judgment would show the required impersonality if the person who made the judgement had to choose . . . in complete ignorance of what his own position (and the position of those near to his heart) would be within the system." (*Journal of Political Economy*, October, 1953: pgs. 434-435.)

Rawls pays *too little* attention to the possibilities of marginal adjustments.

One of the functions of economists, and indeed one of their persistent burdens, is to keep on reminding people that there are *degrees* of almost everything. Thus, most recently, public interest in the environment has raised a cry for the *abolition* of various kinds of pollution, and economists have had to try to teach that there is an optimum *degree* of almost every kind of pollution. And for such improved adjustments we need the discrimination of marginal analysis.

Rawls distinguishes only between the "worst off" group in society and everybody else, giving no indication of the probabilities of being in one or another group, or of the level of income of the rest of the population and its distribution among them. He naturally comes to the conclusion that anybody who is deciding into which kind of world he would rather be born would protect himself by choosing that kind of world in which its worst off are not so badly off. It would, of course, be more desirable to be born into a world where one had a better chance of being better off, or of being much better off rather than not so much better off. But this marginal analysis is played down by Rawls, and in its absence any rational person, if he is forced to be "moral" or "just" (in the sense of "unbiased") by being put in the "original position," would indeed follow Rawls' "maximin" principle. He would maximize the minimum income by choosing to be born into that kind of world in which the worst off group is better off than the worst off group in any other of the possible worlds.

Rawls views the initial or "natural" distribution of material assets, as well as of natural abilities, as "morally arbitrary," so that no person has a moral claim to any particular product share. Those who would be better off (if paid the value of the marginal product of their work and of the material factors they owned) should make transfers to the worse off.

This is only just (i.e., fair) because it brings the distribution closer to what would be chosen if there were no bias, namely, what would be chosen in the "original position."

Rawls also says that the better off in society make transfers to the worse off in order to gain and maintain their cooperation. This argument can be turned around. One can just as logically argue that the worse off give rewards to the better off to induce them to cooperate and make the worse off not as badly off as they otherwise might be. But this argument is not really basic to Rawls' model of the impartiality or morality of the choice in the "original position." It is, of course, a fact that there are transfers from the rich to the poor in the actual world, but that is not because of a widespread acceptance of Rawls' maximin principle. It is more due to the partial but very general acceptance of the utilitarian arguments with the marginal analysis which Rawls plays down.

It would seem that Rawls avoids marginal analysis because of an over-reaction to the Nozick type of error—his misuse of marginalism to defend the final distribution of income in a competitive free exchange economy as morally justified—and he throws out the baby with the bathwater. But by combining marginal analysis and utilitarian theory with the principle of "fairness" (with or without the dramatic device of the "original state") we can improve on Rawls' analysis even though in some cases we get the same conclusions.

III. Distributional and Productive Efficiency

In Bentham's utilitarianism, the general satisfaction of the community is measured by the sum total of the individual satisfactions or utilities. It is democratic and egalitarian, each individual's satisfactions receiving equal weight. Because of some philosophical difficulties in the meaning of *total* utility, some economists have instead spoken of maximizing *average* utility. But this too involves adding the total utilities

of individuals, and philosophical difficulties remain. It is possible to avoid all these difficulties by speaking instead of *marginal* utility to individuals from the consumption of goods and services. To compare the marginal utilities of individuals we do not have to go into the much more troublesome issue of measuring, or even estimating, any individual's *total* utility.³

If a gain in utility by an individual to whom a dollar of income is transferred is greater than the loss of utility by the individual from whom the dollar is taken (i.e., the marginal utility of income is greater to the first individual than to the second), then the shift of the dollar is an improvement because the gain is greater than the loss. If it is considered desirable to distribute a given total income most efficiently—to provide a larger rather than a smaller marginal utility wherever possible—this objective is achieved only by a distribution of income which makes all the marginal utilities of income equal. (It has to be a *given* total income so that the amount of income taken from one is equal

to the amount given to the other.) This is the condition for the efficient distribution of income.

Distributional efficiency, the principle governing the efficient division of income among the different individuals, is exactly parallel to productive efficiency, the principle governing the efficient allocation of resources among their different possible uses. That is obtained by shifting every resource from any point where the value of its marginal product is less to where it is greater. Only when these are equalized everywhere, and such further shifts are therefore no longer possible, is the efficient allocation of resources achieved. The same is true of the allocation of a consumer's income among the different consumption goods available to him. Its efficient allocation requires the equalization of the marginal utilities (which are the marginal products) of his different expenditures on the different consumption goods. This is included in our "productive efficiency."

³Having discovered that *ordinal* welfare functions are sufficient for microeconomic welfare theory in explaining market behavior in terms of observed acts of choice that reveal individual preferences, economists have rejected *cardinalism*, which had been used to explain market behavior in terms of unobservable individual utilities (although it can also be used exactly like ordinal utility).

Comparing the marginal utilities of income of an individual at different income levels (in discussing diminishing marginal utility of income), does not mean resorting to cardinality. Cardinality does imply the possibility of such comparisons, but the converse is not true. It is true that if we could compare the marginal utilities throughout the whole possible range of income, we would be able to reconstruct the whole cardinal utility function and to measure the *total utility*. But the comparison of total utilities does indeed suffer from the meaninglessness that has improperly been attributed to *all* utility comparisons. The comparison of *marginal* utilities is *not* meaningless.

The measurement of temperature has often been suggested as an analogy to the measurement of utility. But I have never been accused of meaninglessness when I said that I feel more keenly the difference between 60° and 65° in the swimming pool than that between 65° and 70°. However, it does not make any sense to

ask how keenly I feel the *total* temperature of the water—whether my zero is Celsius' or Fahrenheit's or Kelvin's.

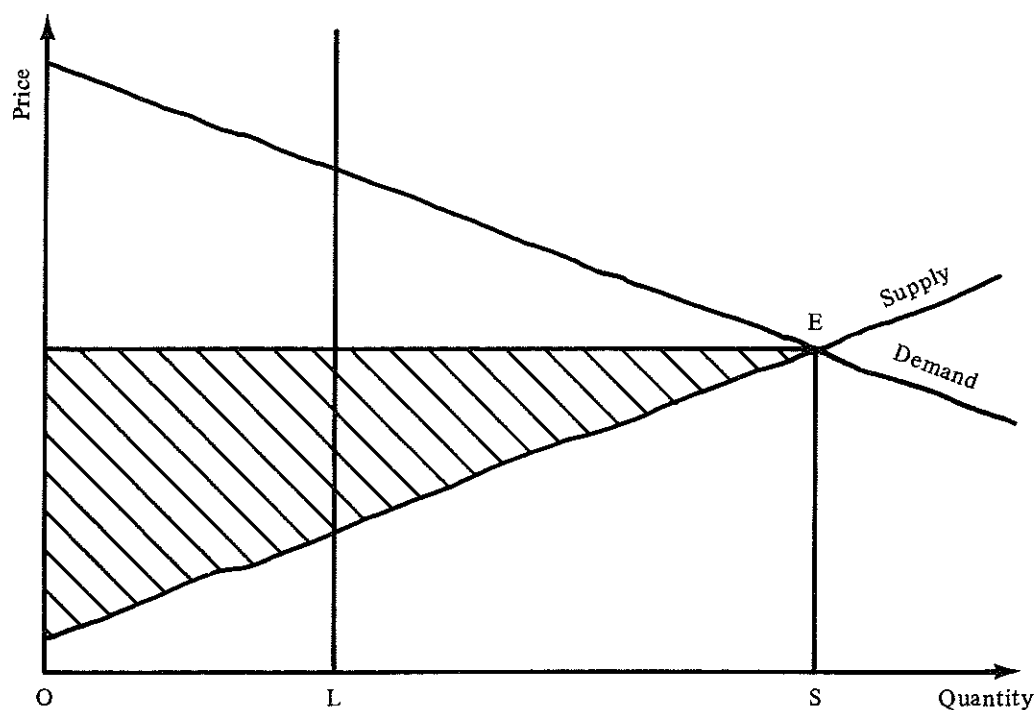
The resistance to the comparison of marginal utilities is very strong, and it seems to rest on the feeling that it is our duty to choose between ordinality and cardinality, and while ordinality does not permit such comparisons, the more permissive cardinality is itself outlawed. But for the comparison of marginal utilities, we do not have to assume *full* cardinality. What we need is a less forbidding name for that *part* of cardinality which is assumed in the comparison of *marginal* utilities. I find "marginalism" to be an adequate name, but that may be only because I was never properly trained in symbolic logic so that I am not tempted to call this cardinality. That may also be why the argument as put forward by me in 1934 was very slow of acceptance—even frowned upon as an unfortunate flaw in my *Economics of Control*—while Harsanyi's use of it in 1953 and 1957 was hailed by Paul Samuelson as a "rehabilitation of cardinal utility" which constituted "a rare quantum-jump breakthrough in welfare economics." But it is not really a rehabilitation of cardinal utility (even though Harsanyi himself says he might as well call it that). It is a smaller jump, and as such, even more to be admired as an important breakthrough.

Allocative efficiency is reached automatically in a free competitive market. If any factor of production has marginal products of different value in different uses, it would pay somebody to shift some of the factor from where the value of the marginal product is less to where it is greater. The difference would be a net gain or profit. But the market does not in the same way equalize the marginal utility of income. In a free competitive market it is the *owner* of every factor who is paid the value of the factor's marginal product. Each individual's income is equal to the marginal product of each factor multiplied by the number of units of that factor that he owns. His income is represented by the area of the rectangle (OE) in the diagram. The height of the rectangle (SE) shows the price or pay per unit of the factor, while its length (OS) shows the quantity he owns. There is no reason for assuming that the distribution of ownership, and consequently of income, would be such, and that the capacity of individuals to

enjoy income would be such, as to make the marginal utility of income to all the different individuals anything like equal. Equalization of the marginal utilities can therefore be achieved only by transferences of income from those with smaller marginal utilities of income to those with greater marginal utilities until the marginal utilities had all become equal.

Some of such transferences are voluntary gifts (charity) to those whose needs seem greater (i.e., who seem to have a greater marginal utility of income). These gifts diminish the inequality but clearly they do not eliminate it. Further equalization of the marginal utility of income can therefore come about only by government redistribution of income—denounced by Nozick (and Kirzner) as coercion and contrary to "justice" since it interferes with the complete liberty of the state of nature.

We have considered these objections to attempts to combine distributive with allocative



efficiency, and there are some more hurdles, but I shall try to show that these can all be overcome.

We have seen that distributive efficiency—the optimal division of income—is obtained only if the marginal utility of income is equal for all individuals. In trying to achieve this by re-distribution of income we immediately come up against a serious difficulty. There is no way of discovering whether any individual's marginal utility of income is greater than, equal to, or less than, that of any other individual! Every individual could declare that he has an exceptionally high capacity for satisfaction and so the marginal utilities of income could be equalized only by giving him more income than anybody else; and there is no way of testing the validity of such claims.

IV. Income Equalization—Maximizing Probable Gain

Nevertheless it is possible to redistribute income so as to maximize the *probable gain*. The probable gain from any redistribution is maximized by dividing the income *equally*. This proposition rests on the following assumptions: (1) other consumers besides me are capable of feeling satisfactions; (2) every individual gets more satisfaction from a larger income than from a smaller income; (3) whenever a consumer can choose among two or more alternatives he chooses the one that yields the greatest satisfaction; (4) the satisfactions experienced by different people are *similar*, so that it is not meaningless to say that the satisfaction which one individual derives from an additional unit of income is greater, or that it is less than, the satisfaction that would be derived from it by another individual.

From these assumptions (essentially, that others have feelings, and some rationality, like me) it follows that the marginal utility of income is diminishing. The *extra satisfaction* the individual gets from a given increase in his

income (the marginal utility of income) is less if his original income is greater.⁴ (This can be represented by a downward sloping curve of the marginal utility of income.)

The establishment of the principle of diminishing marginal utility of income is of crucial importance. If indeed consumers spend their income in a way that maximizes their satisfaction, the things bought give a greater satisfaction than the other things that could have been bought instead with the same income, but were not bought for this very reason. It therefore follows that if income were greater, the additional things that would be bought would be "next best"—the things that are rejected when incomes are smaller because they give less satisfaction. And if income were greater still, even less satisfactory things would be bought. The greater the income, the less satisfactory are the additional things that remain to be bought with the additional increase in income. That is all that is meant by the principle of diminishing marginal utility of income.

I further invoke three more assumptions which somewhat complicate the issue and might seem to cast some doubt on the principle of diminishing marginal utility of income. I will show that even if these three additional assumptions are not warranted, the general conclusion still holds. Initially I assume further (5) that the different satisfactions

⁴My initial statement on the principle of the diminishing marginal utility of income appears in Abba P. Lerner, *The Economics of Control* (New York: The Macmillan Company, 1944, Chapter 3). In that work I discuss both distributive efficiency and productive efficiency but I do not adequately integrate them, leaving the implication that a compromise between two contradictory principles is required i.e., that *both* principles have to be compromised. In the present article I argue that this is not the case. It is possible for only *distributive* efficiency to have to give way, while productive efficiency is untouched; the limitation to the full achievement of distributive efficiency is due not to the necessity of satisfying the basic condition for productive efficiency, but to quite different forces.

obtained by a person from the consumption of different goods are independent of the size of his income and therefore of the other goods consumed. Initially I assume (6) that the experience of having a larger or smaller income does not develop or dull a person's tastes or capacities for enjoying income. Initially I assume (7) that an individual derives more satisfaction from an extra \$1.00 of income for himself than from an extra \$1.00 of income for anyone else.

Consider a transfer of income from a richer to a poorer individual. If both individuals have *the same* overall capacity for deriving satisfaction from income (which would be represented by *identical* downward sloping curves of the marginal utility of income) such a transfer would yield a net gain. For in that case the poorer individual has a higher *marginal* utility of income. (The identical marginal utility of income curve is *higher* at the point corresponding to the smaller income.)

Such overall equality of capacities for satisfaction cannot, however, be assumed to be necessarily the case. The richer individual may have either a greater or a smaller capacity than the poorer one (a higher or a lower marginal utility curve). If the poorer person has the greater capacity, the gain is increased on that account. If the richer person has the greater capacity, the gain is diminished on that account. It may even be converted into a loss. The possibility of an increase in gain offsets the possibility of a diminution of gain since these are equally likely to occur in any particular case. The net gain (which is a certainty in the case of equal capacity) then becomes only a *probable* gain on account of the possible increase or diminution of the gain which arises with unequal capacities. *Every transfer of income from a richer to a poorer individual therefore shows a probable gain.* From this it follows that the rational procedure is to keep shifting from richer to poorer until the income differences have disappeared, i.e., to divide a given total income *equally*.

I shall now consider the complications mentioned above. The argument for diminishing marginal utility of income holds strictly throughout every individual's whole income range only on the assumption (assumption 5) that the satisfactions than an individual derives from the consumption of different goods are independent of each other. If these utilities have complementary relationships it is possible for the marginal utility of income to increase with increased income instead of diminishing. Items that are rejected and not bought at one income level may nevertheless be bought at a greater income not just because they are "next best" although less satisfying than what is bought at a lower income, but because other things are now being consumed by the individual and these increase the utility of the new items.

Complementarities can, however, be either positive or negative, and negative complementarity *reinforces* the principle of diminishing marginal utility of income. Only positive complementarity tends to mitigate, and may occasionally even reverse, the diminishing marginal utility of income. It does this by creating a kind of discontinuity which, if it is powerful enough, can raise a "bump" on the marginal utility curve, possibly making a portion of the curve slope upward with increasing income.

But the "bump" always has a second side which, when the complementarity effect has been exhausted, must make the curve fall *more* than it rose (if it actually did rise). The marginal utility of income will still be diminishing with the same *average* intensity but will be subject to irregularities. The principle of diminishing marginal utility therefore reduces to a probability. But this is all that we needed anyway.

A similar answer can be given to the objection that many decisions are foolish or hasty or based on insufficient or faulty information, so that they do not result in that expenditure being chosen which yields the greatest satisfaction. The results of such decisions are a matter of luck or chance. But as long as a

significant part of the decisions are based on deliberate judgement and reliable information, our conclusion still holds. These decisions will lead to desired results more often than the opposite, and their results in any particular case will *probably* be preferable to the results of the deliberately rejected choices, so that the results of every actual choice will still *probably* be preferred to the results of the rejected choices; and that again is all that we need.

Another problem arises if we relax assumption (6). The experience of having a larger income may develop a person's tastes and capacities for enjoying income so that the marginal utility curves of rich people would be higher than those of poor people. An ideal distribution of income would then give more to people who have been richer. The converse of this too may be argued with about the same plausibility. A person with high income gets used to luxuries so that he hardly notices items that would give a great thrill to a poor man unused to them. This would reverse the first argument and *strengthen* the case for equalizing incomes. However even if the first consideration were known to outweigh the second, and even if of the two arguments only the first were found to have validity, there would still be a *probable* gain from an egalitarian division of income. But this would now be true only in a longer run. If experience of higher income really had the effect of *raising* the curve of marginal utility of income, it would be best *temporarily* to leave more than the average of income with those whose levels of income has been higher. This still does not alter our conclusion. It only addresses itself to how rapid should be the *transition* from a previous unequal to an ideal egalitarian distribution of income, and to the importance of avoiding the establishment of inequalities of income in the first place.

Some might argue that a person feels a greater increase in satisfaction from a given income *increment* than he would feel from the same-sized income *decrement*. But this comes down to the same problem already considered—

that of the rich who are supposed to have acquired a greater sensitivity to income. Greater sensitivity to diminutions of income would merely lead to the conclusion that redistribution should be gradual. It would still follow from our analysis that no income above the average should be increased and that no member of the new generation should be given an income above the average, so that the acclimatization to high income should not arise in the first place. But where it has arisen, the *speed* of redistribution of income from rich to poor must be limited to the rate of redistribution at which the harm by the change itself is just great enough, at the margin, to offset the gain from the improvement in the distribution of income. A more rapid rate of redistribution would do more harm than good at the margin, while a less rapid rate of redistribution would mean the abandonment of some benefit from improved distribution that is greater than the harm that would be done by the greater rate of change.

A qualification might appear to be necessary for differences in income due to greater effort. When an individual applies himself more assiduously to his work this might be held to indicate that he has greater use or need for the extra income derived from the extra effort. Some might contend that such an individual should receive a larger income. This does not follow. A person works hard only when the wage plus the attractiveness of the work itself (or minus the irksomeness of the work), in relation to the opportunity cost of the alternative leisure forgone, is greater than for others. But there is no more reason for believing that a person works harder because an additional dollar is worth more to him than that he works harder because the marginal disutility of work or the marginal utility of leisure is less for him. We know only that these are less in relation to the marginal utility of consumption goods. But it is all of these marginal utilities *combined* that constitute the marginal utility of income and there is no reason here for believing that that is higher or lower than for others.

Assumption (7) is that each individual derives more satisfaction from an increment to his own income than from the same increment to the income of others. If the poor derive more satisfaction from seeing the splendors of the rich increase by \$1.00 than from the alleviation of their own poverty by \$1.00, our argument might be upset, but then they could, and would, donate some of their income to the rich until this were no longer the case. Conversely, the pain that even those who are not poor get from the existence of poverty, and the envy that the poor feel for the rich, would tend to strengthen the argument for equality.

Sometimes the spectacular expenditure of the rich is identified with cultural values. Because cultural values are relative within a society this appears an extremely precarious thesis. In any event, our primary theorem remains intact. The rule for the most efficient distribution of any given income is to distribute it equally.

V. Material Incentives

It might seem that this conclusion, no matter how inescapable it is to all but those who have the arrogance to deny our basic assumptions that others, besides them, have feelings and some rationality, is rendered practically worthless by its complete dependence on the assumption of a *given* income, i.e., that the amount of income to be distributed is independent of how it is distributed. But there would certainly be less income produced if there were no material incentives for work and for efficiency. Payment for work equal to its marginal productivity is also a very important part of the general rule for productive efficiency—that the price paid for every factor of production should be equal to its marginal product. Thus, there appears to be a fundamental conflict between distributive and productive efficiency that would be absent only in extreme situations where the value of the marginal product of all kinds of human effort has fallen to zero.

Examples of this could be a condition of *absolute plenty* (sometimes called “full communism”) where productivity (capitalist or socialist) has outstripped the desire for goods and services (perhaps as a result of a philosophy of simpler living) or a condition of *absolute shortage* of land, tools and materials to work with—as in a besieged fortress (or in other cases of catastrophic overpopulation). We have never experienced absolute plenty, when there would be no distributive problem anyway, and in cases of extreme shortages, when total income is more or less given, income inequality has usually been diminished, in real terms, by egalitarian rationing. This indicates a general recognition of the distributive efficiency of egalitarianism and its dominance where productive efficiency becomes unimportant.

(It is of interest to note that on the assumption of a given total income to be shared, the maximin principle also yields the same result as the utilitarian marginal analysis. This is because the average income is then also the maximin income.)

But we live in a world where productive efficiency is of the utmost importance. For the richest countries it can be argued that greater productivity alone would merely raise “keeping up with the Joneses” to a higher intensity while speeding up the depletion of the planet’s natural resources, and that only an increase in distributive efficiency could improve the condition of man. But for most of the world even the complete equalization of income would merely spread the misery from the vast majority to all. Only an increase in output (per capita, of course) could bring significant relief.

Since material incentives are necessary to induce productive efficiency to increase the output—the income available to be distributed—the rule for the efficient distribution of a *given* income seems supremely irrelevant.

These issues present no problem for Nozick. His assumption of the justice of the distribution of holdings in a regime of free exchange removes any concern for efficiency of distribu-

tion, while his implied assumption that “nature” continually updates the laws that define property rights so as to eliminate all externalities, thus maintaining productive efficiency through the automatic competitive equalization of the prices of all factors to the changing value of their marginal social products.

Rawls, however, recognizing that changes in property rights and other social arrangements affect both the size and the distribution of the social income, is constrained by his maximin principle to approve of changes which increase output only if they improve the condition of the “worst off” group (or in case of a tie, of the next-worst-off group). Utilitarian marginalism also recognizes that differences in property rights and other social arrangements affect the size and distributions of the social income, but it leads to Rawls’ maximin conclusion only on the assumption of a universal absolute and overriding risk aversion. Absolute risk aversion would still lead a perfectly unbiased person to declare that he would have chosen the maximin world to be born in. But it would not make it moral for him to choose that for others unless he could assume that every one else suffers from the same absolute risk aversion. Furthermore, it raises the doubly difficult philosophical question whether the unborn judge in the “original position” should be supposed to know what his risk aversion or risk preference is (or would be?).

Rawls claims, however, that he does not base his maximin principle on marginal analysis and absolute risk aversion but on everybody, including those in the “worst off” group, having some “common property rights” in whatever can be produced in society, and so also in any increases on output. But this, if I understand it correctly, is a complete departure from his basic idea of Justice as Fairness, as guaranteed by the “original position,” and is close to some Nozick-like notion of “entitlement.”

Since the equal income theorem applies only to the distribution of a *given* total income, anything which *increases* somebody’s income with-

out reducing anyone else’s income must be accepted as a pure gain. Utilitarian marginalism here leads to the rule that the price paid for anybody’s extra (marginal) effort should be equal to the extra (marginal) product, bringing us back to the principle of allocative efficiency—the efficient allocation of time between work and leisure. The marginal effort will then be undertaken by everybody as long as this is beneficial to him, while there is just as much product left for everybody else. There is a net gain. Some are better off and none are worse off, even if none of those in the worse off group are among the beneficiaries.

If, as seems very likely, an increase in total income and wealth eases transfers from rich to poor and there is an increase in such transfers, we have an improvement in both productive and distributive efficiency. But is it possible to work more systematically toward achieving both types of efficiency?

VI. Unnecessary Payments

This appears to involve a fundamental conflict of mutually inconsistent principles. The rule called for by the principle of distributive efficiency stands squarely on a *given* or fixed income to be distributed. The rule called for by the principle of productive efficiency rests roundly on an *augmentable* income, maximised by paying the worker the value of his marginal product. How can this circle be squared?

The circle *can* be squared to the degree that *some* of the marginal product need not be paid. That part can then be treated as belonging to a *given* “surplus” which can be subjected to the distributive efficiency rule of equal distribution.

(This use of the surplus, though *possible*, is of course not *necessary*. Surplus has been appropriated throughout the ages for all sorts of other purposes.)

Until recently, it was not uncommon for economists to argue that income taxes and inheritance taxes could improve distribution without interfering with production. As long as the

marginal tax is less than 100%, it was argued, the point of maximum profit is not moved, and production is not affected. This has now gone out of fashion. Income is not pure profit and income taxes reduce the incentive to earn income. Only lump-sum taxes are seen as neutral, and they are not practical and are considered a theoretical curiosum. Also half forgotten is an older unpractical theoretical museum piece—the “faculty tax”—for eliminating income inequalities from work (inherited income having perhaps already been dealt with by progressive taxation of inheritances). The “faculty tax” is a lump sum tax equal to a person’s ability to earn income (his “earning faculty”) minus the average income. (If one’s “earning faculty” is less than the average income, his tax becomes a *negative tax*—a subsidy.)

The impracticability of the “faculty tax,” like the popular demand for abolishing pollution or Rawls’ exclusive concentration on the “worst-off” group, can be alleviated by bringing in the economist’s all-purpose solvent—their gradualness, continuity, or marginalism. What cannot be done in one fell swoop can be done by degrees; and in fact it is being done by degrees all over the place by lump sum taxes in various guises—or is it disguises?

The telephone user pays a lump sum tax (in this case a negative one—a subsidy) in the form of getting a certain number of calls free. As long as he uses more than this number of calls, what he is charged per call by the telephone system could be the marginal cost, which is the proper degree of discouragement. The “small business” may get some loans at specially low rates of interest, but if it still has to borrow some more money at the current rate, this does not cause it to use any of the capital inefficiently. Here again is a hidden lump sum subsidy. When a fixed amount of gas is sold by the producers to the distributors at less than the market price, the difference in money received is exactly equivalent to a lump sum tax paid by the former to the latter. But as long as this does

not apply to marginal production, the effect is again purely distributional. It has no effect on productive efficiency. Farmers in Israel (and elsewhere) are given a certain amount of water at much below cost. But as long as this supply is not enough fully to satisfy the demand, and they buy additional water (at a price that covers the marginal cost) up to the point where the price is equal to its marginal product, and if the farmers would be operating in any case, there is again a pure lump sum subsidy—a transfer of income that does not interfere with productive efficiency—that may be a contribution to distributive efficiency.

The general principle underlying the possibility of increasing distributional efficiency without diminishing productive efficiency is illustrated in our accompanying diagram, which shows the number of hours worked (OS) and the wage rate (SE), E being the equilibrium point at which supply is equal to demand. The rectangle (OE) represents, in its area, the price (SE) times the quantity (OS), the total amount paid.⁵

In this figure the supply curve shows that smaller amounts would be supplied at lower prices. The shaded area within the rectangle (lying above the supply curve) thus represents payments over and above what is necessary to make the supply available. Only the unshaded part of the rectangle, (that part of it which lies below the supply curve) represents the “necessary” payment that the suppliers have to get to be willing to make the supply available. The rest is “surplus” or “unnecessary payment.” The full-price (SE) is what has to be paid for the marginal unit, but all the other units could be obtained for less. This is shown by the supply curve being lower to the left of E.

Paying for *all* the units at the same marginal rate, may seem natural or even necessary, but this is only because it is our habit to start count-

⁵The development of this diagram appears in Abba P. Lerner, “The Economics and Politics of Consumer Sovereignty,” *Proceedings of the American Economic Association*, (May 1972): 261.

ing from zero. But this procedure is arbitrary. The count could begin at any other point—at, say, half the standard number of hours. If the standard number of hours of work in a particular industry is 6 hours, then counting the hours for the purpose of computing the pay would begin only after 3 hours of work.

VII. Degrees of Socialism

The later this count begins, i.e., the greater the number of “unpaid” hours, the more of the product would be left as a part of the “given” social product available for “free income” to be distributed in accordance with distributive efficiency, and the greater will then be the degree to which the society could be said to be socialistic.

This looks rather like a trade-off between allocative and distributive efficiency, but that is not the case. As long as the *marginal units* of factor services are paid their marginal product any increase in “unpaid” hours, and the consequent decrease in “paid” hours, is not at the expense of productive efficiency.

Coming back to our diagram, consider the possible vertical lines, like that at L, which could be drawn parallel to the ordinate as points from which we begin to count the “paid” hours. If we start at O on the extreme left, paying all the units at the marginal rate, we have the pure free exchange system with the allocation favored by Nozick and honored by him as “natural.” Moving the vertical line all the way to the right up to the point of intersection E, would mean having no “paid” hours at all with the whole of the output available for distribution in accordance with allocative efficiency. This is the other limiting case of pure socialism via the lump sum “faculty tax” and the equal distribution of income.

VIII. Resistances

A close approximation to such complete equalization of income is unpractical, but not

because it would interfere with productive efficiency. It is unpractical only because the population is not prepared to accept such extreme equality of income. The supplier will not be willing to supply his services unless the pay he gets for his “paid” hours covers the “necessary payment for all the hours he works.” The “lump sum” tax that takes the form of the “unpaid” hours must not be too great.

If L in our diagram represents the point at which we start the pay, the total pay will be shown by the area of the smaller rectangle LE. If the area LE is less than the unshaded part of the larger rectangle OE, he will not willingly supply the “free” hours of work for the sake of the earnings of the “paid” hours that follow.

If L is the point at which these two areas (LE and the unshaded area of the larger rectangle OE) are just equal, then starting the pay at any point between O and L would provide varying degrees of distributional efficiency but would retain complete allocative efficiency throughout. Starting the pay later, between L and S, would not work. It would be “too socialistic.”

The objective of trying to achieve as much distributive efficiency as possible without sacrificing any allocative efficiency can thus be expressed as trying to move L, the point from which the pay begins, as far to the right toward S as possible. But the area LE cannot be reduced below the “necessary payment.” What determines the necessary payment?

The payment that is necessary to obtain the use of a factor for a particular purpose is what the factor could get in the best paid alternative use. In a society organized with the social purpose of maximizing both kinds of efficiency, the alternative marginal rate of pay is technically determined by the value of the alternative marginal product. But the L point in the alternative occupation, the point at which the pay begins, is under the central government’s control. The total alternative earnings, and thus the necessary payment, can thus be made very low indeed. The L point can be shifted quite far toward S. This is limited only by the re-

maining options that cannot be government controlled. These are the myriad devices that can be invented by workers who do not feel that they are being treated fairly or justly and which will ultimately force the authorities to grant conditions that eliminate such resentment.

High among these resistances to "too much socialism" are the established traditions of *inequalities* of pay for different occupations. It is these feelings of the *propriety* of certain income differences that constitute the resistance to the "socialistic" distributional efficiency, and it is these that have to be weakened if greater distributive efficiency is to be attained while protecting maximum productive efficiency.

Another possible approach would be to set the L point at the same number of "unpaid" hours for everybody, with all workers free to work more hours and be paid the value of the marginal product for the additional hours, or to work fewer hours and be docked the value of the marginal product of the reduced hours from their "free income." This would induce the high marginal product workers to increase their efforts and their hours of work, and their already high total pay, until the increased marginal disutility of work rose to equality with the marginal utility to the worker of the high marginal pay; and conversely it would induce the low marginal product workers to cut down their efforts and their hours of work, and their already low total pay, until the marginal disutility of the work is reduced to equality with the marginal utility of the low marginal pay. Work with high value of marginal product would then become doubly attractive and conversely work with very low value of marginal product would become doubly unattractive.

Workers would then be induced to move from work with low marginal product to work with high marginal product. Both of these responses would tend to equalize the marginal product as well as the marginal pay, the distribution of marginal effort and the number of hours worked, and thus also the individual incomes.

It would improve both allocative and productive efficiency, while at the same time acclimating the public to more egalitarian and more socialistic distribution.

As this develops, the L point can be moved further to the right, and the "unpaid" hours and the "free income" increased. But all the time the increase in distributional efficiency does nothing to diminish productive efficiency. What limits the degree of distributional efficiency and the speed with which additions to it are achieved is not the protection of productive efficiency, but the traditions of legitimate inequalities, the political power of pressure groups, the threats of strikes, and, most of all, the inefficiencies from apathy in work which often cannot be distinguished from the conscious or unconscious sabotage of productive efficiency that grows out of feelings of resentment at whatever is deemed "unjust."

Such a development of increased distributional efficiency can properly be said to make society more socialistic. But whether it makes it less capitalistic is not so clear. It has to be so if capitalism stands for income inequality or distributional inefficiency—as it certainly has come to mean for many people. But if capitalism stands for making use of free exchange and competition in markets for the purpose of maximizing productive efficiency, then there is nothing anticapitalist in the socialism of greater distributional efficiency through "unpaid" inframarginal hours of work (and other productive factors) and "free income."

While such redistributions are usually given names with socialist connotations there are equally appropriate capitalistic words. Instead of being called "social grants" or "welfare payments" or even the hybrid "negative income tax," one could declare every citizen to have become a capitalist—to have inherited shares in "The Grand National Corporation," or even in specific local or multinational corporations, and the "free income" to be nothing but the dividends he is entitled to because they are earned by the shares he owns.

IX. Summary

I have presented three recipes for desirable income distributions, each claiming to yield higher levels of human satisfaction. But only my argument about the tandem nature of allocative and distributional efficiency and the diminishing marginal utility of income is based clearly on an analysis of the expected effects on human satisfactions.

Nozick's theory of *entitlement* justifies the distribution reached in a free exchange economy, presuming that all factors are "naturally" paid the value of their marginal product, and that such payments are "just." Rawls' *Maximin* principle can be derived on utilitarian principles, from an imagined choice in "the original position" only on the assumption of an overriding absolute risk aversion. It can also be reached from a concern limited to the welfare of the "worst off" or by jettisoning Rawls' Justice as Fairness (together with his "original position") and subscribing to a belief in some Nozickian "entitlement" of the "worst off" to a share in any *increase* in total social income.

The basic distributive theorem of utilitarian marginalism—that distributive efficiency calls for income equality—is strictly applicable only for a *given* total social income. Since the total social income is not given but depends on human effort, a compromise has to be made. The equality theorem must be adjusted to induce individuals to cooperate in exercising their full capacities. But it is argued that within a wide range the equalizing redistribution of income can take place without reducing total output. This is shown in an analysis of "necessary" and "unnecessary" payments which is illustrated in our diagram.

Utilitarian marginalism points the way to a combination of the maximization of productive efficiency, given the degree of perfection and competitiveness of the market, with a simultaneous maximization of distributional efficiency, constrained only by social, conventional and political resistances to extremes of egalitarianism. My colleague at Florida State University, Professor E. Raymond Canterbury has suggested that this policy be called the pursuit of the *Maximax*.