DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

# INTERGENERATIONAL JUSTICE AS OPPORTUNITY

Talbot Page


SOCIAL SCIENCE WORKING PAPER 389

Revised June 1982

Abstract

This paper contrasts two views of intergenerational justice. The first view is a global one, which focuses on the aggregate well-being of each generation and discounts future generations' utilities. In this view discounting is not a notion of intergenerational justice; instead it is defended as a necessary condition of intergenerational efficiency. And intergenerational efficiency is not advanced as a notion of intergenerational justice but as a strongly desirable condition of any intergenerational just system.

The appeal of the discounting approach is explained by a set of conditions which define neoclassical utilitarianism. Within the defining conditions, discounting future utilities is "natural," but not necessary for efficiency. The framework of intergenerational social choice also fits the defining conditions of neoclassical utilitarianism, and it is easy to construct choice rules which do not discount future utilities and yet which are intergenerationally efficient.

Although there is room within neoclassical utilitarianism for efficient rules of choice which do not discount future utilities, the second view of intergenerational justice does not appear to fit naturally within the utilitarian system. The second view of intergenerational justice is specialized, and focuses on the preservation of "essential" opportunities.

The second view becomes more appealing when the defining conditions of neoclassical utilitarianism are modified. In modifying the conditions, the notion of intergenerational efficiency becomes weaker, partly because as an ordering principle it becomes less complete and partly because potential Pareto improvements are no longer discretionary from the vantage point of the future. How much, if any, efficiency loss there might be from application of the specialized notion of intergenerational justice depends on the extent of modification of the defining conditions and the structure of institutions spanning generational time. The modified conditions appear to accord more closely with commonsense notions of intertemporal justice than do the original conditions. And thus the opportunity concept of intergenerational justice appears to be closer to our commonsense notions of intragenerational justice than does the global (discounting) concept.

# Intergenerational Justice as Opportunity*

by

Talbot Page

In managing the resource base we have to deal somehow with the potential of very long-lived costs.[1] For nuclear power a principal concern is with the effects of radioactive material; for oil, depletion; for coal, cancer, climate modification, and ultimately depletion. How much emphasis we give to conservation and other alternatives depends on how we think about these long-term costs. In this paper I attempt to distinguish between two views.

In the first view, long-term energy costs should be discounted and treated just like other future costs. This means that present and future energy costs are weighted and then added together. With a positive discount rate future costs are counted less heavily than present costs. If as a matter of equity or justice between generations, it is thought that future generations are going to be made to suffer too much from these costs (or any other costs imposed by the present generation) then this view allows for compensation by a transfer of aggregate wealth across generations. The usually recommended means to such a transfer would be to lower the discount rate uniformly for all investments through the tax structure.[2] This

* Forthcoming in Energy and the Future ed. by Douglas MacLean and Peter G. Brown, Totowa, N.J., Rowman-Littlefield.



appears to be consensus view among neoclassical economists. It is a global approach to handling long-term energy costs.[3]

As a special case, it is sometimes recommended as a matter of intergenerational justice that the intertemporal weights be all set equal — by setting the discount rate equal to zero. This followed in practical decision-making, is a special case of discounting generally because it still treats present and future costs as commensurable — to be combined by a simple (weighted) average. It is a global approach in the other sense as well, as it treats energy and other costs on the same footing.

In the second view, potentially large and very long-term costs of energy alternatives should be treated specially, partly because they are large and long-term and partly because they have to do with the management of the resource base. In this view, the resource base should be preserved "essentially intact," as a matter of justice between generations.

In the first section of the paper, I contrast the second view with the global one. In the second section, I give conditions that define "neoclassical economic utilitarianism," ("neoclassical utilitarianism," for short) which I take to be the philosophical perspective of neoclassical economics. Briefly, a neoclassical utilitarian does not attempt to maximize the sum of utilities over all people, as does his classical utilitarian forebear. Instead, for a neoclassical utilitarian each person maximizes his own utility separately. In the third section, I suggest that a global discounting approach fits naturally but not inevitably with the defining principle

of utilitarianism (both classical and neoclassical). There appears to be room for the alternative view within the neoclassical system, but the fit is partial. And in the fourth section, I consider what happens when the principles defining the neoclassical system are no longer viewed as obtaining. By modifying the principles, we move outside neoclassical utilitarianism, at least as I have defined it. The second view of intergenerational justice seems more consistent with these modified principles. Thus it really belongs outside neoclassical utilitarianism. But as I find the modified principles more "realistic" than the original conditions, I find this an argument in favor of the second view of intergenerational justice. In the fifth section, I consider the principal objection to this alternative, that it is likely to lead to intergenerational inefficiency. Both for empirical and conceptual reasons it appears that the notion of efficiency applies somewhat differently, and with less normative appeal, when there are long time periods and potentially grave harms involved.

Two preliminaries are in order. First, it should be made clear that one does not need to abandon a notion of discounting if one accepts the second view. If one accepts the second view, there is a role for discounting in terms of the opportunity cost of capital for specifying what it means to keep the resource base "essentially intact." And there remains the traditional role for discounting once the prerequisites or constraints of intergenerational justice are met. While the second view does not "abandon" discounting or advocate zero discounting, there are fundamental differences between the two views, which the paper tries to clarify. Second, in attempting to elucidate



the two views, I have drawn what may appear an unflattering portrait of a neoclassical utilitarian and his theory of mind. That is not the intention. As sketched, the theory is simple but simple theories have hard cutting edges. The enormous power and utility of traditional economic analysis are not in question. What is in question is the appropriateness of the theory and the perspective for the problem of very large and very long term costs, where issues of intergenerational justice are fundamental.

## I. Global Versus Special

The most appealing argument for discounting long-term energy costs at the same rate as other costs is only peripherally related to a concept of intergenerational justice. Instead it is based on intergenerational Pareto optimality.[4] The argument goes as follows. In markets, costs and benefits are discounted at a rate equal to the opportunity cost of capital, i.e., the value of alternative uses of invested capital. If energy costs are handled specially — discounted at a lower rate, discounted at a zero rate, or handled in some other way altogether — then the resulting plan and resource use will be intergenerationally inefficient. It will be possible to modify the plan so that some or all generations are made better off without hurting others. And the modification will be made by discounting energy costs "just like" other costs and benefits. Since a Pareto improvement is generally considered good, discounting, as the condition for it, must also be considered good.

An example illustrates the point. Suppose we are to choose among several possible uses of $100,000 worth of resources. The choices are: (A) Use the $100,000 for a short-term safety program with the expectation of saving two lives this year; (B) Invest the $100,000 at the prevailing market rate of interest, let us say 7 percent; (C) Spend $100,000 on incrementally improving the safety features of a nuclear plant with the expectation of saving four lives thirty years from now; and (D) A combination of the above. Suppose further that in thirty years it will still be possible to institute a short-term safety program for $100,000 worth of real resources with the expectation of



saving two lives in the thirtieth year. If we are to treat energy costs specially, particularly ones having to do with costs to life, we might not want to discount the four lives of (C). Treating the date of a life saved as "morally irrelevant," the saving of four lives thirty years from now appears an even better bargain than saving two lives now, and (C) seems better than (A). Because we are treating lives and energy costs specially we might not even consider (B), and make our choice for (C).

But according to the argument for discounting this would be a mistake, and it can be seen to be a mistake by a simple discount calculation. Discounting four lives thirty years from now at 7 percent per year, we have the equivalent of only 0.5 expected lives, at present. Thus (A), which saves two lives now, beats (C) and there must be an intergenerational inefficiency. And so we must be able to find a way of making all generations better off, compared with plan (C). One such way would be to follow (D) by devoting half of the $100,000 to the short-term lifesaving program, with the expectation of saving one life in the present generation (follow (A) with half of the resources). The other half is invested in (B) where it grows to $381,000 in thirty years. The proceeds then are channeled into a short-term lifesaving program with the expectation of saving slightly more than seven lives in that generation. All generations are made better off, compared with (C): the first generation by saving an extra life, the second by saving an extra three lives.

There are some problems with this argument in its trading off lives from one program to another and in its application to the very

long term, but I want to postpone a discussion of these problems until the end, and follow another thread of the argument for now. It is well known that efficient allocations may not be fair or just ones. The present generation might be very well off and future generations starving and cancer-ridden, from radioactive materials and other toxic chemicals. And yet the situation could still be intergenerationally efficient in the sense that the future could not be made better off without making the present worse off. Justice is not efficiency: it may be possible to move from one efficient but unfair allocation to another efficient and fairer allocation.

At this point I will simply stipulate that intergenerational efficiency is a desirable property to be obtained when it can be achieved. Instead I want to focus on the across-the-board, aggregative aspect of the discounting approach. The approach is global in the sense that everything is subject to substitution and trade-off.

This approach can be contrasted with the more specialized or piecemeal approach in a second illustration. Suppose that you are to occupy a friend's house for a month, while the friend is away on vacation. In the course of your stay you make some phone calls, eat some of the staples, and perhaps drink some of the beer in the refrigerator. As your stay draws to a close you restock the refrigerator, replenish the staples, mow the lawn, and generally arrange to put the house back into the condition in which you found it. This is a piecemeal approach. You are not primarily concerned with maximizing the sum of your and your friend's utility nor are you concerned with an efficient allocation between you and your friend.



You are interested in putting things, <u>particular things</u>, back where you found them. Of course the rice you buy is not the identical rice that you ate, it may not even be rice if you can't find it. But it will be a close physical substitute. (How close will depend on how particular you are and how particular you think your friend is.) You might go further and leave some flowers or a house present, but that would be a gift, not a requirement. What is required, on this view, is to leave intact physically what is not yours to run down.

It is also possible to take a more global view. Why should I mow the lawn, you might say. Perhaps my friend will want to reseed the lawn. Why should I do the laundry? Perhaps my friend will want to buy some new sheets. Instead I will leave a generalized transfer. I will leave some money, enough to compensate him if he wants to mow the lawn and do the laundry himself, or hire someone, or help finance some other choice if he wants that. In this way the range of choice is increased and efficiency improved.

Putting the house back into its original conditions is like keeping the resource base intact intergenerationally. Taking the more global perspective and letting the house run down but trading off with generalized compensation is like the discounting approach, but somewhat abstracted from it because I have abstracted from time, productivity, and even discounting itself. Moreover, I have glossed over a couple of problems to be attended to later. I would like to argue that the specialized approach is more appropriate than the global one, as a matter of intergenerational justice, but I do not want to do so within the framework of utilitarianism underlying modern neoclassical

economics. Part of the reason is that in this framework, at least as I will define it below, there is virtually no room for a concept of justice. And part of the reason is that when we depart from the neoclassical system by modifying its defining principles, the specialized view of resource preservation as a notion of intergenerational justice appears to fit rather nicely with the modified principles. Thus the next step is for me to define the central features of the neoclassical system.



## II. Defining Neoclassical Utilitarianism

If you ask an economist what it means to advocate utilitarianism, he might well reply that it means to prescribe behavior that maximizes whatever it is that one values. This definition, of course, it too vague and inclusive. To some extent we are all self-serving and we all practice maximizing behavior. In a sense, even inanimate objects practice maximizing (or minimizing) behavior. Soap films minimize surface area; light bends to minimize travel time through different media; water flows in a way to minimize potential energy. If some form of maximizing behavior is the defining characteristic of utilitarianism, then we are all utilitarians by definition.

But the important thing for utilitarianism is not that we maximize, it is the pervasive nature and the particular conception of utilitarian maximizing behavior. Thus it is useful to define utilitarianism by the principles that make the maximization process universal.

Maximization is difficult when there are many distinctions drawn. It is generally impossible to maximize two separate things at the same time. The principles that I set out below serve to clear away or collapse distinctions that might otherwise interfere with the maximization process. One way of stating the matter is to say that these principles are conflations.[5] My approach is to take conflations as the defining characteristics of utilitarianism. Clearly there are advantages to abstracting from detail and blurring distinctions. In the case of mathematical argument, abstracting and generalizing often

lead to deeper and more powerful insights into structures of ideas. But there can be disadvantages too. When the distinctions are important, glossing them over can lead to a structure of ideas quite different from the world that we are trying to say something about.

## Principles of Neoclassical Utilitarianism

Some of the utilitarian principles can be seen as defining a theory of mind, and a very simple one at that. The neoclassical utilitarian is drawn to this theory for two important reasons. The first is to portray all a person's values as generally comparable, so that some general maximization can make sense. Second, it to base the theory on observable behavior.

The first aspect of this theory of mind can be expressed as a view of how values are formed. What might otherwise be considered heterogeneous things (for example, decision processes, descriptions of conceivable states of affairs, the present and the future) are treated as unified, homogeneous objects through conflating many distinctions. I will call attention to three of these conflations.

1. Only preferences matter. There seem to be several ways of making decisions. For some questions the process of decision might follow moral reasoning, for others it follows maximization calculus. Some decisions might be made on the basis of religious concerns, some on the basis of habit, some on the basis of some automatic code of behavior. Or we might posit two entire preference structures: one appropriate for normal decisions, the other appropriate for moral choices, as was done by Plott.[6] If we make such distinctions there are



obvious problems. We have to explain when each one applies. We need, at least in principle, some way of defining the boundaries among the processes, and we have to be able to say how one process shifts into another. But if we conflate all decision processes into a single process, that of ordering preferences, such difficulties are cleared away. A preference conflation implies that all these processes are fundamentally the same process and can be modeled as though they were just one process, that of preference ordering.[7] For the classical utilitarian this first conflation can be stated by saying that each person has just one utility function; for the neoclassical utilitarian, just one preference ordering.

2. All states are comparable. The idea of a "state" is a very general concept. It is a complete description of reality. This description can include such morally laden possibilities as "John was murdered." The second conflation says that any complete description of reality is directly comparable with any other in the sense that each individual is assumed able to judge whether he prefers the first to the second, the second to the first, or is indifferent between the two. The classical utilitarian states the second conflation by saying that the domain of the utility function is all conceivable states. The neoclassical utilitarian would say that preference orderings is complete over all conceivable states for the world.

3. Future and present states are directly comparable. Now we turn to a matter that was left open by the second principle. This is the question of whether a complete description of the world — a state — includes a description of the future as well as the present. The

third conflation says that a state is a description not only of a conceivable present but also of a conceivable entire future. We can think of a state not just as a snapshot of the present moment, but as a whole movie film of an entire possible present and future, where the first frame is a complete description of the present and each successive frame is a complete description of a possible future day, or generation. Thus the utilitarian chooses among whole possible movie films, not just single frames of the movie, snapshots of the present moment.[8]

A logical implication of this conflation is that if we are really choosing among whole movie films there is really just one choice, now, for all time. Many economic models are of this form. Dynamic programming models, or control theory models, collapse the future and the present together into a single shot choice. This point is clear when we realize that for control theory problems there is a single-valued functional, ranging over all time, being maximized just once, from the vantage point of the present moment. Thus a value is being put upon the entire movie film and not just a single frame.

The collapse of time is held across the life of the individual as well as across generations. It is as though one could put a thermometer into a person's mouth and out would pop rankings of entire life plans. These plans would of course be done from the vantage point of the present moment. In many models, the only problem is to find the top element of the ranking — the problem is one of maximization. It seems apparent, however, that 10 or 15 years from now a person will have interests differing from those of today (perhaps because he has



grown and changed or perhaps merely because of the shift in vantage point in time). If we accept the third conflation in full force, we concern ourselves only with what the thermometer says today.[9] Backing off from this full collapse of time, we might be concerned with the different readings of the thermometer in different vantage points in time. Then we would have the problem of "justly" taking into account potentially conflicting interests. In the language of social choice, there is an "aggregation" problem (the problem of resolving conflicting interests), in addition to the maximization problem (the problem of picking off the top element once the aggregation problem is resolved). The framework of intertemporal social choice, which later will be used to discuss the problem of justice across generations, can also be applied to the problem of justice between the earlier and later selves of a particular individual.

The third conflation says that we concern ourselves only with what the thermometer says today. (Today's reading takes into account the contemplation of future utilities and preferences, but only insofar as the contemplation adds to present utility.) For a classical utilitarian, time-dated states are legitimate arguments for each individual's utility function. For the neoclassical utilitarian, preference orderings are complete over time-dated states.

The second feature of what I am calling the neoclassical utilitarian theory of mind, the inclination to behaviorism, is expressed in our fourth utilitarian principle.

4. Utilities of different individuals are not directly comparable. This principle can be stated as a conflation that divides

classical and neoclassical utilitarianism. The classical utilitarian accepts comparability, but the neoclassical utilitarian does not (except sometimes in the intertemporal case). For the classical utilitarian, utility is a measurable quantity, at least in principle, and a quantity of utility for one individual can be added to a quantity of utility from another individual. For a neoclassical utilitarian, utility is not a "real" quantity — it cannot be measured even in principle — and thus there is no way to add one person's utility to another's.

Interestingly, however, in the intergenerational case, neoclassical utilitarians often cross the line and act like classical utilitarians. While in economic analysis there is great reluctance to add the utility of one person to that of another, within a generation, utilities of different people are commonly added across time. One story that formally avoids adding utilities across time is to assume that each person lives forever. Another story that permits adding utilities across generations is to assume that our heirs are "just like" ourselves. They are extensions of ourselves, share the same interests, and one utility function fits all (per extended family), Neither story is very satisfying.

A colorful way of stating how the neoclassical utilitarian came to reject the fifth conflation is to say that it was killed by the possibility of a utility monster. The utility monster is a sensitive fellow who can squeeze more utility out of a given resource than an ordinary person can. If we are classical utilitarians who attempt to maximize the sum of utilities across people, then we should give a



larger share of a resource to someone who has greater capacity to squeeze more utility out of it. But many economists are unwilling to give the larger shares of the economic pie to those most efficient in converting utility. One might argue to the contrary, that the less efficient converters should be given larger shares in compensation.

Added to this is the incentives problem, which is of great concern to modern economists. Even if utility "existed" it would be non-observable, and if we were to maximize the sum of utilities there would be obvious incentives for each one of use to claim that he is a utility monster (the youngest child of each family is sometimes tempted toward such claims.)

To state the matter a little more soberly, many economists rejected classical utilitarianism in favor of its neoclassical version when they decided that utility was entirely non-observable. At the same time it became clear that most of the structure in economics could be preserved by thinking in terms of preference orderings as opposed to quantitative utilities. Preference orderings have the advantage of being, at least in principle, observable by choices actually made. This rejection of classical, quantitative utility has two repercussions noteworthy for our purposes.

First, if interpersonal comparisons of utility are impossible, then we are no longer able to maximize the sum of utilities across people. So the neoclassical utilitarian defends a weaker kind of maximization process in which each one maximizes his own utility. The classical utilitarian's moral principle, which says to maximize the sum of utilities, is strong in the sense that it sometimes directs people

to act against their own selfish interests. The corresponding weaker neoclassical utilitarian's moral principle says that we should move toward Pareto optimality. This principle is weaker in the sense that it does not require individuals to act against their own selfish interests. It is also weaker because in many situations it does not tell us what to do (it is a partial ordering).

Second, the rejection of unobservable utilities leads toward a behaviorist or black-box theory of the mind. The only evidence allowed for interferences about happiness or satisfaction is taken to be observable behavior: for example, actual purchases in markets. Thus, evidence from introspection is looked upon with suspicion, as are surveys of stated preferences. The situation is a little like trying to infer the structure of a car's motor by observing the car's behavior. With this black-box approach it is not surprising that we might be limited to simple concepts of the motor.

The theory of mind for a neoclassical utilitarian may be a very primitive theory, but the overall view is consistent and coherent so there is little chance of refuting it internally. In that sense it is a comprehensive theory and can explain almost anything.[9a] It is similar in this respect to other comprehensive theories of mind and human behavior. For example, another alternative explains every human action by "God willed it." The theories are different but each difficult to refute. The mere fact that things are explainable within a system is not an argument for that system. Another way of choosing among systems involves the appeal to some outside criteria, for



example, Occam's razor, predictive success, introspection, or ethical considerations.

The view of the mind that is presupposed by an economic theory will of course be very important in assessing the moral implications of using that theory for making choices. There is, in addition, one further explicitly moral principle that is central to neoclassical utilitarianism. It is a thesis about rights.

5. Property rights must be well-specified. This principle tends also to be a conflation because, in its extreme version, it can be taken to mean that the only important rights for the neoclassical utilitarian are property rights, and that anything that can be valued should be privately owned.[10]

The motivation here is that free transactions through the market are the best way of revealing preference orderings and also of arriving at allocations of goods and services — states of affairs — that are Pareto optimal. A primary way to low transaction costs and few conflicts among various individual rights is to completely specify ownership rights and make them tradable.

## III. Discounting Within the Utilitarian System (Classical and Neoclassical)

The purpose of this section is to suggest that discounting fits easily but not inevitably with principles 1-5 or a subset of 1-5. Indeed, one is struck by the number and variety of the arguments that take these principles as background and lead to discounting. The impression is that all roads lead to Rome. I will mention four of these arguments.

All four approaches incorporate principles 1-3. Two of the roads are in the classical utilitarian tradition and incorporate principle 4; two are in the neoclassical tradition and reject 4.

We will find that all four approaches are "institution free." This means that they define criteria, but not constitutions for achieving the criteria. Conflation 5 concerns property rights, an institutional structure. Thus the four approaches below are compatible with but do not directly incorporate the notion that all rights should be construed as property rights.

For purposes of the paper as a whole, it is more important to show that discounting does not inevitably fit than to show it fits easily into the landscape of principles 1-5. To show that some roads do not lead to Rome, all I need is a counterexample. The counterexample provided below shows that the choice of intergenerational decision rules, even within the confines of principles 1-5, is much broader than that of choosing between discounting at a positive rate or discounting at a zero rate (adding up utilities or some other measure across time).



Nevertheless, 1-5 do shape perspective and they do not appear to be a compatible background for a specialized conception of justice as opportunity. In the following section, I attempt to modify 1-5 to develop a background more conformable to this latter notion.

### Four Roads That Lead to Rome

With this itinerary in mind I begin now with four of the arguments leading to discounting. They all involve a "planner" who trades off present and future generations' utilities somewhat as he would trade off his own present and future utilities in the third conflation. But here the planner is assumed to be, in some sense, intertemporally neutral, or sympathetic with the interests of all generations.[11]

A. The planner who maximizes the sum of present and future utilities. This is a planner who is not selfish because he weights other generations' utilities as heavily as his own generation's. To do this the planner must accept the 4th conflation (he is a classical utilitarian). But the planner sees no sense in allocating equal weight to a distant generation if it may not exist. So the planner discounts each generation's utility by the probability that it will not exist. To arrive at a constant discount rate (Rome) it need only be further assumed that the probability of extinction during the course of one year, given that extinction has not already taken place, equals the probability of extinction during the course of another year, given that extinction has not already taken place before that other year.[12]

From the point of view of this discussion, the most fundamental normative problem is that this approach treats the probability of extinction as a fixed parameter outside the system, unaffected by this generation's actions. But the probability of the next generation's survival is strongly affected by the present generation's actions, and a sufficient guarantee of an "adequate" level of survival is the central question of intergenerational justice.

B. The selfish planner whose self-serving tendencies are blocked by a veil of ignorance. This time we need not subscribe to principle 4. The planner is only looking out for his own welfare and is not comparing it with others (he is a classical utilitarian). But even though the planner only wants to maximize the utility of his own generation, he does not know to which generation he belongs. Thus he maximizes the expected value of his own utility, weighting each generation's utility by the probability that he attaches to being in that generation. (For simplicity we imagine that each generation contains only one person, which may or may not be the planner.) The planner's utility function takes into account his own (selfish) risk aversion to being caught short in a particularly barren generation. As in the previous argument, the planner posits the increasing likelihood of eventual extinction and we end up again with utility discounting, but a different route and a different interpretation of the utility function. For this Rawlsian gambler, or more accurately Harsanyian gambler as interpreted by Dasgupta and Heal, the key assumption leading to discounting is again the declining certainty of future existence.



C. The egalitarian planner who is worried about productivity. In simple models which allow for capital productivity, if we are adding up utilities across time (in the classical tradition) to achieve equal utilities across time, then we need to discount by the marginal productivity of capital. In this type of model, if we simply maximized the sum of utilities, discounting at a zero rate, early generations would sacrifice to invest more so that later generations could feast off time-delayed yields of capital. To achieve an egalitarian sharing across time, the productivity of capital needs to be offset by discounting future utilities. In more complicated models egalitarianism is not achieved so simply, but the flavor of egalitarianism remains in allowing discounting to offset productivity.[13]

D. The planner who is fair because his preferences are generated by fair axioms. This is a more complicated path that is based on some important work that attempts to find a social choice rule for aggregating individual utility orderings. If we allow, as a simplifying assumption, that each generation's utility can be treated as a separate preference ordering and thus combined by an aggregative social choice rule, then the fairness of this rule would seem to be deducible from the fairness of each of the axioms that describe it. This is consistent with a neoclassical utilitarian perspective.

In a pioneering set of papers, Tjalling Koopmans proves a theorem which can be reinterpreted in an intergenerational context.[13a] The theorem depends on a set of axioms, each of which appears neutral, innocuous, and fair. In its reinterpretation the theorem proves that

an intergenerational planner who adopted these axioms must be led to a social choice rule that discounts the utilities of future generations. The proof is mathematically complicated, and we will not try to reproduce it here.[14]

It is possible, however, to choose a set of axioms also appearing neutral, innocuous, and fair, that lead to a different social choice rule. Kenneth Arrow's well-known axioms, applied to the intergenerational context, generate a social choice rule that strongly favors the future over the present. Arrow's collection of axioms is in some ways similar to majority rule voting, and the infinite majority of future generations dominates the minority of the present. Yet Koopmans's axioms, which are also applied to an infinity of generations, yield a quite different time bias. Most interestingly, if we take the crucial axiom from Koopmans' set, the axiom of stationarity, and combine it with Arrow's three axioms, we get a still different result: "dictatorship of the present."[15] This term has a technical meaning in social choice theory. It means that whatever the first generation prefers is the intergenerational social choice.

Thus, an axiom that seems plausible can be combined with other plausible axioms to yield controversial results about discounting, and even to yield different results under different combinations. Time bias is not apparent from looking at axioms singly, but depends on sensitive interaction among the axioms taken together.

A further observation is that the Koopmans axioms and the Arrow axioms illustrate the weakness of efficiency (Pareto optimality) as an ethical principle. Koopmans's axioms lead to discounting for a social



choice rule; Arrow's do not. But both satisfy Pareto optimality. It is an axiom in both systems. Thus, the Pareto principle cannot be used to choose between them.

Finally, this social choice rule framework illustrates that there are three possible levels at which discounting can take place. Each generation, individually, may discount to reflect its own time preference, because each generation determines its own preference ordering over the entire time path. This is not true for both Koopmans's and Arrow's axioms. Second, discounting can also show up, in both systems, in the definition of the feasible states, by taking account of capital productivity in determining what is feasible. Thus, the opportunity costs of capital can be embedded in the definition of feasibility. The difference between the Koopmans and Arrow axioms appears at the third level. For the Koopmans axioms discounting is also the form of the aggregation rule. For Arrow's axioms it is not.

We conclude, then, that within the framework of neoclassical utilitarianism, many paths lead to a discounting formulation. They exhibit a rich variety of assumption and interpretation. But there are also paths that do not lead to discounting, <u>at the level of intergenerational choice</u>. This third level is the level of our main concern. Arrow's axioms, reinterpreted intergenerationally, show that it is possible to not have Pareto optimality without discounting at the level of social choice, yet with discounting for personal time preference and the opportunity cost of capital at the other two levels.

## IV. Outside the Neoclassical System

In the preceding section we looked inside the neoclassical system. We found room for aggregation of intergenerational interests without discounting them. However, in the neoclassical system any notion of justice would have to be built on the utilitarian principles 1-3. Within this theory "preferences are all." They soak up and explain all forms of choice and behavior at the individual level. It may be possible to develop within this system a satisfactory notion of a fair or just aggregation of intergenerational preferences. Indeed, we have shown there are alternative conceptions of intergenerational fairness inside the neoclassical system. But the utilitarian principles 1-3 are confining, as are 4-5. In this section we enlarge the inquiry.

I want to develop a conception of justice that is based on opportunity rather than utility. To do this, I must move outside the neoclassical system by modifying its defining principles. Why opportunity rather than utility, why move outside the neoclassical system? The motivation is as follows. Inside the neoclassical system there appears to be little room for a concept of justice at the individual level. Outside the system there is no unified concept of utility (or preferences). A simple solution is to move outside the system and base a notion of justice on something other than utility (or preferences). Brian Barry suggests in his chapter that opportunity is a more sensible base than utility (and I have put forward a similar suggestion).[16]



There is another reason for moving outside the neoclassical system: it may be "unrealistic," too simple to describe adequately how our values determine which choices are made and which actions undertaken, and too simple to incorporate our considered judgments about rights and property. If the world is "really" more complicated, then to capture the most important complications it becomes necessary to draw some distinctions.

So now we move outside the utilitarian system (both classical and neoclassical) and attempt to draw some distinctions that might be considered realistic and important. These distinctions lead toward a commonsense notion of justice generally, intragenerationally as well as intergenerationally. But in the intergenerational context they appear to pick out the resource base as a special concern of justice. The conception of intergenerational justice constructed below is not "inevitable." Other conceptions are possible. But the idea is to base this intergenerational concept of justice on its relationship to a commonsense notion of justice intragenerationally.

The way to proceed is as follows. I make some "relevant" distinctions in the neoclassical principles 1 through 5, as they lead toward a commonsense concept of justice intragenerationally. Then I apply these distinctions to the intergenerational case. The appeal for this concept of justice is then grounded in the independent reasonableness of the distinctions.

### A. All ownership rights are not on a par.

I adopt here the Lockean notion of "just acquisition." The most absolute claim of just acquisition by an individual is the claim to

one's work when it is wholly created by oneself. Thus, Byron had a right to burn his books, but his wife did not, without his permission. (The classical utilitarian would not see the point of this distinction and might indeed deny Byron himself the right to burn his books.) The next strongest claim of just acquisition is by an individual who "produces" an object by mixing his labor with a resource of which there is "enough and as good" left for others. The least claim, in fact no claim at all, of just acquisition concerns the resource base as a whole from the point of view of the present generation. For the resource base passes into the hands of the present generation by the mere passage of time alone, willy-nilly, without any effort by the present generation. Shakespeare's plays are a part of this resource base. They were not produced by the present generation, hence this generation does not have the right of ownership over them in a sense that would justify doing what it wants to them, including destroying all records of them.

By this distinction, ownership is a relative not an absolute concept and is based on a relative notion of just acquisition. This notion of acquisition sharply distinguishes the resource base, including the cultural and technological heritage of past generations, from the capital stock produced by this generation. The distinction between present capital and the resource base is not admitted in the neoclassical utilitarian system. In fact, within the neoclassical utilitarian system, it is likely to be argued that such a distinction will lead to large intertemporal inefficiencies. I will discuss this in section V.



The distinction between just ad unjust acquisition is constantly made in everyday life. Even if you innocently buy stolen goods, your ownership is not secure because the goods were not justly acquired by the thief and thus not the thief's to sell or yours to own. In the United States we allow rather absolute ownership rights over particular natural resources. One interpretation of this approach is that it has historically been believed in this country that resources are sufficiently extensive that there are "enough and as good" remaining for others, including later generations. In other countries, where resources have been more obviously limited, ownership of natural resources is more circumscribed. With the growing concern that the Lockean proviso is not satisfied, there is greater concern about how absolute the ownership of natural resources should be. For example, the trend toward increasing severance taxes can be viewed as a limitation on the absolute ownership of materials extracted from the environment.

The distinction — between what is acquired more through our own efforts and what less — leads to a commonsense notion of intergenerational justice. By this notion, it would be unjust to future generations if we were to run down the resource base when we have the opportunity to treat it on a sustainable basis: since the resource base was not justly acquired by us, it would be unjust to run it down. By the same token, it would be unjust to run down the previous generation's capital and cultural accumulation. It might be ungenerous if the present generation chose to add nothing to the future heritage, but it would not be considered unjust (like the house

sitter's non-obligation to provide flowers). Thus this commonsense notion of justice is a kind of minimum of moral responsibility.

Clearly, the idea of preserving what is not justly acquired needs to be made more practical. The present generation cannot be required to preserve every obscure, minor literary work that all previous generations might have produced. Nor can it be required to preserve every tree, or oil and coal deposit. In the example of the house guest in the first section, the replacement of the basic stocks was not precise. Little things need not be restored to their original position, only the more essential. Thus, to make this notion of justice practical, we need some notion about what is more and what is less essential.

B. Not all states or goods are comparable.

The second distinction is to say that some things are more essential than others. Obvious candidates for essentiality are basic health and liberty. Essential goods appear to correspond with Rawls's primary good. the idea of essentiality also appears in Adam Smith's diamond and water paradox. Smith thought it a paradox that though water was much more valuable (essential) than diamonds, diamonds had a much higher price per unit than water. A way of distinguishing the essential from the non-essential is to note that we might consider trading essential goods near the margin, but not far inside the margin. For example, suppose that you are wrongly convicted and imprisoned. After a day in jail, the authorities realize their mistake and set you free. Unlike the current system, the authorities attempt to make complete restitution for you. They ask you how much money you would



need to have in order to make you feel completely indifferent between the day's imprisonment and the compensation and neither the imprisonment nor the compensation. Even though you are trading off liberty in this case, you might be able to name a figure that would indeed make you indifferent between having the mistake and the compensation and neither. (Whether or not you would honestly reveal this figure is a separate question.)

However, suppose that you are falsely convicted and imprisoned for twenty years before the authorities realize their mistake. At that time, the authorities again pose the same question to you. What compensation would you require to make you feel indifferent between the twenty-year mistake and the compensation and neither the mistake nor the compensation? In this case, even if you attempt to address this question honestly, you might have no way of dealing with it. You may find you have no basis for the comparison, no way of naming such a figure, even a very high figure. Similarly, one can ask you what is the premium wage payment you are willing to accept in order to live with a slightly higher risk of cancer. You might be quite willing to make this trade-off on the margin. But suppose you, in fact, got cancer. Is there then some compensation that could make you feel indifferent to it? Or suppose you are asked to work in a hazardous occupation with an 80 percent probability of cancer.

Turning to the intergenerational case, it is plausible to argue that the resource base as a whole is more essential than this generation's capital stock accumulation. For example, Japan and Germany got along with more than one generation's worth of capital

stock destroyed and both countries were able to rebuild their capital stock rather quickly. But it is clear that neither country could have survived without their own or imported energy and metals and other materials from the resource base. A particular metal may not be essential but metals as a group are. Similarly, a single source of energy may be inessential but the entire energy sector is essential. A sufficient condition for sustainability, and one that is perhaps unnecessarily strong, is to keep the cost of extraction from the resource base roughly constant, major sector by major sector. This criterion allows substitution within sectors. It allows destruction of some resources. Elsewhere I have discussed the role of severance taxes in creating new technologies and new substitutes for the depleting resources and thus promoting sustainability.[17]

C. Offsetting harms with benefits.

Attitudes toward the distinction between doing good and avoiding harm is a litmus test for utilitarians. Within the utilitarian system we can't distinguish between avoiding harm and doing good; one is the opportunity cost of the other. MacLean offers the example of randomly killing a person in order to extract two kidneys in order to save two lives of kidney sufferers. Within utilitarian systems this might seem like a net gain. But most commonsense notions of justice would consider it unjust to kill the one person in order to save the two.

In the intergenerational context, we are currently harming future generations by damaging the resource base through the dispersion of radioactive wastes and toxic chemicals and physical depletion. At



the same time, we are benefiting the future by increases in the capital stock, technological understanding, and cultural accumulation. However, under this distinction, there is not a simple one-to-one trade-off.

This distinction leads to an asymmetrical treatment of the resource base. How do we draw a distinction between allowable harms and unallowable harms? We can't prevent all harms to the environment. But we can "protect" and "renew" essential goods. We can reduce our releases of radiation toward the background levels of release that would occur through natural erosion; we can stabilize the population to maintain the resource base on a constant per capita basis.

The ethical choice for the present generation is to move in one of two directions: The present can manage the resource base on a sustainable basis or it can let the base slide into an irreversible decline. (If there were no way of preventing the latter option, the choice would lose its moral relevance.) Within the utilitarian system the latter choice is viewed simply as a preference of the present generation. Its consequences would be unfortunate for the future. In the alternative view the latter choice is unjust as well as unfortunate.

D. Opportunity vs. utility.

It seems sensible to focus on and limit our responsibility to what we can foresee and control. As future opportunity is more in our control than future utility, it would seem that the former is a more sensible predictor for a notion of intergenerational justice. With some effort we can control the form of the heritage to be passed on to

the next generation. It is beyond the control of the present generation to ensure that the next one will be happy or hard-working. It is beyond our control to increase their welfare; we can only assure them of certain opportunities for happiness that we can foresee will be essential. But we can preserve certain essentials, such as the valuable parts of the cultural and natural resource base. If we cannot ensure that these will in fact be passed on to the more distantly future generations, we can at least keep from ensuring that they well not be passed on.[17a]



## V. The Inefficiency Issue

From the perspective of neoclassical economics the most obvious objection to a special treatment of the resource base is that attempting to preserve it "essentially" intact could conceivably lead to very large inefficiencies. A great deal of effort and sacrifice could be spent preserving some part of the resource base which no one in the future would want. The objection suggests both empirical and conceptual considerations. First the empirical.

There is of course the possibility that the present will go to great effort to preserve something that the future does not want. But is this probable for the likely candidates for essential goods — conditions of basic health, alternative provision of energy sources, water, soil, space per capita, etc.? In the case of radioactive waste, the notion of intergenerational justice developed here suggests that aggregate exposure be kept near natural background levels. Natural releases are to be diminished to compensate for releases from energy production. Whether or not this standard can be met, and if so at what cost, is an empirical matter. But it is unlikely in the foreseeable future that people will come to be indifferent about cancer. Even if a cancer cure is found, many millions of people with little or no medical care will not benefit from it.

As an empirical matter, it appears that with the present accumulation of man-made capital, dependence on the physical resource base is growing, not shrinking. It is conceivable that we might someday free ourselves from our dependence on (say) metals. In that case metals would become "inessential" and their preservation, in an

opportunity sense, would no longer be considered a matter of intergenerational justice. But year by year annual rates of extraction for aggregate resource groups go up, not down.

Also as an empirical matter, we can ask how much it would cost to satisfy a notion of justice as equal opportunity. Consider a switch from present depletion allowances to severance taxes. Compared with the present tax system such a switch appears to impose few or no aggregate costs upon the present and yet to produce net benefits to the future.[18] In other words, implementation of this notion of justice may even coincide with a step toward intergenerational efficiency.

And finally, as a quite different and more conventional approach toward intergenerational efficiency, we may consider the kind of compensating investments contemplated in section I. Suppose, for example, we calculate that there is a 1 percent chance of large-scale radioactive contamination following uncontrolled nuclear proliferation leading to a 10 percent excess risk of cancer worldwide, 100 years from now. Suppose further that this risk could be eliminated by a present investment in safe-guards of $5 billion. Are we to decide against the safeguard if the expected number of deaths, discounted at the marginal rate of productivity (say 10 percent) is less than the $5 billion? In this case the rationale for the comparison falls apart because the compensating investment is not sustainable for a century or more at a 10 percent marginal rate, when the entire economy is growing at substantially less than that. For such a long period a substantial marginal investment is not a real option, because it would dwarf the



economy, in a 100 year interval (as it would for this example if economy as a whole were growing 3 percent per year).

These empirical inquiries, by no means settled but at least partially identified, are useful, but less fundamental than conceptual considerations. Suppose for example, contrary to empirical likelihood, it were possible to make large compensating investments over a century or more. Obviously, if the compensating investment is not made in the present, the compensation is not an option in the later period, because the investment has to grow in the intervening years to become available in the later years.

The impossibility of later compensation through redistribution stands in stark contrast to the conventional notion of potential Pareto improvement. In the standard example, a dam is constructed that floods the land of some farmers. But so much benefit is created from the dam as a whole that there are enough proceeds for the winners to compensate the losers so that everyone comes out ahead. In the conventional case actual Pareto improvement is possible in the second period. Thus, the compensation choice does not have to be taken in the first period. But in the intergenerational case either the compensating investment has to be made in the first period or it becomes irrelevant from the point of view of the second period because it is not an option by that time.

The standard argument for discounting says that it is all right to harm the future, as long as it might be possible to benefit the future on net balance by a compensating investment, even though the investment is not taken. To abstract from time, this is like saying that it is all right for me to harm you if I had the option, which I

did not take, of aiding you on balance. If the harms are minor, and there are many interactions, and on balance I am aiding you, then we might overlook the non sequitur. But when grave harms are involved the argument has less appeal.

Moreover, when grave harms are involved, we may not be willing to trade off, period. If we reject the first and second conflations, and conclude that not all things are comparable, the demands of efficiency become weaker. For some cases there may be no way of deciding when someone or some generation is better or worse off. Perhaps nothing can be said "on balance," only that in some ways a person or generation is better off, in some ways worse off.

And finally, neoclassical utilitarianism is unable to distinguish or choose between two very different intergenerational rules of choice (e.g., the implications of Arrow and Koopmans axioms) because both satisfy the condition of intergenerational efficiency in principle. Thus, intergenerational efficiency is hardly an adequate or sufficient notion for the long run.

In the intergenerational case, we should start with a notion of a just protection of fundamental opportunities and from this initial starting point encourage steps toward intergenerational efficiency. This would mean establishing institutions that in some sense permit one generation to "communicate" with another. Common law, perhaps, is one such institution.

It is not possible to establish trades among generations in the same way that trades take place intragenerationally, but it is possible to establish institutions whereby one generation anticipates the needs



of another. To the extent that we are successful in establishing such institutions the cost of providing justice as equivalent opportunity will be diminished. There will be greater efficiency in the conventional sense.

## VI. Conclusion

Neoclassical utilitarians make no distinctions between natural resources and man-made capital. These are highly substitutable. The focus is on highly aggregative concepts complete preference orderings for the neoclassical utilitarian; utility for his classical forbear.

In this essay we move outside the utilitarian tradition to make several distinctions which appear to lead toward a commonsense notion of intergenerational justice. These distinctions support a specialized notion of justice focused on the preservation of opportunities arising from the resource base and the past accumulated cultural heritage. Not all opportunities demand preservation, only the most essential opportunities.

I am suggesting that if the present generation provides a resource base "essentially" the same as it inherited (including the same lack of contamination), it has satisfied a notion of intergenerational justice. "Essential," of course, is the key word, and I am construing it perhaps more narrowly than some would. This notion of intergenerational justice appears to be a sufficient one in the sense that if the present generation gives the next an equal chance at what is jointly shared across time, the requirements of intergenerational justice have been fulfilled.



## NOTES

1. All the participants in the Working Group on Energy Policy and Our Obligations to Future Generations contributed my ideas, one way or another, and I would like to thank them all, and especially to thank Douglas MacLean, for his excellent and thoughtful editing. I would also like to thank Will Jones and Ed Green for helpful comments.

2. See, for example, Joseph Stiglitz: "The appropriate instruments to use for obtaining a more equitable distribution of welfare (if one believes that the present distribution is not equitable) are general instruments, for example monetary instruments directed at changing the market rate of interest." "A Neoclassical Analysis of the Economics of Natural Resources," in Scarcity and Growth Reconsidered, edited by V.K. Smith (Baltimore: Johns Hopkins Press, 19790, p. 61.

3. How "the discount rate" and hence all interest rates are to be manipulated is usually left unclear. Presumably adjustments are to be done through the tax structure, or perhaps through monetary policy. There does seem to be a "targets and instruments" problem, because manipulation of interest rates is suggested for several purposes (inflation control, stimulation of certain sectors of the economy, balance of trade, etc.)

4. Some definitions. An intergenerational Pareto improvement is a move in which at least one generation is made better off without making any other generation worse off. A Pareto optimal plan is

one where no Pareto improvements are possible.

"Intergenerationally efficient" is used synonymously with "intergenerationally Pareto optimal," and "inefficient" synonymously with "not Pareto optimal."

5. John Rawls uses this term in passing in A Theory of Justice (Cambridge, Mass.: Harvard University Press, 1971), p. 27.

6. Charles Plott, "Ethics, Social Choice Theory and the Theory of Economic Policy," Journal of Math. Sociology 2 (1972): 181-208.

7. Stephen Marglin discusses the possibility of having two different types of valuation processes, one appropriate for market decisions and the other appropriate for the political arena. While he says he has strong sympathy for the distinction, which he calls the schizophrenic answer, he does not appeal to this approach in his paper. Instead he develops his argument on the basis of a single preference ordering both public and private, intertemporal and intratemporal. Stephen Marglin, "The Social Rate of Discount and the Optimal Rate of Investment," Quarterly Journal of Economics 77 (1963): 95-111.

8. We might imagine that the first frame, which describes the present, would be in much sharper detail than the other frames, which describe the future. But in most models the whole film is in equal color and detail. The film does not represent what we know or forecast about the present and future, in which case later frames would rapidly blur. The film represents a conceivable present and future, and each possible conception of the future can be in as much detail as a possible conception of



the present.

9. Economic models become more difficult when we admit that in the future we can have different interests from our current ones, and that these differences in interests depend upon the shifting vantage point in time. In such models time "really" evolves. In a pioneering paper Strotz analyzes one of these problems, which is known as the problem of intertemporal consistency. Strotz notes that this problem of inconsistency disappears if individuals have utility functions of a discounting form. Strotz believes that not everyone would have such a utility function and there could be an "intertemporal tussle." But Strotz's resolution of the intertemporal tussle is really one of imposition by power as opposed to a solution by "justice." The idea is that if a person's utility function is of a discounting form, he will constrain future opportunities in such a way that later there will be no way to depart from today's plan, to the advantage of the future self. Robert Stroz, "Myopia ad Inconsistency in Dynamic Utility Maximization," Review of Economic Studies 23 (1955-56): 165-180.

9a. It is not a tautology, however. Preference theory usually includes axioms such as transivity and the "weak axiom of revealed preference," and with these or other axioms there is the possibility of counter-evidence and refutability. For an example of counter-evidence see David Grether and Charles Plott, "Economic Theory of Choice and the Preference Reversal

Phenomenon," American Economic Review, Vol. 69, No. 4, Sept. 1979, 623-38.

10. For some further discussion, see Hillel Steiner, "The Rights of Future Generations," chap. 9.

11. Dasgupta and Heal discuss three of the four approaches in P.S. Dasgupta and G.M. Heal, Economic Theory and Exhaustible Resources (Digswell Place, Welwyn: Cambridge University Press, 1979), chap. 9.

12. This is the condition defining a Poison process. See Dasgupta and Heal, Economic Theory and Exhaustible Resources, pp. 260-5, for further discussion; pp. 269-75, for discussion of B; pp. 275-81, for D.

13. This road combines arguments (3), (4), and (5) of Derek Parfit, "Energy Policy and the Further Future, Part I," chap. 2.

13a. This reinterpretation was made, independently, in Ferejobn and Page as in Dasgupta and Heal.

14. The ambitious reader can consult Tjalling Koopmans, "Representation of Preference Orderings Over Time," in Decision and Organizations, edited by C.B. McGuire and R. Radner (Amsterdam: North Holland Publishing Co., 1972).

15. Alas, this proof is also technical. It can be found in Ferejohn and Page.

16. Brian Barry, "Intergenerational Justice in Energy Policy," chap. 1 and Talbot Page, Conservation and Economic Efficiency (Baltimore: Johns Hopkins Press, 1977).



17. Page, Conservation and Economic Efficiency, especially chapter 8.

17a. The same idea of foreseeability can be applied to the means of ensuring opportunity. To some extent, we can foresee ways that new capital and technology can substitute for inherited resources. But many conceivable substitutions (for example, the potential substitution of fusion energy for fossil fuel) are highly speculative. Similarly we might speculate that there will be a cure for cancer and thus it is not so important to contain radioactive materials. But it is much easier to foresee that cancers from radiation will be prevented if the radioactive materials are contained. Thus it would seem that our responsibilities are more direct in preventing radiation releases than in working for a cure, because of the greater uncertainties of the speculative cures for genotoxins.

18. See Conservation and Economic Efficiency, chap. 6 for further discussion of inefficiencies in the present taxation of the resource sector.